

# Photometric wine color measurement: new applications based on machine learning for wine quality management

Vom Fachbereich Chemie der Rheinlandpfälzischen Technischen Universität  
Kaiserslautern-Landau zur Verleihung des akademischen Grades „Doktor der  
Naturwissenschaften“ genehmigte Dissertation

DE-386

**R**

**TU**

Rheinland-Pfälzische  
Technische Universität

**P**

Kaiserslautern  
Landau

vorgelegt von

M. Sc. Marcel Hensel  
geboren in Gelnhausen

Betreuer: Prof. Dr. Dominik Durner  
Kaiserslautern, den 26.08.2024

---

## Eidesstattliche Erklärung

Hiermit versichere ich, dass ich die vorliegende Arbeit eigenständig verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel verwendet, sowie Literaturzitate kenntlich gemacht habe. Die Arbeit liegt weder in gleicher noch in ähnlicher Form in einem anderen Prüfungsverfahren vor.

---

Ort, Datum

---

Marcel Hensel

---

Wann ess Juhr seim Enn Zougitt,  
kahn Schdomp mieh uff de Felder schditt,  
eann de Loft die Roawe fläije,  
Keann als geann de Houste kräije,  
wann die Weiwer beim Erwache  
schnäll eam Owe Feuer mache,  
unn die Menner inner fluche  
ihrn ahle woame Wamst sich suche  
unn ess Wärrer iwerzwerch,  
dann ess Mährt enn Otteberch.

-Oberhessisches Mundartgedicht aus meiner Heimat-

---

## Lebenslauf

Name: Marcel Hensel

Geburtsort: Gelnhausen (Hessen)

## Berufliche Laufbahn

- 12/2020 - heute: Wissenschaftlicher Mitarbeiter, DLR Rheinpfalz, Neustadt a.d. Weinstraße  
Projekte:  
"Entwicklung eines Smartphone-Analysensystems zur Prozesskontrolle in der Weinherstellung und Biotechnologie"  
"SmartGrape: Anwendung KI basierter Messsysteme zur Charakterisierung der Rohstoffe im Weinbau"
- 04/2023 - heute: Informatikdozent am Weincampus Neustadt a.d.W.
- 02/2024 - heute: Freiberuflicher Programmierer für AI-Chatbot Datenannotation, Data Annotation tech, New York, USA

## Ausbildung

- 12/2020 - heute: Promotion: "Photometric wine color measurement: new applications based on machine learning for wine quality management", Rheinland-pfälzische Technische Universität Kaiserslautern-Landau/Institut für Weinbau und Önologie, DLR Rheinpfalz, Neustadt a.d.W.
- 06/2018 - 12/2020: Studium Lebensmittelchemie, Technische Universität Kaiserslautern Abschluss: Master of Science
- 10/2015 - 06/2018 Studium Lebensmittelchemie, Technische Universität Kaiserslautern Abschluss: Bachelor of Science



---

Die vorliegende Arbeit wurde im Zeitraum 12/2020 – 08/2024 in der Arbeitsgruppe von Herrn Prof. Dr. Dominik Durner am DLR Rheinpfalz in Neustadt a.d.W. im Institut für Weinbau und Oenologie durchgeführt und von Herrn Prof. Dr. Jörg Fahrer im Fachbereich Chemie der RPTU betreut.

### **Prüfungskommission**

Vorsitzender: Prof. Dr. Wolfgang Kleist

1. Berichterstatter: Prof. Dr. Jörg Fahrer

2. Berichterstatter: Prof. Dr. Dominik Durner

Prüfer: Prof. Dr. Dr. Gereon Niedner-Schatteburg

Dekan: Prof. Dr. Antonio Pierik

---

## Danksagung

Als Erstes möchte ich Prof. Dr. Dominik Durner danken für die Bereitstellung eines sehr interessanten und interdisziplinären Themas. Die Forschung bezüglich des Themas weckte in mir eine neue Leidenschaft – das Programmieren. Vielen Dank für die zahlreichen Diskussionen, fachlichen und persönlichen Ratschläge.

Weiterhin möchte ich Prof. Dr. Jörg Fahrer danken für die Übernahme der Betreuung sowie die Erstellung des Erstgutachtens. Der Arbeitskreis Fahrer hat mich bei meinen Besuchen an der RPTU Kaiserslautern-Landau immer mit Freude aufgenommen. Vielen Dank dafür auch an Max Carlsson und Simon Wittmann. Auch vielen Dank an Prof. Dr. Dr. Gereon Niedner-Schatteburg für die Ergänzung der Prüfungskommission, an Prof. Dr. Wolfgang Kleist für die Übernahme des Vorsitzes der Prüfungskommission und an Dr. Tina Kostka für die Protokollführung.

Dr. Sarah Di Nonno und Prof. Dr. Roland Ulber möchte ich für die herausragende Zusammenarbeit im Projekt "Smartphone Photometrie" sowie für die zahlreichen Lösungsansätze während unserer Kaffeepausen über Zoom danken.

Das gleiche gilt für Dr. Jochen Vestner, der für mich ein Mentor im Bereich Machine Learning und Python-Programmierung war. Ohne ihn wären Teile dieser Doktorarbeit nicht möglich gewesen. Danke auch an Dr. Patrick Nickolaus und Dr. Pascal Wegmann-Herr dafür, dass sie mich immer zum kritischen Hinterfragen angeregt haben oder aber auch bei önologischen, branchenspezifischen sowie labortechnischen Fragen immer weitergeholfen haben. Ein herzliches Dankeschön auch an Dr. Thi Nguyen für die exzellente sprachliche sowie fachliche Begutachtung meiner Publikationen.

Weiterhin möchte ich allen Absolventen danken, die ich betreuen durfte. Allen voran Marina Scheiermann, deren Arbeit unverzichtbar für mich war. Genauso möchte ich Tobias Baumgärtner und Johannes Hirner danken.

Ich bedanke mich auch bei allen Doktoranden und dem Arbeitskreis Durner, mit denen ich diese etwas mehr als drei Jahre verbringen durfte. Die Grillabende und alle anderen außerberuflichen Aktivitäten waren immer ein Fest. Namentlich möchte ich mich bedanken bei Svetlana Cvetkova, alias die beste Büropartnerin, die man sich wünschen kann. Weiterhin Sandra Feifel, Fabian Marnet, Daniel Zimmermann, Julian Döbler, Caroline Dietzel, Louis Backmann, Marc Weber, Jörg Gottmann, Caterina Szmania, Lisa Käßler und allen anderen.

Einen ganz besonderen Dank möchte ich an meinen ehemaligen Chemielehrer Matthias Decker richten, der mich innerhalb eines Jahres von einer 4- auf eine 2+ in Chemie gebracht hat. Ohne ihn hätte ich gar nicht Lebensmittelchemie studiert. Auch ein Dankeschön an Antje Salatzkat für die englische Korrektur meiner Dissertation.

Zuletzt möchte ich meiner wundervollen Verlobten Jeanine Bernhardt danken, die mich in der kompletten Zeit durch alle Höhen und Tiefen begleitet hat, die immer für mich da war und durch die ich immer wieder neue Blickwinkel auf eine Fragestellung bekommen habe.

---

## Publications

### Peer-reviewed Publications

**Hensel, Marcel;** Vestner, Jochen; Fahrner, Jörg; Durner, Dominik (2024): Evaluation of different machine learning algorithms to build an application classifying Blanc de noir wine with spectrophotometric data (submitted to American Journal of Enology and Viticulture).

**Hensel, Marcel;** Scheiermann, Marina; Fahrner, Jörg; Durner, Dominik (2023): New Insights into Wine Color Analysis: A Comparison of Analytical Methods to Sensory Perception for Red and White Varietal Wines. *Journal of Agriculture and Food Chemistry*

**Hensel, Marcel;** Di Nonno, Sarah; Mayer, Yannick; Scheiermann, Marina; Fahrner, Jörg; Durner, Dominik; Ulber, Roland (2022): Specification and Simplification of Analytical Methods to Determine Wine Color. *Processes*, 10(12): 2707

Becker Daniela; Bakuradze Tamara; **Hensel Marcel;** Beller Simone; Yélamos Carolina C., Richling Elke (2021): Influence of Brewer's Spent Grain Compounds on Glucose Metabolism Enzymes. *Nutrients*

### Technical Publications

**Hensel, Marcel;** Scheiermann, Marina; Fahrner, Jörg; Durner, Dominik: Fotometrische Farbanalytik und die sensorische Wahrnehmung der Weinfarbe. *Lebensmittelchemie*, 77, (2023)

Durner, Dominik; **Hensel, Marcel:** Is there a need to re-define the methods to evaluate wine color?- BIO Web of Conferences 68, 02008 (2023)

Di Nonno, Sarah; Scharfenberger-Schmeer, Maren; Durner, Dominik; **Hensel, Marcel;** Ulber, Roland: Photometrisches Analysesystem für die Weinanalytik; *Der Lebensmittelbrief* (2023) (34) 2, 26-27

### Oral Presentations

**Hensel, Marcel;** Di Nonno, Sarah; Scheiermann, Marina; Mayer, Yannick; Ulber, Roland; Durner, Dominik: Neue Impulse und Einblicke in die Farbbestimmung von Rot- und Weißweinen - 10. Awendertreffen der Weinanalytik (2022), Neustadt a. d. W., Germany

**Hensel, Marcel;** Scheiermann Marina; Durner, Dominik: New Insights into Wine Color Analysis: A Comparison of Analytical Methods and their Correlation with Sensory Perception – In vino analytica scientia (IVAS)(2022), Neustadt a. d. W, Germany

**Hensel, Marcel;** Vestner, Jochen; Durner, Dominik: Entwicklung und Anwendung einer Machine Learning Methode zur Unterscheidung von Blanc de Noir Weinen und Weißweinen auf Grundlage der Farbe – 63. Arbeitstagung vom Forschungsring des Deutschen Weinbaus (2023), Veitshöchheim, Germany

---

Durner, Dominik; **Hensel, Marcel**: Is there a need to re-define the methods to evaluate wine color? – 44. World Congress of Vine and Wine, Cardiz-Jerez, Spanien

**Hensel, Marcel**; Vestner, Jochen; Durner, Dominik: A Machine Learning Application to Differentiate White Wine, Blanc de noir, and Rosé Wine Based on CIEL\*a\*b\* – 74. National Conference of the American Society of Enology and Viticulture (ASEV) (2023), Napa, California, USA (**Ausgezeichnet mit dem "Best student presentation Award - Enology"**)

**Hensel, Marcel**; Vestner, Jochen; Durner, Dominik: Entwicklung und Anwendung einer Machine Learning Methode zur Unterscheidung von Blanc de Noir Weinen, Rosé, und Weißweinen auf Grundlage der Farbe – Oenologisches Symposium (2023), Neustadt a. d. W., Germany

### **Poster Presentations**

Di Nonno, Sarah; Könnel, E.; **Hensel, Marcel**; Scharfenberger-Schmeer, Maren; Durner, Dominik; Ulber, Roland: How to implement simple and affordable process analysis for specific applications?; Himmelfahrtstagung on Bioprocess Engineering 2024 - Novel Strategies and Technologies for Sustainable Bioprocesses and Bioproducts; Regensburg, Deutschland

## Contents

<b>Eidesstattliche Erklärung</b>	<b>I</b>
<b>Lebenslauf</b>	<b>III</b>
<b>Publications</b>	<b>VI</b>
Peer-reviewed Publications . . . . .	VI
Technical Publications . . . . .	VI
Oral Presentations . . . . .	VI
Poster Presentations . . . . .	VII
<b>Zusammenfassung</b>	<b>1</b>
<b>Abstract</b>	<b>2</b>
<b>1 Introduction</b>	<b>3</b>
<b>2 The Color of Wine</b>	<b>4</b>
2.1 The Chemistry of Wine Color . . . . .	5
2.1.1 Anthocyanidins and Anthocyanins . . . . .	5
2.1.2 Tannins and Proanthocyanidins . . . . .	7
2.1.3 Copigmentation . . . . .	8
2.1.4 Hydroxycinnamates and Quercitin derivatives . . . . .	9
2.1.5 Effects of enological factors on the wine color . . . . .	10
2.2 Introduction about spectrometry, spectroscopy, and photometry . . . . .	12
2.2.1 Principles of Molecular Spectroscopy . . . . .	12
2.2.2 Lambert Beers Law . . . . .	13
2.2.3 Photometer systems . . . . .	14
2.2.4 Monochromator systems . . . . .	15
2.2.5 Detector Systems . . . . .	15
2.3 The Quantification of Wine Color . . . . .	18
2.3.1 Theory of colors . . . . .	18
2.3.2 Glories' color evaluation . . . . .	20
2.3.3 CIE Color Spaces . . . . .	20
2.4 Calculation of color distances . . . . .	25
2.4.1 Euclidean color difference . . . . .	26
2.4.2 Development of the CMC( <i>l:c</i> ) Color Distance Formula . . . . .	26
2.4.3 Development of the CIE94 Color Distance Formula . . . . .	27
2.4.4 Development of the CIEDE2000 Color Distance Formula . . . . .	28
2.5 Human perception of Wine Color . . . . .	32
2.5.1 Descriptive Analysis . . . . .	32
2.5.2 Sensory threshold Analysis . . . . .	33

---

<b>3</b>	<b>A detailed description of statistical modeling, machine learning, and Artificial Intelligence</b>	<b>33</b>
3.1	Logistic Regression . . . . .	34
3.2	Support Vector Machine . . . . .	35
3.2.1	Linearly separable data . . . . .	35
3.2.2	Non-linear separable data . . . . .	38
3.2.3	Soft margin SVM . . . . .	39
3.3	Extreme Gradient Boosting . . . . .	41
3.3.1	Objective function . . . . .	41
3.3.2	Gradient Boosting . . . . .	42
3.3.3	Tree pruning and other mechanism to prevent overfitting . . . . .	45
3.4	Neural Networks, Multilayer Perceptron Models, and Deep Learning . . . . .	46
3.5	Model Performance Measures . . . . .	50
3.5.1	Accuracy, Precision, Recall, F1-Score . . . . .	50
3.5.2	Receiver Operating Characteristics Curve . . . . .	50
<b>4</b>	<b>Scope and aim</b>	<b>52</b>
<b>5</b>	<b>Cumulative part of the dissertation</b>	<b>54</b>
5.1	Specification and Simplification of Analytical Methods to Determine Wine Color . . . . .	54
5.2	New insights into Wine Color Analysis: A Comparison of Analytical Methods to Sensory Perception of Red and White Varietal Wines . . . . .	82
5.3	Conference Proceeding- Is there a need to redefine the methods to evaluate wine color? . . . . .	103
5.4	Evaluation of different machine learning algorithms to build an application classifying Blanc de noir wine with spectrophotometric data . . . . .	116
<b>6</b>	<b>Additional relevant unpublished results</b>	<b>152</b>
6.1	Blanc de Noir Check Web Application . . . . .	152
6.2	The Classification of Red wines . . . . .	153
6.2.1	Samples and photometric measurement . . . . .	154
6.2.2	Support Vector Machine Modeling . . . . .	154
6.2.3	Neural Network architecture . . . . .	155
6.2.4	Results and Discussion . . . . .	156
<b>7</b>	<b>Concluding remarks</b>	<b>159</b>
<b>8</b>	<b>References</b>	<b>161</b>
<b>I</b>	<b>Appendix – Basic Framework of the Web Application in Full display</b>	<b>170</b>
<b>II</b>	<b>Appendix – Used Red Wines for section 6.2</b>	<b>171</b>

## List of Figures

1	Base structure of flavonoid polyphenols (modified according to Belitz et al. 2001) . . . . .	5
2	Base structure of anthocyanidins/anthocyanins (modified according to Belitz et al. 2001)	5
3	pH dependent forms of anthocyanins with the blue colored chinoidic form, the red colored flavylium cation, the colorless carbinol bases, and the pale yellow chalcone form (modified according to Ribéreau-Gayon 2006) . . . . .	6
4	Structure of the most abundant acylated anthocyanin glucosides. $R_1$ and $R_2$ can either be a methyl group ( $-\text{CH}_3$ ), a hydroxy group ( $-\text{OH}$ ), or an hydrogen ( $-\text{H}$ )(modified according to García-Beneytez et al. 2003). . . . .	7
6	Structure of flavan-3-ols (modified according to Ribéreau-Gayon 2006). . . . .	8
8	Initial structure of substances that form pigments in white wine (modified according to Singleton et al. 1979; Belitz et al. 2001) . . . . .	9
9	Structures of higher molecular flavonoid polyphenols as a result of aging. Modified according to Asenstorfer, Hayasaka, et al. 2001; Asenstorfer, Lee, et al. 2006 ( $R_1, R_2 = -\text{H}, -\text{OH}, -\text{OCH}_3$ ; $R_3 = -\text{H}, \text{O-Glucose}$ , $R_4 = -\text{H}$ (Vitisin A), $-\text{COOH}$ (Vitisin B)) . . . . .	10
10	Structure and reaction of anthocyanins with added sulfur dioxide (modified according to Jurd 1964) . . . . .	11
11	Depiction of the excitation of electrons (absorption), radiation-free transition back from excited singlet state to ground state (emission), fallback from excited singlet state to ground state with emission of radiation after internal conversion (fluorescence), and fallback from excited triplet state to ground state with emission of radiation after intersystem crossing from singlet to triplet state (phosphorescence) . . . . .	13
12	Optical arrangement of single beam and double beam photometers (modified according to Matissek et al. 2018) . . . . .	15
13	Schematic assembly and measurement mechanism of photomultiplier tubes (modified according to Löffler-Mang 2012; Shimadzu 2023) . . . . .	16
14	Three dimensional crystal structure of silicon as well as a p-type and n-type silicon semiconductor (modified according to Riedel and Janiak 2022) . . . . .	17
15	Schematic assembly and mechanism of a PIN-junction in a PIN-photodiode . . . . .	18
16	Depiction of the vector combination regarding the additive color mixing (modified according to Lübbe 2013) . . . . .	19
17	Comparison of the additive color combination and the subtractive color combination (modified according to Lübbe 2013; Bühler et al. 2018) . . . . .	19
18	Spectral power distribution of the standard illuminant A, D50, and D65 (modified according to CIE 2019c) . . . . .	21
19	Schematic depiction of the original experiment of Wright and Guild and the resulting color matching function that contains the imaginary negative parts (W. D. Wright 1929; Guild J. 1931). . . . .	22
20	Color matching function of the CIE XYZ color space (CIE 2019a). . . . .	22
21	CIE Chromaticity diagram according to CIE and the irregularities discovered by MacAdam	24
22	Depiction of the CIE $L^*a^*b^*$ color space, where $L^*$ represents the lightness, $a^*$ either red ( $a^* \in \mathbb{R}^+$ ) or green ( $a^* \in \mathbb{R}^-$ ), and $b^*$ either yellow ( $b^* \in \mathbb{R}^+$ ) or blue ( $b^* \in \mathbb{R}^-$ ) . . . . .	24

23	Color ellipses plotted by Luo and Rigg in the $a^*b^*$ color plane from experimental data (modified according to Schanda 2007; Luo and Rigg 1986; Luo, Cui, et al. 2001). . . . .	29
24	Dependence of $G$ with increasing chroma (from Schanda 2007) . . . . .	29
26	Linear line scales used for the description of red wine and white wine. The first scale evaluates the lightness of a wine. The second scale evaluates the color hue of a wine (modified according to Durner et al. 2010) . . . . .	32
27	Visual description of the basic ideas of logistic regression (modified according to Harrell 2001) . . . . .	35
28	The basic principle of SVM. The most suitable hyperplane has to be identified. With the Hyperplane and the support vectors the margin can be derived. . . . .	36
29	Example for a dataset that is not linear separable. The kernel function transforms the two-dimensional input space in a three-dimensional feature space (from Jordans 2004) . . . . .	38
30	Training dataset with outliers and depiction of the effect of the slack variable $\xi$ . The slack variable is the distance between the misclassified point and the margin that borders the respective class. . . . .	39
31	The summary of the main ideas of Extreme Gradient Boosted trees (modified according to Vestner 2021) . . . . .	42
32	General architecture of a multilayer perceptron neural network (modified according to Bhowmik 2019) . . . . .	46
33	Mode of operation of a single neuron in the neuronal network (modified according to Sonnet 2022). . . . .	47
34	Depiction of the Heaviside, ReLU, and Sigmoidal activation functions of the neuron Kruse et al. 2015; Sonnet 2022. . . . .	48
35	The process of forward propagation in a three layer neural network with one input layer, one hidden layer with two neurons and an output layer (modified according to (Starmer 2022)). . . . .	48
36	Example ROC curve for two classifiers with different performances. Performance of Classifier 1 is better than Classifier 2. The $45^\circ$ line is considered a random guessing classifier. . . . .	51
37	Flowchart that describes the workflow of the web application. . . . .	152
38	Output messages of the web application depending on the user interaction and the prediction result of the SVM . . . . .	153
39	Data pipeline starting from the transmission spectra as raw data through the different computation steps. Color indicates different spectrophotometric input data. All input datasets were used to fit and optimize an SVM. The optimized models were then evaluated with performance parameters. . . . .	155
40	Depiction of the data processing starting by scaling the data and using PCA for dimensionality reduction. The principle components of the PCA are then used as input for the MLP model with one hidden layer. . . . .	156
41	Depiction of the ROC and for the test dataset (A) and training dataset (B). . . . .	157
42	Depiction of the mean absorption spectra (A) and the mean transmission spectra (B) in the dataset ( $n=366$ ). The unsaturated areas are the standard deviation. . . . .	158
43	Basic framework of the web application after completely loading. . . . .	170



## List of Tables

1	Names and corresponding residues of different anthocyanins (modified according to Belitz et al. 2001) . . . . .	5
2	Names and corresponding residues for of catechin derivatives . . . . .	8
3	Names and corresponding residues for of epicatechin derivatives . . . . .	8
4	Names and corresponding residues for of Quercitin derivatives . . . . .	9
5	Kernel function of the most commonly used kernels. . . . .	39
6	Performance parameters precision, recall, and F1-score of SVM on absorption spectra, transmission spectra, and CIE L*a*b* coordinates as well as MLP on absorption spectra for training and test dataset. . . . .	157

## Abbreviations used

Abbreviations	Description
$a^*$	Hue parameter for red and green in the CIE $L^*a^*b^*$ color space
AI	Artificial Intelligence
AUC	Area under curve
$b^*$	Hue parameter for yellow and blue in the CIE $L^*a^*b^*$ color space
C	Penalty for miscalsification
$C^*$	CIE Chroma
CI	Color Intensity
CIE	Internationle Beleuchtungskommission (franz.: Commission Internationale de l'Eclairage)
CIE XYZ	Tristimulus color space
CMC	Color Measurement Committee
$dA(\%)$	Brillance
$\Delta E_{00}$	CIEDE2000 Color distance
$\Delta E_{ab}$	Euclidean color distance
eV	Electron Volts
FN	False negative count
FP	False positive count
FPR	False positive rate
$h^0$	CIE hue angle
$H_0$	Null-hypothesis
$H_1$	Alternative Hypothesis
HOMO	Highest occupied molecular orbital
INOX	Steel Tank
JND	Just noticeable difference/ visual color threshold
$L^*$	Lightness parameter in the CIE $L^*a^*b^*$ color space
L1	Lasso Regression
L2	Ridge Regression
$\lambda$	Wavelength
LED	Light Emitting Diodes
LUMO	Lowest unoccupied molecular orbital
ML	Machine Learning
MLP	Mutlilayer-Perceptron Model
N	All negative Instances
OIV	Internationale Organisation für Rebe und Wein (franz.: Organisation Internationale de la Vigne et di Vin)
P	All positive instances
PCA	Principle Component Analysis
$\phi(\lambda)$	color stimulus function
$\phi(x_i)$	feature space

---

Abbreviation	Description
PIN-Photodiode	Positive/Intrinsic/Negative Photodiode
PMT	Photomultiplier tube
PRD	Plotting Reference Dataset
RBF	Radial basis function
ReLU	Rectifier Linear Unit Function
RGB	Red-Green-Blue
ROC	Receiver operating characteristics
SDC	Society of Dyers and Colorists
STRESS	Standardized Residual Sum of Squares
SVM	Support Vector Machine
T	Transmission / Glories Hue
TD	Training dataset
$\Theta$	Threshold neural networks
TN	True negative count
TP	True positive count
TPR	True positive rate
UV	Ultraviolet
Vis	Visible light
XGBoost	Extreme Gradient Boosted Trees

---

## Zusammenfassung

Die Farbe von Wein ist ein wichtiger Qualitätsparameter, der für den ersten Eindruck der Verbraucher entscheidend ist. Die *Organisation Internationale de la Vigne et du Vin* (OIV) empfiehlt zwei Methoden zur Beschreibung der Weinfarbe: die Farbberechnung nach Glories und die Bestimmung der Koordinaten im  $L^*a^*b^*$ -Farbraum der *Commission Internationale de l'Eclairage* (CIE). In dieser Arbeit wird der Einfluss von Photometereinstellungen auf die Berechnung des CIE  $L^*a^*b^*$ -Farbraumes ermittelt. Es konnte gezeigt werden, dass die Einstellungen des Photometers die Reproduzierbarkeit der Messung beeinflussen. Weiterhin wurde die Farbmessung nach Glories mit dem CIE  $L^*a^*b^*$ -Farbraum verglichen. Die Ergebnisse zeigen eine schwache Korrelation bei hellem Rotwein und Weißwein. Daher können die Glories-Methode und der CIE  $L^*a^*b^*$ -Farbraum nicht austauschbar verwendet werden. Um festzustellen, welche der Methoden für weitere Untersuchungen besser geeignet sind, wurde die Farbmessung nach Glories und dem CIE  $L^*a^*b^*$ -Farbraum mit der visuellen Wahrnehmung der Weinfarbe bei 112 Rot- und Weißweinen verglichen. Die Ergebnisse deuten auf eine bessere Eignung des CIE  $L^*a^*b^*$ -Farbraums hin, da er die vom Menschen wahrgenommene Farbe besser abbildet. Die euklidische Farbdifferenz ist die derzeit von der OIV empfohlene Formel für den Vergleich von Weinfarben. Die CIE empfiehlt jedoch die CIEDE2000-Farbabstandsformel, die sich als präziser erwiesen hat. Daher wurde die visuelle Wahrnehmungsschwelle mit der CIEDE2000-Farbabstandsformel durch Dreieckstests neu bewertet. Da CIE  $L^*a^*b^*$  eher der menschlichen Wahrnehmung entsprach, sollte dieser Farbraum der Methode nach Glories vorgezogen werden. Die visuellen Farbschwellen wurden mit CIEDE2000 besser ausgedrückt, variierten aber immer noch je nach Farbbereich im CIE  $L^*a^*b^*$ -Farbraum. Die Ergebnisse dieser Studien zeigen, dass der CIE  $L^*a^*b^*$ -Farbraum für weitere Untersuchungen besser geeignet ist.

Maschinelles Lernen (ML) und statistische Modellierung haben sich zu wichtigen Innovationen in der Wissenschaft entwickelt. In der Weinforschung wird ML häufig eingesetzt, um abstrakte Parameter wie die Weinqualität auf der Grundlage komplexer instrumenteller chemischer Analysen vorherzusagen. Die vorliegende Studie verwendet spektrophotometrische Daten und CIE  $L^*a^*b^*$ -Koordinaten von 176 kommerziellen Weinen, um Blanc de noir von Roséwein und Weißwein zu unterscheiden. Die Transmissionsspektren wurden verwendet, um extreme gradient-boosted trees und eine Support Vector Machine (SVM) zu trainieren. Die CIE  $L^*a^*b^*$ -Koordinaten wurden zum Trainieren von SVM und logistischer Regression verwendet. Nach Parameter-Hypertuning lieferte die Kombination von SVM auf CIE  $L^*a^*b^*$ -Daten die optimale Klassifizierung mit einer kreuzvalidierten Genauigkeit von 0,88 und einem F1-Score von 0,93. Das endgültige Klassifizierungsmodell wurde in einem browserbasierten, benutzerfreundlichen Dashboard für Winzer und andere Nutzer, wie z. B. Weinlabore, bereitgestellt. Ein weiterer Teil dieser Arbeit ist die Klassifizierung von im Fass gereiftem Rotwein. Die Transmissionsspektren von 363 Rotweinen wurden gemessen und in Absorptionsspektren und CIE  $L^*a^*b^*$  Koordinaten umgewandelt. Transmissionsspektren, Absorptionsspektren und CIE  $L^*a^*b^*$ -Koordinaten wurden zum Trainieren einer SVM verwendet. Darüber hinaus wurden die Absorptionsspektren auch zum Trainieren eines Perceptron-Modells verwendet. Die Spektren wurden skaliert und mittels Hauptkomponentenanalyse (PCA) transformiert, um die Dimensionalität zu reduzieren. Die Leistung der SVM bei den Transmissionsspektren wurde von der SVM bei den Absorptionsspektren und den CIE- $L^*a^*b^*$ -Koordinaten übertroffen. Die beste Leistung erzielte das neuronale Netz mit einem F1-Wert von 0,75.

## Abstract

The color of wine is an important quality parameter essential for the first impression of consumers. The *Organisation internationale de la Vigne et du Vin* (OIV) recommends two methods to describe wine color: color calculation according to Glories and the determination of coordinates in the  $L^*a^*b^*$  color space of the *Commission Internationale de l'Eclairage* (CIE). In this work, the influence of photometer settings on the calculation of the CIE  $L^*a^*b^*$  color space were determined. It was shown that the photometer settings influence the reproducibility of the measurement. Furthermore, the color measurement according to Glories is compared to the CIE  $L^*a^*b^*$  color space. The results show a weak correlation in the light red wine and white wine color area. Therefore, Glories' color measurement and the CIE  $L^*a^*b^*$  color cannot be used interchangeably. To determine, which of the methods is more suited for further investigation, the color measurement according to Glories and the CIE  $L^*a^*b^*$  color space were compared to the visual perception of 112 red and white wines. The results indicate that the CIE  $L^*a^*b^*$  color space is better suited to depicting the color perceived by humans. Since its development, the CIE color spaces have undergone various changes. The possibility of comparing colors has been no exception. The Euclidean color difference is the formula currently recommended by the OIV to compare wine colors. However, the CIE recommends the CIEDE2000 color distance formula, which has been proven to be more precise. The reason why the Euclidean color difference is still used in wine research is the absence of reference values calculated with the CIEDE2000 color distance formula for the just noticeable difference (JND), or the visual color threshold, the minimum difference in color hue that is visible by the human eye. Therefore, the JND was re-evaluated with the CIEDE2000 color distance formula via triangle testing. Compared to Glories' color measurement, CIE  $L^*a^*b^*$  more closely match the human perception, elevating the use of CIE  $L^*a^*b^*$  over the use of the Glories method. Visual color thresholds were better expressed with CIEDE2000 but still varied depending upon the color area in the CIE  $L^*a^*b^*$  color space. The results of these studies indicate that the CIE  $L^*a^*b^*$  color space is better suited for further investigation.

Machine learning (ML) and statistical modeling have emerged as important innovations in science. In wine research, ML is often used to predict abstract parameters such as wine quality based on complex instrumental chemical analysis. The presented study used spectrophotometric data and CIE  $L^*a^*b^*$  coordinates from 176 commercial wines to distinguish Blanc de noir from rosé wine and white wine. The transmission spectra were used to train extreme gradient-boosted trees (XGBoost) and a support vector machine (SVM). CIE  $L^*a^*b^*$  coordinates were used to train SVM and logistic regression. After parameter hypertuning, the combination of SVM on CIE  $L^*a^*b^*$  data provided the optimal classification with a cross-validated accuracy of 0.88 and a F1 score of 0.93. The final classification model is deployed in a browser-based, user-friendly dashboard for winemakers and other users, such as wine laboratories. SVM was also applied in the context of classification of barrel-aged red wine. The transmission spectra of 363 red wines were measured and transformed into absorption spectra and CIE  $L^*a^*b^*$  coordinates. Transmission spectra, absorption spectra, and CIE  $L^*a^*b^*$  coordinates were used to train an SVM. Furthermore, the absorption spectra were used to train a multilayer perceptron model. The spectra were preprocessed and transformed with principal component analysis (PCA) to reduce dimensionality. The performance of SVM on transmission spectra was outperformed by SVM on absorption spectra and CIE  $L^*a^*b^*$  coordinates. The best performance was achieved by the neural network/MLP, with an F1 score of 0.75.

# 1 Introduction

Color is one of wines most important quality parameters as it raises expectation about olfactory and gustatory properties. Consumers come to anticipate certain things based on experience and regular encounters with specific combinations of color and aroma. Such associations are presumably acquired through a process of learning, given their variability across diverse cultural contexts (Shankar et al. 2010). These associations are described as "prototypes" of wine. The existence of prototypes implies that human evaluation of wine aroma may not be based on specific sensorial properties, but rather on cognitive association stemming from previous encounters and experience with wine (Brochet and Dubourdieu 2001). Consequently, the perception of wine aroma can be altered by color, because consumers search for familiar the aromas they most associate with certain colors, even though these aromas may be absent. Generally, white wine is described with yellow colored or colorless objects such as lemon, honey, hazelnut, and butter. Red wine, however, is often described with red or dark-colored objects like dark cherry, blackberry, or cassis (Ballester et al. 2009).

Numerous studies with dyed wines have been carried out to evaluate the cognitive nature of wine perception. A study conducted in 1963 showed an increase in perceived sweetness if white wine was dyed pink (Pangborn et al. 1963). Another study showed that white wines dyed red were perceived with typical descriptors of red wines, whereas the undyed white wine was perceived with typical descriptors of white wine (Morrot et al. 2001). In more recent studies, white wines were dyed to resemble the color of rosé wines inducing aromas like rose or strawberry (Wang and Spence 2019). In 2023, Nguyen and Durner demonstrated how many descriptors of wine aroma are color-driven. Switching from black to clear classes, led a panel of trained judges to perceive an increase of the odors green fruit, citrus fruit, and stone fruit in white wine, whereas the aromas dark fruit, red fruit, and oak decreased. The reverse was true in red wine (Nguyen and Durner 2023).

Besides the alteration of the perceived aromatic profile, color also influences perceived quality impression. Parpinello et al. 2009 compared the perceived typicality of Italian Novello wines. The majority of inexperienced consumers in this study preferred wines with high color intensity and linked the color to a higher perceived quality (Parpinello et al. 2009). Wine color has also an impact on the perceived typicality of Provence rosé wines: the lightest, clearest wines were generally considered the most typical, whereas wines with a redder character were considered the least typical rosé wine (Coulon-Leroy et al. 2018).

The color of a wine depends on various factors, some of which can be controlled by the winemaker altering the wine color. For example, the extraction of the colorants can be influenced by fermentation temperature (Girard et al. 1997), maceration time (Gómez-Plaza et al. 2001; Kovac et al. 1992), or oxygen uptake via micro-oxygenation or barrel aging (Durner et al. 2010; Oberholster et al. 2015). Other parameters cannot be influenced, e.g. grape variety, and weather conditions. Hence, the more cognitive than sensorial evaluation of wine color (Nguyen and Durner 2023) as well as the impact of wine color on perceived quality (Parpinello et al. 2009), typicality (Coulon-Leroy et al. 2018), and buying decisions (Issa-Issa et al. 2021). Sometimes state regulations are tied to the color of a wine (Bundesrat Drucksache 175/21 2021) which implies the need for an analytical evaluation that still resembles the human perception and therefore objectifies it.

The OIV released a compendium containing protocols recommended for the use by winemakers and wine laboratories (OIV 2021). Two photometric methods are recommended for the wine color. The first method was developed in France by Yves Glories for dark red wine and consists of absorbance values at three wavelengths. With these primary absorbance values, parameters like color intensity, hue and red content can be evaluated. The second, more complex, method is the  $L^*a^*b^*$  color space, according to the CIE, which consists of the lightness parameter  $L^*$  as well as the hue parameters  $a^*$ , and  $b^*$  (CIE 2019b). To calculate these parameters a complete transmission spectrum is needed. These methods differ in complexity, and scope of definition. Glories' method was developed specifically for wine using dark red grape varieties in the process, while the CIE  $L^*a^*b^*$  color space was developed for color measurement in general. Wine color as a parameter is mostly used for descriptive reasons. Numerous studies have measured wine color with the sole premise to validate the perceived color obtained by sensory evaluations (Renner et al. 2022; Durner et al. 2010; Alcalde-Eon, Ferreras-Charro, et al. 2019). Machine learning could help to expand the application spectrum of wine color.

With the recent developments in Artificial Intelligence (AI), ML and other methods for statistical models the focus of this research has shifted to more practicability and user-oriented applications. Given the recent established digital infrastructure the acceptance of computer assisted evaluations is growing. However, it is unclear how these models perform in wine analysis or, to be more specific, how these models perform using transmission and color data. Countless ML models, which are all vastly different operating with different levels of complexity as well as different mathematical backgrounds can be used for classification or regression purposes. For example, logistic regression uses a probabilistic approach to classify the given data (Harrell 2001). Other algorithms like SVM use linear algebra for the same purpose (Boser et al. 1992; Bennett and Campbell 2000). The more complex algorithms use different combinations of mathematical methods e.g. extreme gradient boosting, which uses a combination of numerical and analysis approaches (Chen and Guestrin 2016).

It was imperative to this work to refine our understanding about wine color and the link to human perception of wine. To achieve this, the protocols to obtain CIE  $L^*a^*b^*$  coordinates were refined by investigating the impact of photometer settings. Glories' color measurement and CIE  $L^*a^*b^*$  were compared to each other and lastly, it was investigated if it is possible to extract more information from color data.

## 2 The Color of Wine

The color of a wine depends primarily on the grape variety. Based on this, wines can be divided into red wines and white wines. The production process of these wines are very different. After harvesting, red grapes are fermented on the mash and are pressed after fermentation. White wines on the other hand are pressed first and the obtained juice is then fermented. The pigments responsible for the wine color are usually located in the berry skins (Jackson 2008; Ribéreau-Gayon 2006). In red wine production wines with high color intensity are desired, so the extraction of the pigments from skin to wine is crucial, leading to long maceration times to extract as many pigments as possible. In white wine, a dark color is not desired, so the berries are pressed directly and then fermented. A special case of red wines are rosé wines, which stand out due to their light red to salmon color. They are also obtained from red grapes, however the maceration time is not as long or none at all as in normal red wine (Jackson 2008). Another special case is Blanc de noir. Here, the goal is to produce a wine from red grapes that has the light appearance more closely depicting white wine. Both, rosé and Blanc de noir are produced

more like a white wine than a red wine. In Germany, to sell a Blanc de noir wine the wines undergo a quality assessment by wine control entities and it must have a color typical for white wine (Bundesrat Drucksache 175/21 2021). Therefore, the color of a wine is a crucial parameter when it comes to quality assessments and has to be closely observed during the production process. The chemical composition of the colorants is key to the lightness and hue of the wine. The following section describe the chemical characteristics of wine color, the analysis, and the quantification of the human perception of wine color.

## 2.1 The Chemistry of Wine Color

To understand the complexity of wine color, it is imperative to understand the underlying chemistry. An important class of compounds that determine wine color are the flavonoid polyphenols. The basic structure of these compounds is a benzopyran group attached to a phenyl group. The aromatic part of the benzopyran structure is called the A-ring, whereas the phenyl group is called the B-ring. The remaining non-aromatic part of the benzopyran structure is the C-ring (Figure 1). Flavonoid polyphenols can be divided into six subgroups: flavone, flavanols, flavonols, isoflavonoids, and anthocyanins. The polyphenols directly related to color are anthocyanins (Belitz et al. 2001).

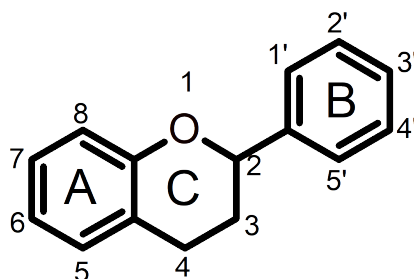


Figure 1: Base structure of flavonoid polyphenols (modified according to Belitz et al. 2001)

### 2.1.1 Anthocyanidins and Anthocyanins

The flavonoids most important to red wine color are anthocyanins, due to their colorful appearance. They are located in the berry skin and in the leaves of the vine in significant quantities by the end of the growing season (Ribéreau-Gayon 2006). They are composed of an anthocyanidine aglycone linked with a glucosyl group (Figure 2).

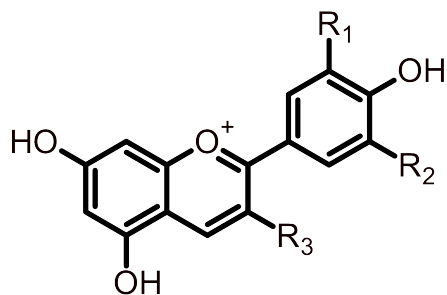


Table 1: Names and corresponding residues of different anthocyanins (modified according to Belitz et al. 2001)

<b>Anthocyanidins</b>	R <sub>1</sub>	R <sub>2</sub>	R <sub>3</sub>
Malvidin	-OCH <sub>3</sub>	-OCH <sub>3</sub>	-OH
Delphinidin	-OH	-OH	-OH
Peonidin	-H	-OCH <sub>3</sub>	-OH
Petunidin	-OCH <sub>3</sub>	-OH	-OH
Cyanidin	-OH	-H	-OH
<b>Anthocyanins</b>	R <sub>3</sub> =O-Glucose		

Figure 2: Base structure of anthocyanidins/anthocyanins (modified according to Belitz et al. 2001)





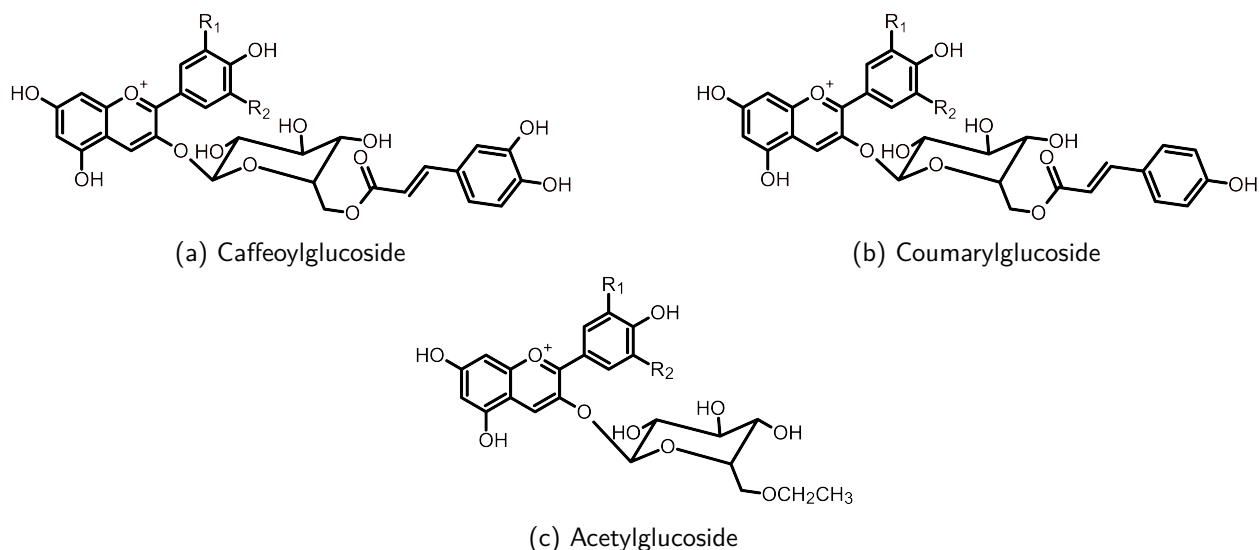


Figure 4: Structure of the most abundant acylated anthocyanin glucosides.  $R_1$  and  $R_2$  can either be a methyl group ( $-\text{CH}_3$ ), a hydroxy group ( $-\text{OH}$ ), or an hydrogen ( $-\text{H}$ ) (modified according to García-Beneytez et al. 2003).

where cyanidin and peonidin glucosides are the most abundant anthocyanins. This is a distinctive feature as in all other groups the malvidin-O-glucoside is the most prevalent anthocyanin. The representatives of this group include the eponym Trollinger as well as Sangiovese, which is used for the Chianti production in Italy, and Pinotage among others. The amount of acylated anthocyanins is relatively low, with acetylated and coumarylated anthocyanins averaging 3%. The largest group is the malvidin group, where malvidin-O-glucoside is the most abundant anthocyanin. The difference to the burgundy group is the high amount of acylated anthocyanins. Grape varieties in this group include Cabernet Sauvignon, Cabernet Franc, Lemberger, and Syrah among others. Due to the high amount of malvidin and other anthocyanins with three substituents at the phenol ring the musts and wines from these grapes exhibit a bluer coloration that is more intense than that of the Trollinger group. The final group is the hybrid group, where 3,5-Diglucosides are present in the wines. Under normal conditions grapes from *vitis vinifera* species do not produce anthocyanin diglucosides. However, the grape varieties that contain these diglucosides are hybrids from *vitis vinifera* and *vitis rupestris* species. *V. rupestris* is a grape variety that is used to produce fungus-resistant grape varieties. One consequence of this combination is that the majority of hybrids contain anthocyanin diglucosides inherited from the wild type (Wenzel et al. 2015).

### 2.1.2 Tannins and Proanthocyanidins

Tannins are brown, yellowish polymeric polyphenols, thereby directly influencing the color. They are present in both the berry seeds and the berry skins and differ in constitution, olfactive, and gustative properties. Two types of tannins can be observed in wine: gallo- and ellagitannins, and flavonoidic tannins. Gallo- and ellagittannins are known to release gallic and ellagic acid as a result of acidic hydrolysis. The second class of tannins is composed of flavan-3-ols including (+)-catechin and (–)-epicatechin (Figure 6). Additional constituents of tannins include the gallo- and epigallocatechins as well as the catechin- and epicatechin-3-O-gallate (Figure 6). Tannins have the ability to form stable complexes with proteins and other plant polymers (Hagerman and Butler 1978). In order for this reaction to occur, the molecular weight of the tannin must be between 600 and 3500 g/mol. If the molecular weight is below this threshold the complexes are not stable enough to form. Conversely, if the molecular weight exceeds

this threshold, the complexes will not form, as the protein sites are too distant (Ribéreau-Gayon 2006).

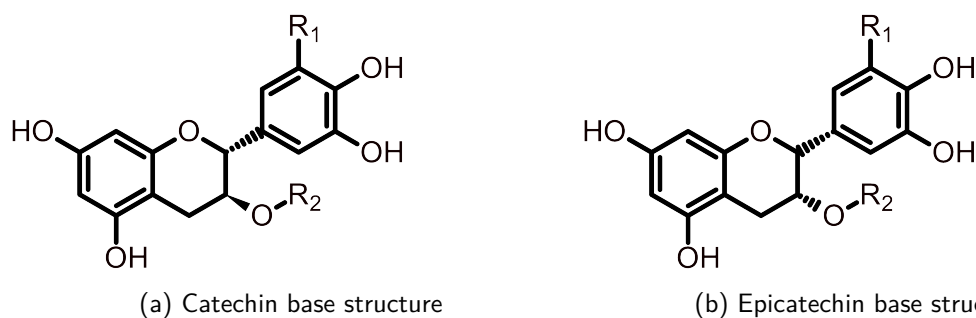


Table 2: Names and corresponding residues for of catechin derivatives

Name	R <sub>1</sub>	R <sub>2</sub>
Catechin	-H	-H
Galocatechin	-OH	-H
Catechin-3-O-gallate	-H	Gallic acid

Table 3: Names and corresponding residues for of epicatechin derivatives

Name	R <sub>1</sub>	R <sub>2</sub>
Epicatechin	-H	-H
Epigallocatechin	-OH	-H
Epicatechin-3-O-gallate	-H	Gallic acid

Figure 6: Structure of flavan-3-ols (modified according to Ribéreau-Gayon 2006).

The heating of flavanoidic tannin solutions in an acidic environment leads to the release of cyanidin or delphinidin, depending on the constitution of the terminal flavan-3-ol unit (Bate-Smith 1954), which leads to the alternative name proanthocyanidins. Proanthocyanidins derived from the grape seeds tend to exhibit a higher concentration of epicatechin-gallate, as well as a higher concentration of galloylated units, whereas in berry skin prodelphinidin units are present (Brossaud et al. 2001).

### 2.1.3 Copigmentation

The phenomenon of copigmentation occurs when a non- colored substance (copigment) forms a complex with a pigment, thereby altering the absorption properties of the pigment in question. This can result in a bathochromic or hypsochromic shift in the absorbance maximum or a hyperchromic shift in the absorbance intensity (Boulton 2001). To investigate copigmentation it is essential to differentiate between copigmentation and self-association. In the latter, anthocyanins complexes exhibit comparable, yet not identical outcomes as a anthocyanin copigment complex. Copigmentation typically results in a bathochromic shift, accompanied with an increase in absorbance values (hyperchromic effect), whereas self association usually results in a hypsochromic shift (Boulton 2001). Numerous studies in different model solutions, in wine, and in juice have identified numerous substances as potential copigments, with polyphenols being one example. In model solutions, hydroxybenzoic acids, including syringic, protocatechuic, and gallic acid have been identified as potential copigments (Zhang et al. 2015). Other studies have demonstrated that hydroxycinnamic acids, such as ferulic acid and caffeic acid in model solution can lead to copigmentation (Eiro and Heinonen 2002; Gómez-Míguez et al. 2006). In wine, the substitution of caffeic acid in a young wine has been found to enhance color (Darias-Martin et al. 2001). Naturally colorless flavonoid polyphenols including rutin, quercetin, epicatechin, and catechin, could be identified either in model solutions or added to wine as copigments (Schwarz et al. 2005; Lambert et al. 2011; Mirabel et al. 1999; Kopjar and Piližota 2011). Contrarily, another study identified catechin and quercetin to decrease the copigmentation effect (Darias-Martin et al. 2001; Rustioni et al. 2012). In addition to polyphenols, other studies have identified pH-value, potassium content, temperature, and molar ratio between different copigments as an influence on copigmentation (Miniati et al. 1992;

Czibulya et al. 2012; Gordillo et al. 2012). The reaction mechanism leading to copigmentation is not yet fully understood. Various theories have been proposed to explain the phenomenon, the most common being the hypothesis that hydrophobic  $\pi$ - $\pi$  interactions between the chromophore of an anthocyanin and the unsaturated part of the copigment are responsible (Boulton 2001; Heras-Roger et al. 2016). Despite this, the impact of copigmentation accounts for a significant proportion of the color of young wines, with estimates ranging 30% to 50%.

#### 2.1.4 Hydroxycinnamates and Quercetin derivatives

The color of white wine differs from that of red wine in the absence of red-colored anthocyanins. Nevertheless, white wines contain a light to bright yellow color. One of the main aspects of the coloration of white wines is in the presence of catechin and hydroxycinnamates, especially caftaric and coumaric acid (Figure 7a and Figure 7b). Initially colorless, these two compounds undergo an enzymatic oxidation reaction during the wine production process, resulting in a yellow/brown tint (Singleton et al. 1979; Ribéreau-Gayon 2006). The browning reaction starts when the grapes are crushed. Enzymes are released, that rapidly oxidize the hydroxycinnamates forming quinones (Waterhouse 2002). Additional coloration of white wines is derived from the flavonones quercetin, quercitrin, and isoquercitrin (Figure 7c)(Singleton et al. 1979).

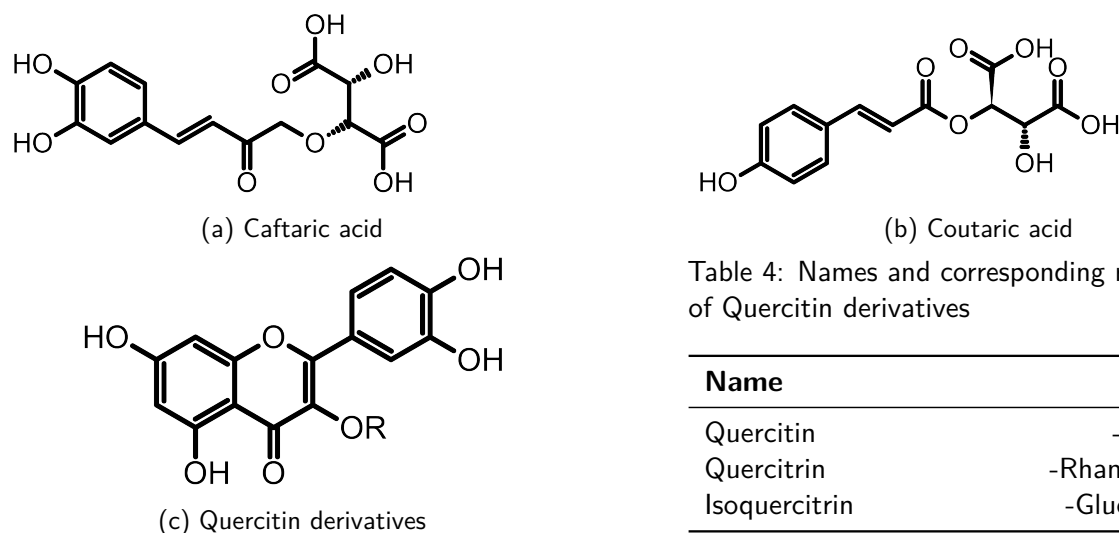


Table 4: Names and corresponding residues for of Quercetin derivatives

Name	R
Quercetin	-H
Quercitrin	-Rhamnoside
Isoquercitrin	-Glucoside

Figure 8: Initial structure of substances that form pigments in white wine (modified according to Singleton et al. 1979; Belitz et al. 2001)

In recent years, it was discovered that the color alteration of white wines inflicted by xanthylum derivatives. These yellow to orange pigments are formed by flavanols linked with a glyoxylic bridge. Although not stable they can affect the color of a wine and they might have an impact on the formation of intermediate products and polymerization (Bührlé et al. 2017).

### 2.1.5 Effects of enological factors on the wine color

The color of wine or must can be influenced by several enological parameters. Studies show that increasing the fermentation temperature leads to a higher color intensity in Shiraz and Semillon wines. Pomace pre-fermentation also increased the color (Reynolds et al. 2001). Fermentation temperature has also an effect on the growth of microorganisms during fermentation, which on their own can alter the color. Due to the higher fermentation temperature, the yeasts produces an increased amount of metabolites that promotes the vitisin production (Asenstorfer, Hayasaka, et al. 2001; Asenstorfer, Lee, et al. 2006; fulcrand et al. 1996). It furthermore increases enzymatic activities regarding glucosidases and pektinases, which can alter the color via enzymatic browning. Unfortunately, microorganisms can impact the color negatively by the adsorption of tannins and other polyphenols to the cell walls, leading to a loss of color (Tofalo et al. 2021). In Pinot Noir, a generally more lighter grape variety, the fermentation temperature was found to be very important for the formation of polymeric pigments. One study investigated the influence of fermentation temperature was in dependence of the storage capability. Pinot Noir grapes fermented at 20°C were initially darker than those fermented at 30°C. However, over a period of 225 days after bottling the wines obtained from grapes fermented at 20°C were lighter than those obtained from grapes fermented at 30°C. After aging, a lower fermentation temperature yielded a more yellow/brown color, while higher fermentation temperature yielded a redder color (Gao et al. 1997).

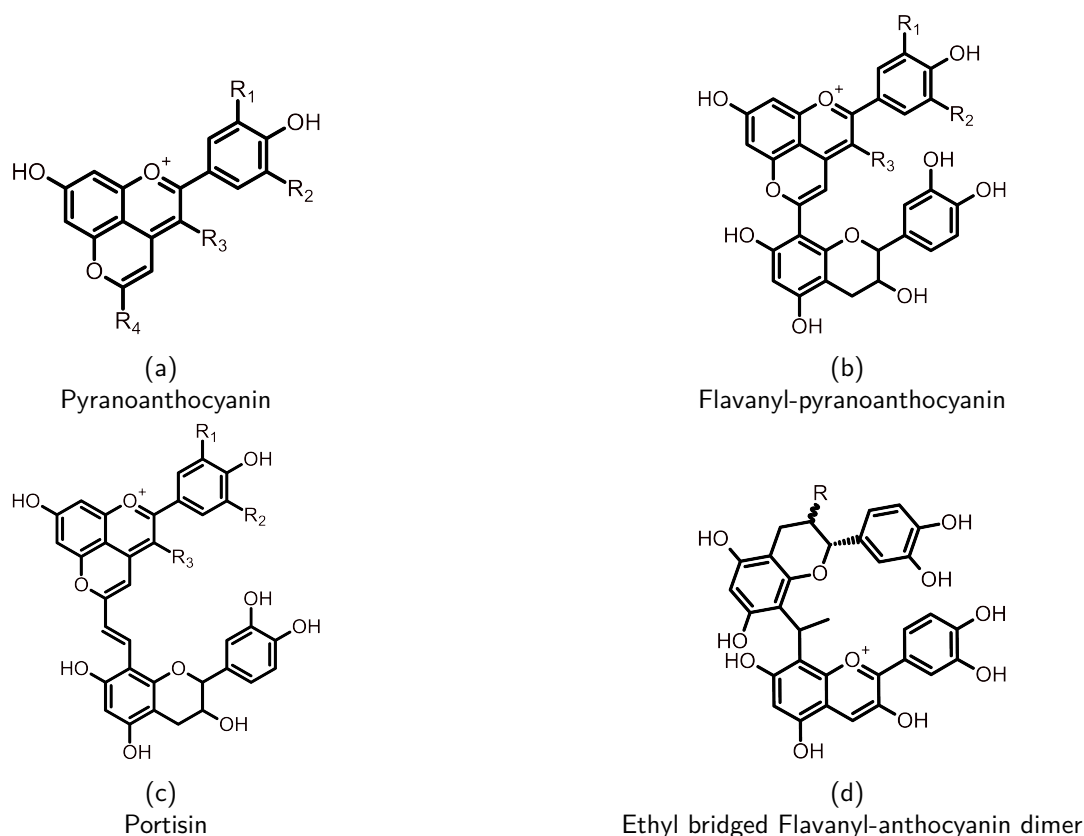


Figure 9: Structures of higher molecular flavonoid polyphenols as a result of aging. Modified according to Asenstorfer, Hayasaka, et al. 2001; Asenstorfer, Lee, et al. 2006 ( $R_1, R_2 = -H, -OH, -OCH_3$ ;  $R_3 = -H, O\text{-Glucose}$ ,  $R_4 = -H$  (Vitisin A),  $-COOH$  (Vitisin B))

Time represents also a notable parameter that changes the wine color. Monomeric anthocyanins are, due to their charge, fairly reactive. With time, numerous products of reactions using monomeric anthocyanins have been identified. The new substances result in an orange/brown color shift that is typical for aged wine. Many of the aging reactions derive from the formation of acetaldehyde, that is built via the Fenton reaction from ethanol (Fenton 1894; Danilewicz 2003). Flavonoid polyphenols and acetaldehyde can react either to a carbocation or to 4-vinyl-catechin. 4-Vinyl-catechin on its own is needed for the formation of pyranoanthocyanins with anthocyanins as reaction partner as well as portisins (Figure 9c). The carbocation derived from the reaction of catechin with acetaldehyde can also react with other flavonoid polyphenols forming ethyl linked dimeric or trimeric pigments (Figure 9d) (Es-Safi et al. 2002). For most wine styles the increase of color intensity is desired. However, this is not the case for all wine styles. e.g. rosé or Blanc de noir wines. These two wine styles are closely related as they derive from red grapes and are produced like a white wine. Contrary to red wine, these wines are not macerated and the fermentation takes place after pressing. The color of rosé wines range from light red, to orange to almost colorless. Blanc de noir, on the other hand, is known for its colorless appearance. In Germany, the color of Blanc de noir is regulated by §32 of the wine regulation act. Herein, the color of Blanc de noir has to be "like a white wine" (Bundesrat Drucksache 175/21 2021). In order to achieve this colorless or light yellow appearance, winemakers are encouraged to use only intact, handpicked berries. However, it is very difficult to hold these high standards during harvest season. Therefore, Blanc de noir production is especially difficult and it is not uncommon that a small amount of color pigments is transferred into the wine. To mitigate the color of Blanc de noir to some extent it is possible to bleach monomeric anthocyanins with sulfur, to be specific  $\text{SO}_2$ .

The reaction takes place at the electrophil centers at  $\text{C}_2$  and  $\text{C}_4$  of the anthocyanin (Figure 10). This disturbs the aromatic system and the molecule loses its color in the process. This type of reaction is not evident with polymeric pigments for sterical reasons.

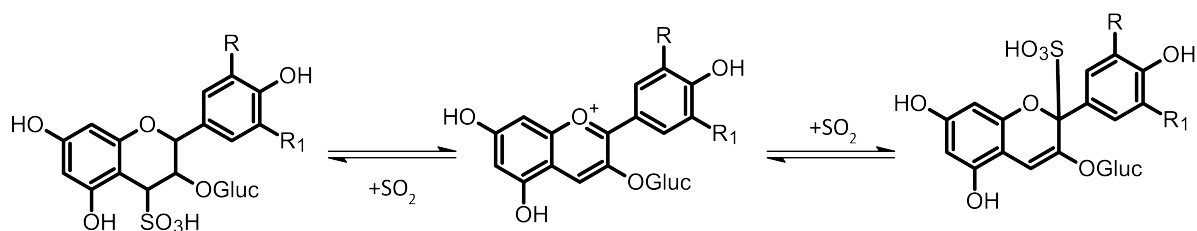


Figure 10: Structure and reaction of anthocyanins with added sulfur dioxide (modified according to Jurd 1964)

The whole chemistry of wine color is extremely complex, as these are only the substances that are identified and quantifiable. However, these do not include oligomeric or polymeric pigments. Therefore, referencing the perception of wine color to the colorants of wine is not target-oriented. It is simply not possible to measure every colorant, as well as to calculate every possible impact of copigments, matrix effects etc. Besides that, the time effort and money cost would not justify the cause. To mitigate these problems a more macroscopic approach is used by winemakers, technicians, and researchers using photometry. The advantage of this is, that matrix effects, copigmentation, and oxidation reactions are included in the measurement, meaning the information is accessible, but not quantifiable.

## 2.2 Introduction about spectrometry, spectroscopy, and photometry

Spectroscopy and spectrometry are often used interchangeably. While both investigate the energy spectrum of a sample, they do not have the same definition. Spectroscopy is normally the evaluation of optical spectra, whereas the term spectrometry is rather used for quantification purposes. Photometry on the other hand is just the measurement method using ultraviolet (UV) and visible light (Vis) (Matissek et al. 2018). Generally, spectroscopy can be divided into atom- and molecular spectroscopy. Molecular spectroscopy is collective term that encompasses UV/Vis-, infrared-, or fluorescence spectroscopy (Gey 2015). The scope of this work focuses on UV/Vis-spectroscopy.

### 2.2.1 Principles of Molecular Spectroscopy

Molecular spectroscopy in UV ( $\lambda = 200 - 380$  nm) and Vis area ( $\lambda = 380 - 800$  nm) of the light spectrum is based on the excitation of electrons from the ground state to excited states between the molecular orbitals. In many cases this transition happens between the highest occupied molecular orbital (HOMO) and the lowest unoccupied molecular orbital (LUMO) by interactions with incident light (Hinderer 2020; Gey 2015). The electron transfer occurs between a binding orbital or an orbital with a pair of electrons and a non-binding or anti-binding electron pair. For transitions needing UV-light the  $\sigma$ -orbitals are essential, as the energy gap between binding and anti-binding orbital ( $\sigma \rightarrow \sigma^*$ ) is relatively big. For colorful molecules especially the transitions between binding  $\pi$ -orbitals and anti-binding  $\pi$ -orbitals ( $\pi \rightarrow \pi^*$ ) and non-binding orbital to anti-binding  $\pi$ -orbital ( $n \rightarrow \pi^*$ ) are especially important, since the energy gap between these orbitals is lower. The conjugation of  $\pi$  electron systems also has an impact on the absorption properties of molecules. Due to the delocalisation of the electrons, the energy gap between binding and anti-binding orbital decreases and the energy of light of the visible spectrum is enough to excite the electrons (Matissek et al. 2018; Gey 2015; Atkins et al. 2022). The wavelength at which the absorption occurs correlates with the distance between the energy states. This allows conclusion about the structural elements of the target molecules in the sample (Kortüm 1962; Williams 1975; Maier 1974). Figure 11 shows the process of absorption or the transition of the electron from the ground state to the excited singlet state. From the excited state the electron falls back to the ground state. The energy loss is in most cases radiation free. The energy is released in form of heat. In some cases the electron can switch between excited singlet states by inter conversion. If the fallback to the ground states coincides with light irradiation and the transition is from an excited singlet state to the ground state, the phenomenon of fluorescence occurs. A far longer luminescence is described with phosphorescence. By intersystem crossing the spin of the excited electrons is transformed from anti-parallel to parallel, resulting in a fallback from excited singlet state to excited triplet state without light irradiation. The fallback from triplet state to ground state causes light irradiation (Atkins et al. 2022; Gey 2015).

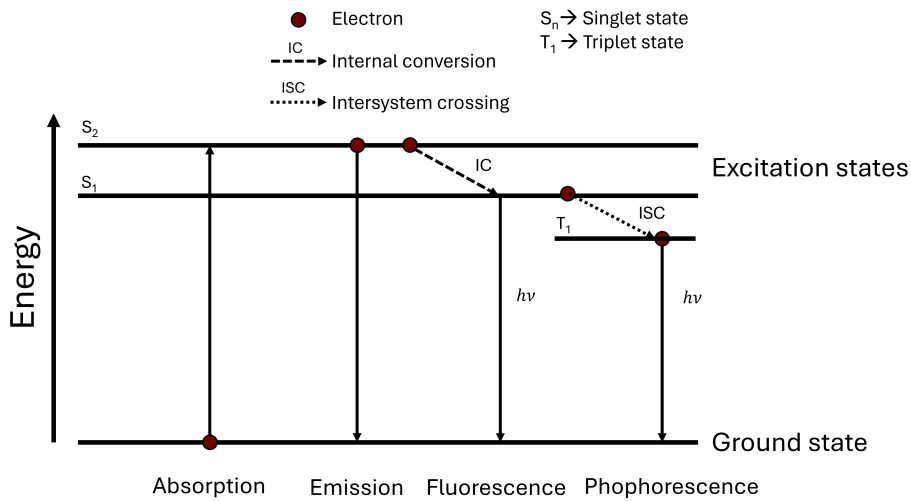


Figure 11: Depiction of the excitation of electrons (absorption), radiation-free transition back from excited singlet state to ground state (emission), fallback from excited singlet state to ground state with emission of radiation after internal conversion (fluorescence), and fallback from excited triplet state to ground state with emission of radiation after intersystem crossing from singlet to triplet state (phosphorescence)

### 2.2.2 Lambert Beers Law

Photometric methods are based on Lambert Beers' law, which can be derived in three steps. In the first step the model has to be constructed. For this, the sample is comparable to volume elements stacked behind each other. The pathlength of a volume element is defined as  $dx$ , whereas the change of intensity is defined as  $dI$ , which is proportional to the pathlength of the volume element, the concentration of the target absorbing species ( $J$ ), as well as the remaining intensity ( $dI \propto c(J)I dx$ ). Due to the fact that the change of intensity  $dI$  is always negative the following expression can be written:

$$dI = -\kappa c(J)I dx \quad (1)$$

By dividing through  $I$  on both sides the expression gets transformed to:

$$\frac{dI}{I} = -\kappa c(J) dx \quad (2)$$

Here,  $\kappa$  is a coefficient of proportionality.

For the next step the absolute reduction of the continuous decrease of initial intensity ( $I_0$ ) when passing through pathlength  $L$  has to be calculated. The sum of an infinitesimal small change is defined as an integral and can be written as:

$$\int_{I_0}^I \frac{dI}{I} = -\kappa \int_0^L c(J) dx \quad (3)$$

On the left side of the equation the following standard integral can be observed:

$$\int_{I_0}^I \frac{dI}{I} = \ln \left( \frac{I}{I_0} \right) \quad (4)$$



Given the condition that the concentration of  $J$  is the same throughout  $L$ . The right side of equation 3 can be written as follows:

$$-\kappa \int_0^L c(J)dx = -\kappa c(J) \int_0^L dx = -\kappa c(J)L \quad (5)$$

In the last step the expression has to be converted from natural logarithm to decadic logarithm with the following operations:

$$\ln\left(\frac{I}{I_0}\right) = \ln(10) \cdot \log\left(\frac{I}{I_0}\right) \quad (6)$$

$$\log\left(\frac{I}{I_0}\right) = -\frac{\kappa c(J)L}{\ln(10)} \quad (7)$$

$$\log\left(\frac{I}{I_0}\right) = -\epsilon c(J)L \quad (8)$$

Solving the expression with respect to  $I$  leads to what is commonly known as Lambert Beers law:

$$I = I_0 \cdot 10^{-\epsilon c(J)L} \quad (9)$$

According to Lambert Beers's law the amount of light reaching the detector, known as the transmission is the quotient of intensity and initial intensity and the absorbance is defined as the decadic logarithm of the transmission (Atkins et al. 2022; Matissek et al. 2018).

$$T = \frac{I}{I_0} \quad (10)$$

$$A = \log(T) \quad (11)$$

### 2.2.3 Photometer systems

Photometers differ in many ways, but the basic principle is the same. Photometers consist of at least one light source, which is usually an illuminant with a very broad spectral power distribution, a monochromator unit, sample mount, and a detector unit. With a single beam photometer the light only passes the sample directly to the detector unit, whereas the light is split up between a reference cuvette and a sample cuvette with a double beam photometer. After passing the reference and the sample cuvette the rays are combined again and are guided to the detector unit (Figure 12). The advantage of a double beam photometer is, that the solvent can be placed in the reference sample mount, eliminating the need for a blank measurement with the solvent.

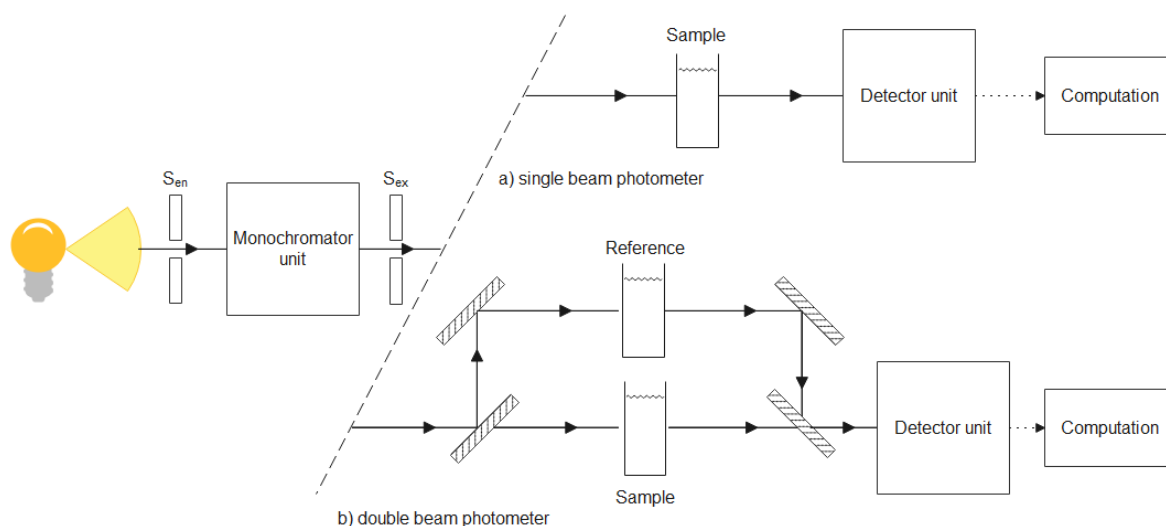


Figure 12: Optical arrangement of single beam and double beam photometers (modified according to Matissek et al. 2018)

#### 2.2.4 Monochromator systems

In order to obtain exact results the emitted light is decomposed by a prism or a grating into single wavelengths. A typical prism mounting is the Wadsworth prism mounting where the light is collimated at a concave mirror, reflected by a planar mirror, and guided through a prism. The deconstructed light of a single wavelength is then collimated again at a second concave mirror and leaves the construction (Czerny and Turner 1930). The disadvantage of the Wadsworth prism mount is that it shows an increased amount of astigmatism, as the reflection at the Wadsworth mirror causes the inside and outside of the wave to be exchanged (Czerny and Turner 1930). Astigmatism is an optical aberration that occurs when an oblique incident beam is refracted to different degrees in meridional and sagittal plane. This leads to two different focal points of the light beam (F. E. Wright 1911). A better and more modern solution to prism mountings are monochromators with diffraction gratings, like the Rowland and Czerny-Turner monochromators. Monochromators using a Rowland circle mitigate astigmatism by dispensing concave mirrors completely. Instead collimation and decomposition of the light is both achieved by diffraction gratings. Czerny-Turners monochromator on the other hand use a second concave mirror to reverse the error of the first concave mirror without interchanging of the waves as it is the case with the Wadsworth mounting (Czerny and Turner 1930; More 2005).

#### 2.2.5 Detector Systems

The biggest difference between photometers is their detector unit, where different systems are capable to detect the incoming light. One of the most sensitive detector systems are photomultiplier tubes (PMT). PMT's work on the premise of the extrinsic photoelectrical effect, where photons transfer their energy onto electrons. Part of the energy is needed to release the electron from the cathode material. The remaining energy of the photon is transferred onto the electron that moves towards the anode as kinetic energy. The initial electron collides with dynodes releasing secondary electrons. These create an avalanche reaction due to the repeated collision of electrons with the dynodes, thus amplifying the signal (Figure 13).

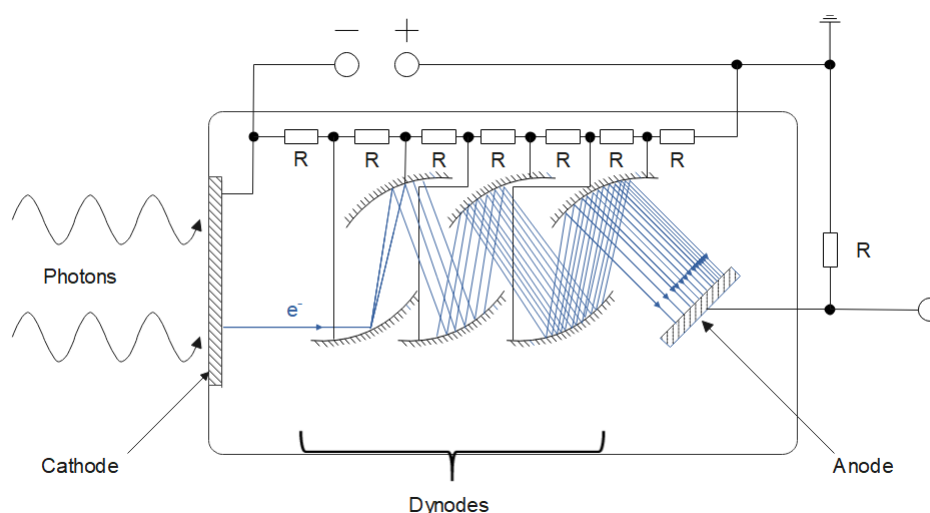


Figure 13: Schematic assembly and measurement mechanism of photomultiplier tubes (modified according to Löffler-Mang 2012; Shimadzu 2023)

Photodiodes on the other hand are doped semiconductors, capable of producing a signal dependent on incident light. The electrical conductivity of semiconductors can be explained by the electronic band structure. In a conductor, e.g. metals, the orbitals of the valence band are always overlap with the next higher energy band. When a voltage is applied to the conductor electrons can move to the next higher energy band without restriction. In contrast, the valence energy band and the conductive band of the isolator is separated by a band gap, that prohibits electron movement. If the band gap is small enough, semi-conductivity is reached. An example for an isolator is the diamond with a fully packed valence energy band and a band gap of 5 eV. The valence band of silicon, a substance that crystallizes homogeneously to the diamond (Figure 14a), is separated from the conductive band by a band gap of 1.1 eV. Electrons can move from the valence band to the conductive band with increasing temperature. Due to the absence of the electrons, a positive charge is building up in the valence band, creating so called holes, which then move in the opposite direction. The silicon lattice shown in Figure 14a can be spiked with impurities that alter the conductivity. For example, the elements of the 15<sup>th</sup> group contain one more electron than the silicon and can be spiked into the silicon lattice (Figure 14c). In the energy band model the energy levels of these donor elements lay in the band gap right under the conductive band, creating an n-type semiconductor. If the silicon lattice is spiked with elements of the 13<sup>th</sup> group a p-type semiconductor is created (Figure 14b). These elements contain one electron less than the silicon and therefore the remaining electron bonds with an electron from a neighboring silicon atom. The creation of these defect electrons are the reason for the conductivity of these p-type semiconductors. The energy level of the acceptor elements is right above the valence band of the silicon. Electrons from the valence band can move up to the energy level of the acceptor with very little energy supply. In conclusion, n-type semiconductors have positive stationary charge and movable negative charge carriers (electrons), whereas p-type semiconductors contain a stationary negative charge and movable positive charge carriers (holes)(Riedel and Janiak 2022). If a p-type semiconductor and a n-type semiconductor come into contact, the free electron moves from the n-zone into the p-zone and holes move from the p-zone to the n-zone. By recombination the electrons are filling up the holes and the charge carriers disappear in the boundary region. Near the contact zone the stationary charges remain and are responsible for the fact that not all electrons in the n-zone fill with all holes in the p-zone. By repulsive force the remaining mobile charge carriers are prevented from entering the other zone. Because of this the so called space

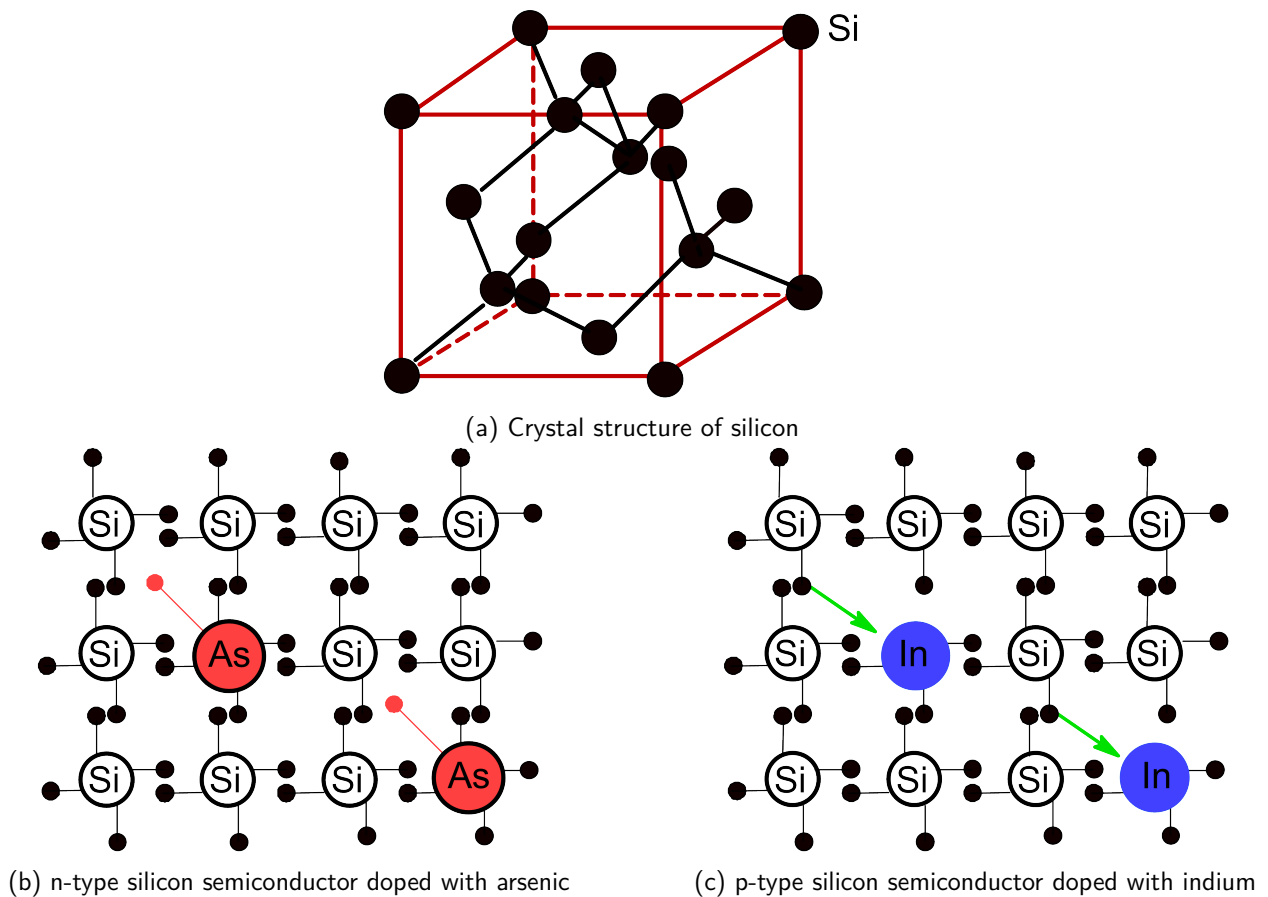


Figure 14: Three dimensional crystal structure of silicon as well as a p-type and n-type silicon semiconductor (modified according to Riedel and Janiak 2022)

charge zone is created, in which there are no mobile charge carriers present. The resulting pn-layer is non-conductive without external electrical impulses. There are two ways to introduce an external voltage. Forward-bias wiring, which is used in light emitting diodes (LED), introduces new mobile charge carriers on the positive pole of the pn-junction, forcing positive charge carriers into the space charge zone. The positive charge carriers then attracts a negative mobile charge carrier and recombination occurs, which then ultimately leads to a current and the emitting of photons. The second way is the reverse-bias wiring. By illumination, additional electron/hole pairs are created due to the introduced energy, which then flow off in the direction of the external voltage, leading to a current that is proportional to the irradiated light. The pn-junction is now working as a pn-photodiode. By applying a voltage in reverse direction, the space charge zone widens and the field strength in the space charge zone increases. The maximal strength of the electric field must be under the dielectric strength or a complete electrical breakdown occurs. The dielectric strength is limited by the capabilities of the semiconductor material, e.g. silicon has a dielectric strength of 25kV/mm. To increase the dielectric strength of a photodiode an intrinsic layer can be introduced between the p- and the n-zone (Figure: 15a and 15b). The intrinsic layer inherits an intrinsic conduction and due to manufacturing processes it is usually doped with a small amount of negative mobile charge carriers. Not only does this progress increase the dielectric strength, it also leads to a quicker response time (Löffler-Mang 2012; Specovius 2010).

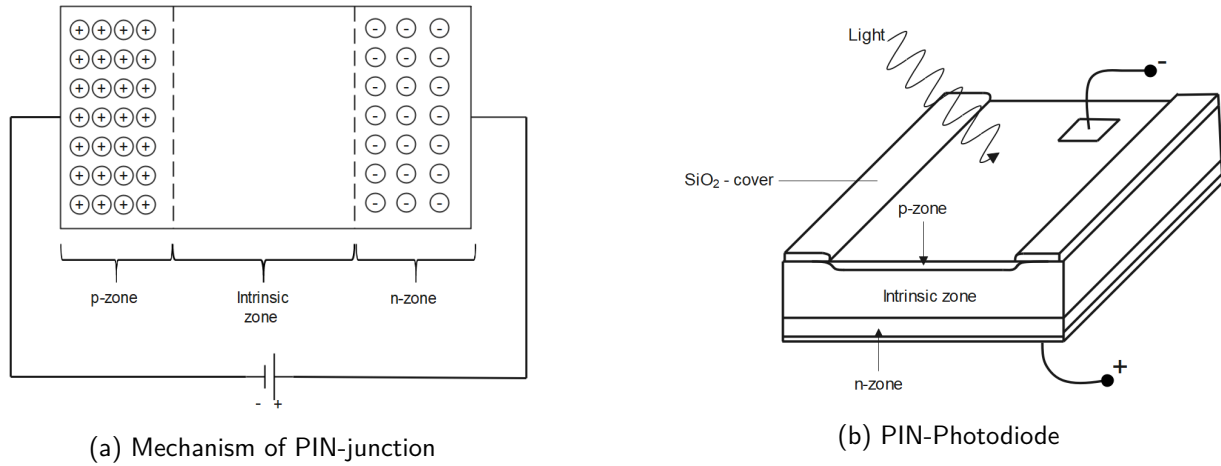


Figure 15: Schematic assembly and mechanism of a PIN-junction in a PIN-photodiode

So, photomultiplier and photodiodes have the same purpose, but with very different modes of operations. The photomultiplier tube relies on the extrinsic photoelectric effect and signal amplification, leading to a huge sensitivity when the intensity of the incident light is very small. The pn- or PIN-photodiode does not amplify the signal. For this purpose the avalanche photodiode was created. However, despite amplification the photodiodes are always less sensitive than the photomultiplier (Löffler-Mang 2012). The advantage of the photodiodes is lower maintenance, a better quantum efficiency as well as the insensitivity to magnetic fields. In the context of this work, it is unknown if the monochromator and detector system have an influence of the measurement of the wine color.

## 2.3 The Quantification of Wine Color

### 2.3.1 Theory of colors

The focus of colorimetry or the theory of colors describes the connection between a light spectrum and a color stimulus or perceived color. Color stimuli are defined as radiation that leads to a direct stimulation of the retina and results in a perception of color. A color stimulus can be triggered in various ways. If an object is an illuminant the color stimulus can be triggered directly. Therefore, the perceived color contains information regarding the light source. Another indirect trigger for the color stimulus is via reflection or transmittance. The result of the combination of different colors can be explained by the three Grassmann's laws. The first Grassmann law states that the color of an object can be completely depicted by three variables (Equation 12).

$$C = \begin{pmatrix} R \\ G \\ B \end{pmatrix} \quad (12)$$

The additive combination of different colors, or the second Grassmann law, can be displayed unambiguously by the linear combination of three primary valences. The additive color addition follows the rules of vector addition (Lübbe 2013; Bühler et al. 2018).

$$C_1 + C_2 = \begin{pmatrix} R_1 + R_2 \\ G_1 + G_2 \\ B_1 + B_2 \end{pmatrix} \quad (13)$$

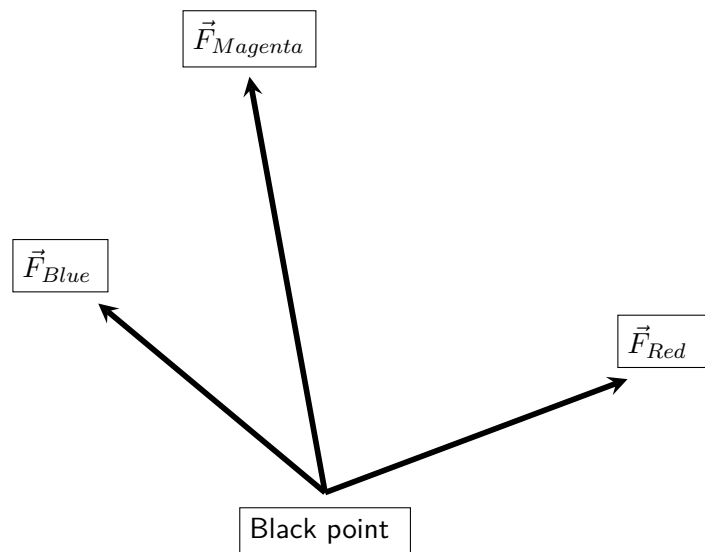


Figure 16: Depiction of the vector combination regarding the additive color mixing (modified according to Lübke 2013)

For example, the linear combination of one vector, representing the blue content, with a vector, that represents a red content, results in a combined magenta vector. The length of the vector corresponds to the color value (Figure 16). Additive combination of color results in an increase of light radiation, leading to white. Another option to obtain different colors is subtractive color combination. Due to absorbance and light scattering the lightness of the new color decreases leading ultimately to black (Figure 17).

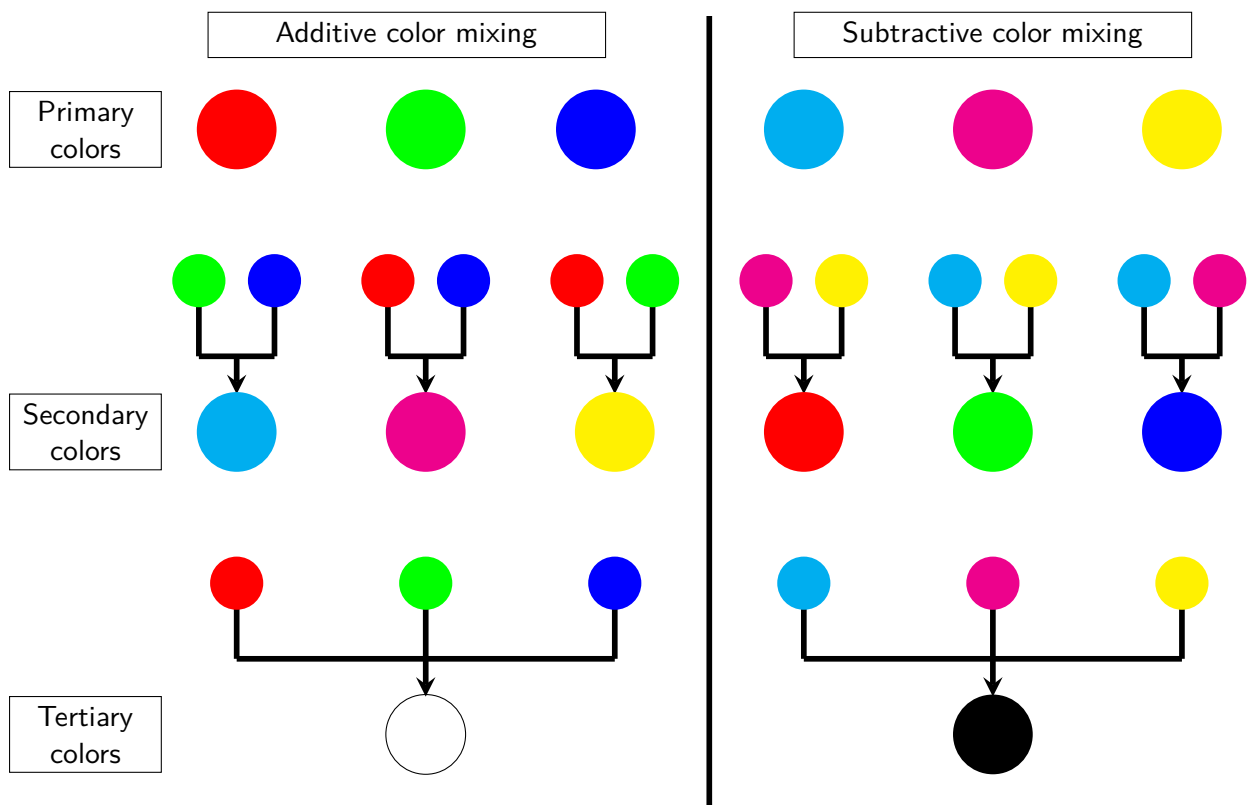


Figure 17: Comparison of the additive color combination and the subtractive color combination (modified according to Lübke 2013; Bühler et al. 2018)

The last Grassman law states that hue of a color addition depends only on the color impression of the underlying original colors but not on the spectral composition, meaning that two metameric color stimuli can be described completely by the hue without knowing the spectral composition.

### 2.3.2 Glories' color evaluation

Glories' color evaluation is a well established and recommended method to quantify the color of a wine (OIV 2021). It involves measuring the optical density of a wine at three different wavelengths. To capture the yellow component of a wine the optical density at 420 nm is measured. To capture the red and blue component of a wine the optical density at 520 and 620 nm is measured. The measurement is set to take place in a 1 mm cuvette and, due to copigmentation, is only applicable to undiluted wine. Starting from the optical densities other parameters like the color intensity (CI), the hue (T), and the *Brilliance* (dA(%)) can be calculated. The color intensity is a parameter indicating the lightness of a wine and can be calculated by summation of the three optical densities (Equation 14). The CI value can vary between 0.3 and 1.8 depending on grape variety, maturity, origin, or enological parameters.

$$CI = OD420 + OD520 + OD620 \quad (14)$$

T is a parameter that indicates the orange color development of a wine and is calculated by the quotient of OD420 and OD520 (Equation 15). It is usually used as a maturity parameter for red wine as it ranges between 0.5 and 0.7 for young wines, whereas mature wines can reach a tint up to 1.3.

$$T = \frac{OD420}{OD520} \quad (15)$$

The parameter *Brilliance* is only usable in red wine as it calculates the percentage of the red color in a wine (Equation 16). It is also used as a maturity parameter as dA(%) should range from 40 – 60 for young wines (Ribéreau-Gayon 2006).

$$dA(\%) = \left(1 - \frac{OD420 + OD620}{2 \cdot OD520}\right) \cdot 100 \quad (16)$$

### 2.3.3 CIE Color Spaces

While Glories' color measurement is specifically used for wine, the CIE color spaces can be used for color measurement in general. Like Glories' method, the CIE color spaces are recommended by the OIV for wine (OIV 2021). The CIE color spaces rely on the idea of the additive color. In 1861, Maxwell provided proof of Grassmann's theory by experimentally demonstrating that every color can be reproduced approximately by light of three wavelengths. Maxwell introduced the initial color space that is now known as RGB color triangle (Schanda 2007). Herein, the primary colors red, green, and blue are placed in the corners of an isosceles triangle. The light intensity decreases evenly until it reaches the zero point of its intensity on the opposite side. However, there appears a decisive disadvantage: although the positions of the spectral colors are correct, the midpoint between blue and green is not saturated enough to represent spectral cyan. This leads to the conclusion that there are colors existing outside the RGB triangle, that can only be reproduced by adding imaginary colors according to Grassmann's law, e.g. negative red to achieve spectral cyan. To mitigate these problems and introduce a color space

that is more coherent with the visual perception the CIE published in 1931 the CIE XYZ color space. This color space depends on the transmission spectrum of the object, in this case wine, the spectral power distribution of a standard illuminant, and the so called color matching function of the standard observer. The perception of color is very dependent on the illumination. If a light source contains a high amount of blue, the object would also have a slight blue tint. The CIE uses multiple standardized light sources, the best practice here is the D65 standard illuminant, which represents natural daylight on a sunny day. Other standard illuminants are the D50 standard illuminant, which represents natural daylight on a cloudy day, and the standard illuminant A. This light source is similar to the commonly used tungsten filament in light bulbs (Figure 18). The color temperatures ranges from 6500 K of D65 standard illuminant to 2856 K, a higher amount of red color irradiance, of the standard illuminant A. The D50 standard illuminant irradiates a slightly less blue light than the D65 standard illuminant, yielding a color temperature of 5000K (Schanda 2007; CIE 2021). The basis of the standard observer are the

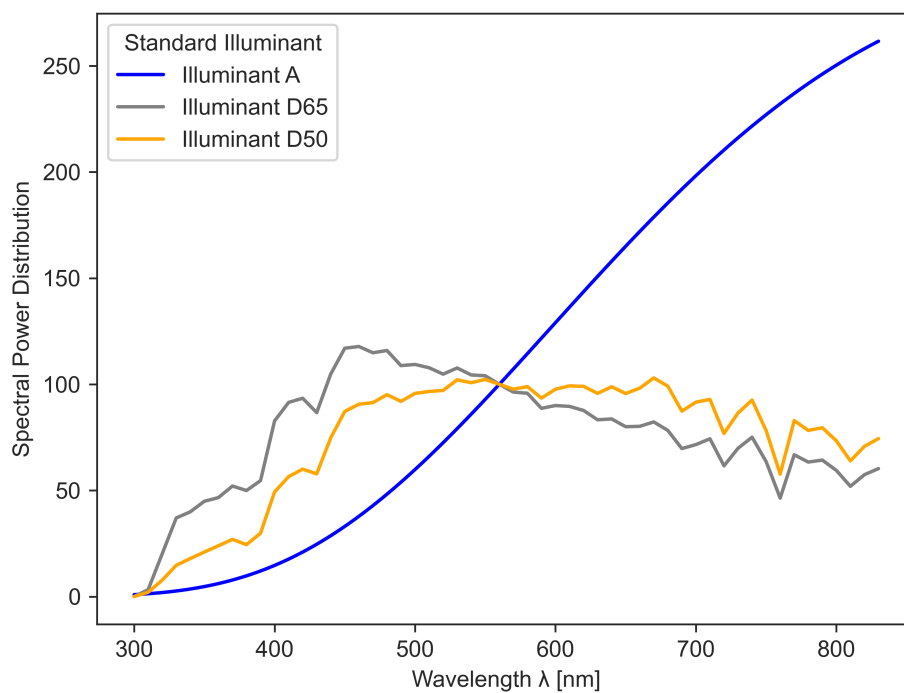
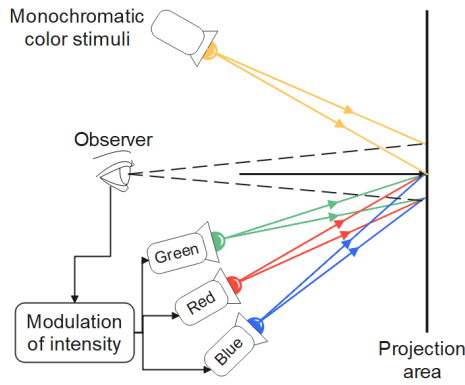


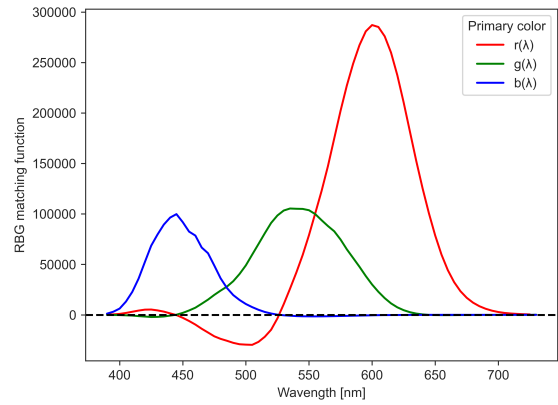
Figure 18: Spectral power distribution of the standard illuminant A, D50, and D65 (modified according to CIE 2019c)

experiments conducted by Wright and Guild. A panel of 17 individuals with normal color vision were asked to recreate a given color stimulus by adjusting the intensity of a red (700 nm), green (546.1 nm), and blue (435.1 nm) illuminant (Figure 19a)(W. D. Wright 1929; Guild J. 1931).





(a) Experimental setup to obtain the color matching function



(b) Resulting color matching function of the CIE RGB system

Figure 19: Schematic depiction of the original experiment of Wright and Guild and the resulting color matching function that contains the imaginary negative parts (W. D. Wright 1929; Guild J. 1931).

This resulted in the so-called color matching function of the CIE RGB system (Figure 19b). To mitigate the need for negative color values the CIE transformed the color matching function of the CIE RGB system to obtain the color matching function for the imaginary primary valences X, Y, and Z, which is also known as the CIE 1931 2°-standard observer (Figure 20). This standard observer represents the averaged vision of a person with normal color vision in a field of view of 1 to 4°, as a 2° visual field lays within Maxwell’s spot, where the density of the *macula lutea* is the most constant. A field of view of 1 to 4° is equal to an object with the dimension of 17 mm seen from a distance of 50 cm. Because of the development of the D65 standard illuminant, the CIE 1964 10°-standard observer was introduced to represent a field of view of >4°. This corresponds to viewing an object with a diameter of approximately 90 mm from a distance of 50 cm. The CIE 1964 10°-standard observer was derived from the experimental works of Stiles and Burch, who recreated the original experiments of Wright and Guild with 67 observers and a range of stimuli between 390 and 830 nm (Figure 20)(Schanda 2007; CIE 2019c; CIE 2021; Stiles and Burch 1959).

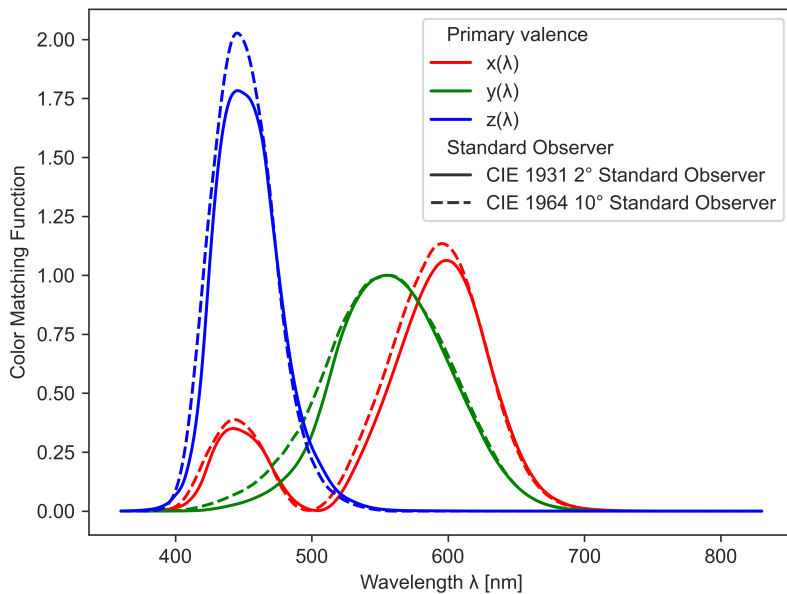


Figure 20: Color matching function of the CIE XYZ color space (CIE 2019a).

To calculate the tristimulus values  $X$ ,  $Y$ , and  $Z$  of a color stimulus, the color stimulus function  $\phi(\lambda)$  has to be calculated with the following equation:

$$\phi(\lambda) = \tau(\lambda) \cdot S(\lambda) \quad (17)$$

with:

$\phi(\lambda)$  = Color stimulus function at  $\lambda$   
 $\tau(\lambda)$  = Transmission at  $\lambda$   
 $S(\lambda)$  = Spectral power distribution at  $\lambda$

The tristimulus values are defined as the area under curve of the color stimulus function multiplied with the corresponding color matching function. To obtain the tristimulus values from experimental data from a photometer a numeric integration has to be performed, with  $\Delta\lambda$  being the data interval set during the measurement of the spectra (CIE 2019a).

$$X = k \cdot \int_{380nm}^{780nm} \phi(\lambda) \cdot \bar{x}(\lambda) d\lambda = k \cdot \sum_{380nm}^{780nm} \phi(\lambda) \cdot \bar{x}(\lambda) \cdot \Delta\lambda \quad (18)$$

$$Y = k \cdot \int_{380nm}^{780nm} \phi(\lambda) \cdot \bar{y}(\lambda) d\lambda = k \cdot \sum_{380nm}^{780nm} \phi(\lambda) \cdot \bar{y}(\lambda) \cdot \Delta\lambda \quad (19)$$

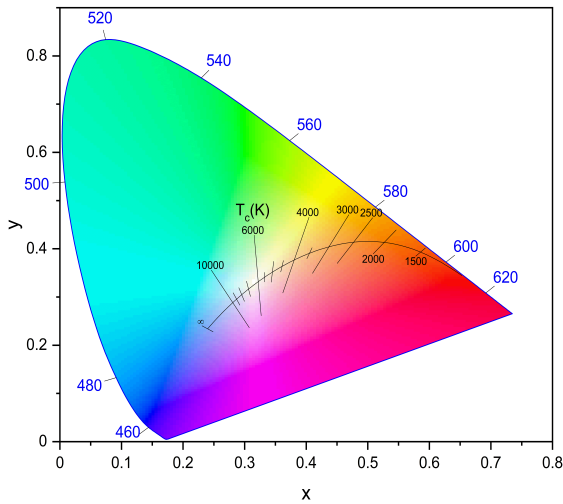
$$Z = k \cdot \int_{380nm}^{780nm} \phi(\lambda) \cdot \bar{z}(\lambda) d\lambda = k \cdot \sum_{380nm}^{780nm} \phi(\lambda) \cdot \bar{z}(\lambda) \times \Delta\lambda \quad (20)$$

The factor  $k$  (equation 21) here is a normalization factor and defined as the sum product of the spectral power distribution and the color matching function  $\bar{y}(\lambda)$

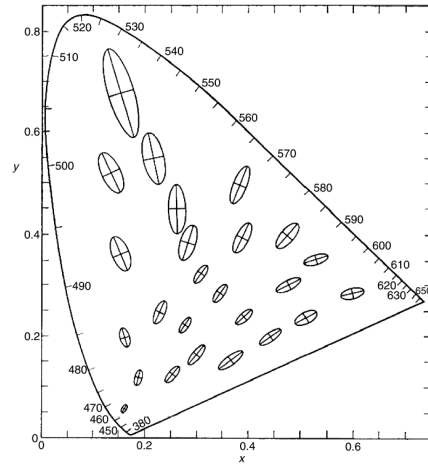
$$k = \sum_{380nm}^{780nm} S(\lambda) \cdot \bar{y}(\lambda) \cdot \Delta\lambda \quad (21)$$

with:

$\phi(\lambda)$  = Color stimulus function at  $\lambda$   
 $\bar{x}(\lambda)$ ,  $\bar{y}(\lambda)$ ,  $\bar{z}(\lambda)$  = Color matching functions at  $\lambda$   
 $S(\lambda)$  = Spectral power distribution of the standard illuminant at  $\lambda$   
 $d\lambda$ ,  $\Delta\lambda$  = Wavelength interval used to record the spectrum photometrically.



(a) CIE Chromaticity diagram derived from CIE XYZ values (CIE 2019b)



(b) MacAdam ellipses ten times enlarged (Judd and Wyszecki 1963; MacAdam 1942)

Figure 21: CIE Chromaticity diagram according to CIE and the irregularities discovered by MacAdam

One major requirement of the CIE XYZ color space was to develop an equidistant color space (Figure 21a), where mathematical color differences reflect the visual color difference. However, MacAdam could prove, that the CIE XYZ color space is not entirely equidistant (MacAdam 1942). The so called MacAdam ellipses represent areas in the CIE XYZ color space where mathematical color differences are not visible (Figure 21b). The color ellipses are bigger in high saturated areas and grow smaller the closer the color is to the origin point. These findings led the CIE to the development of the CIE  $L^*a^*b^*$  color space. A three-dimensional Cartesian coordinate system, where the  $L^*$  coordinate represent the lightness of an object ( $L^* \in \mathbb{R} \mid L^* = [0, 100]$ ), the  $a^*$  coordinate either the red color component ( $a^* \in \mathbb{R}^+$ ) or the green component ( $a^* \in \mathbb{R}^-$ ) of an object, and the  $b^*$  coordinate either the yellow ( $b^* \in \mathbb{R}^+$ ) or blue ( $b^* \in \mathbb{R}^-$ ) color component of an object (Figure 22). Using a non-linear transformation seen in equations 22 – 24 the CIE  $L^*a^*b^*$  coordinates can be calculated from the CIE XYZ tristimulus values (Schanda 2007; CIE 2019b).

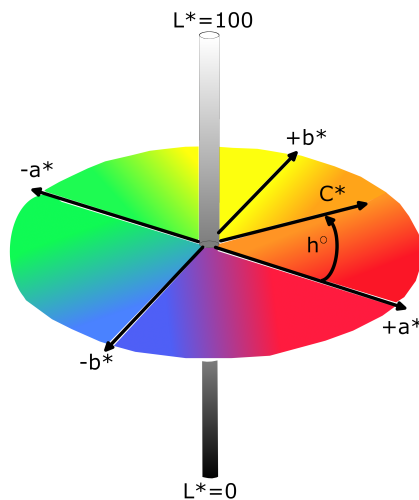


Figure 22: Depiction of the CIE  $L^*a^*b^*$  color space, where  $L^*$  represents the lightness,  $a^*$  either red ( $a^* \in \mathbb{R}^+$ ) or green ( $a^* \in \mathbb{R}^-$ ), and  $b^*$  either yellow ( $b^* \in \mathbb{R}^+$ ) or blue ( $b^* \in \mathbb{R}^-$ )

$$L^* = 116 \cdot f\left(\frac{Y}{Y_N}\right) \quad (22)$$

$$a^* = 500 \cdot \left( f\left(\frac{X}{X_N}\right) - f\left(\frac{Y}{Y_N}\right) \right) \quad (23)$$

$$b^* = 200 \cdot \left( f\left(\frac{Y}{Y_N}\right) - f\left(\frac{Z}{Z_N}\right) \right) \quad (24)$$

where

$$f\left(\frac{X}{X_N}\right); f\left(\frac{Y}{Y_N}\right); f\left(\frac{Z}{Z_N}\right) = f\left(\frac{P}{P_N}\right) = \begin{cases} \left(\frac{P}{P_N}\right)^{\frac{1}{3}}; \left(\frac{P}{P_N}\right) > \left(\frac{6}{29}\right)^3 \\ \left(\frac{841}{108}\right) \cdot \left(\frac{P}{P_N}\right) + \left(\frac{4}{29}\right); \left(\frac{P}{P_N}\right) < \left(\frac{6}{29}\right)^3 \end{cases} \quad (25)$$

and the ideal values of X,Y, and Z with total reflection:

P =Placeholder for X,Y, and Z

$X_N=94.825$

$Y_N=100$

$Z_N=107.381$

Next to the Cartesian projection of the CIE L\*a\*b\* color space, another commonly used variation is the CIE L\*C\*h<sup>0</sup> color space that uses polar coordinates. Herein, the L\* coordinate is the same as in the CIE L\*a\*b\* color space and also represents the lightness of an object. The C\* coordinate or *Chroma* is a color vector that represents the saturation and the h<sup>0</sup> coordinate is the so called hue angle. At 0° and 360° the hue angle is red. The other colors can be represented as follows: 90° represents yellow, 180° represents green, and 270° represents blue. Chroma and hue angle can be calculated with the following equations (CIE 2019b):

$$C^* = \sqrt{(a^*)^2 + (b^*)^2} \quad (26)$$

$$h^0 = \arctan\left(\frac{b^*}{a^*}\right) \quad (27)$$

## 2.4 Calculation of color distances

One major advantage of the CIE L\*a\*b\* color space over the color measurement method of Glories' is the ability to calculate the distance between two colors. For this in 1976, the CIE used the Euclidean color distance as at this time the CIE L\*a\*b\* color space was thought to be completely equidistant (CIE 2019c).

### 2.4.1 Euclidean color difference

$$\Delta E = \sqrt{(\Delta L^*)^2 + (\Delta a^*)^2 + (\Delta b^*)^2} \quad (28)$$

where:

$$\Delta L^* = L_2^* - L_1^* \quad (29)$$

$$\Delta a^* = a_2^* - a_1^* \quad (30)$$

$$\Delta b^* = b_2^* - b_1^* \quad (31)$$

### 2.4.2 Development of the CMC(l:c) Color Distance Formula

With its introduction, the CIE L\*a\*b\* color space was rigorously tested in the textile industry, where inhomogenities were discovered (McLaren and Rigg 1976; McLaren 1980; Kuehni 1976). The irregularities show that the contours of equal tolerance were rather ellipsoids than spherical. Changes were applied accordingly, leading to the color difference formula presented by the Color Measurement Committee (CMC) of the Society of Dyers and Colorists (SDC). The CMC(l:c) color difference formula introduced weighting functions ( $S_L$ ,  $S_C$ ,  $S_H$ ) to the CIE L\*a\*b\* color space to mitigate the inhomogenities. (Clarke et al. 1984; Luo and Rigg 1986).

**CMC( $l : c$ ) color distance formula**

$$\Delta E_{CMC} = \sqrt{\left(\frac{\Delta L_{ab}^*}{l S_L}\right)^2 + \left(\frac{\Delta C_{ab}^*}{c S_L}\right)^2 + \left(\frac{\Delta H_{ab}^*}{S_H}\right)^2} \quad (32)$$

with:

- $\Delta L_{ab}^*$  =CMC lightness difference  
 $\Delta C_{ab}^*$  =CMC chroma difference  
 $\Delta H_{ab}^*$  =CMC hue angle difference  
 $S_L, S_C, S_H$  =CMC weighting functions  
 $l, c$  =CMC parametric factors

**CMC weighting functions**

$$S_L = \frac{0.04975 L_1^*}{(1 + 0.01765 L_1^*)} \quad (33)$$

unless  $L^* < 16$ , when  $S_L = 0.511$ 

$$S_C = \frac{0.0638 C_{ab,1}^*}{(1 + 0.0131 C_{ab,1}^*) + 0.638} \quad (34)$$

$$S_H = S_C [(0.36 + |0.4 \cos(h_{ab,1} + 35))| \left( \sqrt{\frac{(C_{ab,1}^*)^4}{(C_{ab,1}^*)^4 + 1900}} \right) - (0.36 + |0.4 \cos(h_{ab,1} + 35)))] \quad (35)$$

The weighting function  $S_L$  reduces the effect of lightness differences with increasing lightness beyond  $L = 16$ , whereas the weighting function  $S_C$  reduces the effect of chroma differences. This is needed as the size of the irregularities in the  $a^*b^*$ -color plane increases with higher chroma. The weighting function  $S_H$  is more complex than the other two. The reason for this is the general size-dependence between chroma and the hue angle. The parametric factors  $l$  and  $c$  control the relative sensitivities for lightness and chroma. For textiles a ratio of 2:1 is defined as standard (Schanda 2007).

**2.4.3 Development of the CIE94 Color Distance Formula**

The irregularities in the CIE color spaces led the CIE to the definition of a new color distance function. THE CIE94 color distance formula followed the general form of the CMC( $l:c$ ) (Equation 32):

**CIE94 color distance formula**

$$\Delta E_{94}^* = \sqrt{\left(\frac{\Delta L_{ab}^*}{k_L S_L}\right)^2 + \left(\frac{\Delta C_{ab}^*}{k_C S_C}\right)^2 + \left(\frac{\Delta H_{ab}^*}{k_H S_H}\right)^2} \quad (36)$$

with:

 $\Delta L_{ab}^*$  =CIE94 lightness difference $\Delta C_{ab}^*$  =CIE94 chroma difference $\Delta H_{ab}^*$  =CIE94 hue angle difference $S_L, S_C, S_H$ =CIE94 weighting functions $k_L, k_C, k_H$  =CIE94 parametric factors**CIE94 weighting function**

$$S_L = 1 \quad (37)$$

$$S_C = 1 + 0.045C_{ab}^* \quad (38)$$

$$S_H = 1 + 0.015C_{ab}^* \quad (39)$$

$k_L, k_C, k_H$  are used as parametric factors that can be changed by users to comply with the CIE reference conditions and are normally set to 1. The weighting functions  $S_L$  was set to 1 as the results of the experimental studies leading to the CIE94 formula did not show a difference in the calculated color distance and the perceived color difference (Schanda 2007). The weighting functions  $S_C$ , and  $S_H$  are thought to cope with the different effects of increasing chroma on the perceived color difference. Rigg 1995 concluded that the CIE94 color distance formula yields similar results to the CMC color distance formula (Rigg 1995), whereas other studies showed the opposite (McDonald and Smith 1995; Oglesby 1995; Heggie et al. 1996). However, both formulae were not able to depict color differences in the yellow color area.

**2.4.4 Development of the CIEDE2000 Color Distance Formula**

Later, Luo and Rigg showed that previous color difference datasets were flawed and therefore re-evaluated the results of the previous experiments (Luo and Rigg 1986). For that, they plotted the experimental color discrimination datasets in the  $a^*b^*$  color plane (Figure 23) and found ellipses that were similar to the MacAdam ellipses seen in Figure 21b. Two other observations could be obtained from this projection. First, the size of the ellipses is the smallest near the origin point of the Cartesian coordinate system ( $a^* = 0, b^* = 0$ ) and appear to be larger with increasing chroma. Second, the orientation of the main axis of the ellipses is pointed approximately to the origin point, with exception of the blue color region, where  $b^* \in \mathbb{R}^-$  and  $a^* \in \mathbb{R}^+$  is close to 0.

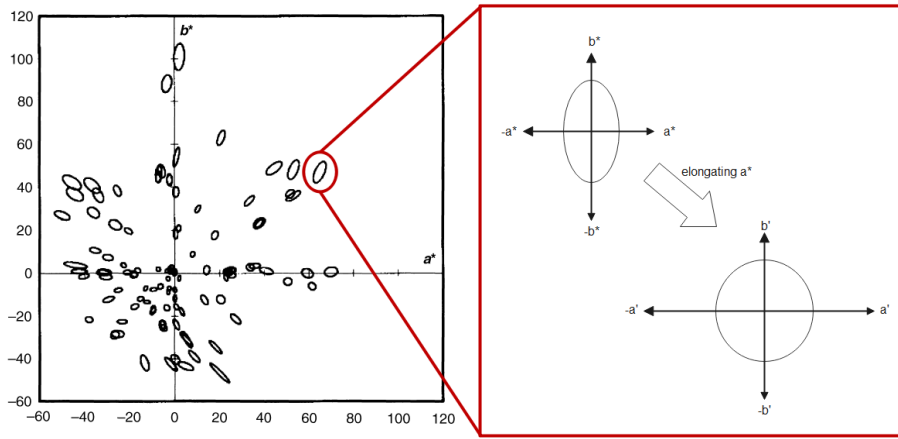


Figure 23: Color ellipses plotted by Luo and Rigg in the  $a^*b^*$  color plane from experimental data (modified according to Schanda 2007; Luo and Rigg 1986; Luo, Cui, et al. 2001).

To cope with the irregularities new color coordinates  $L'$ ,  $a'$ , and  $b'$  are calculated from the existing CIE  $L^*a^*b^*$  calculation. Herein, the  $L^*$  and  $b^*$  underlie no further transformation (Equation 41 and 42). However, to transform the near-achromatic ellipses in the CIE  $L^*a^*b^*$  color space the  $a^*$  coordinate is elongated with equation 43 (Figure 23). The effect of the transformation decreases with higher Chroma and has no impact when  $C_{am}^* \approx [30, 40]$  (Figure 24).

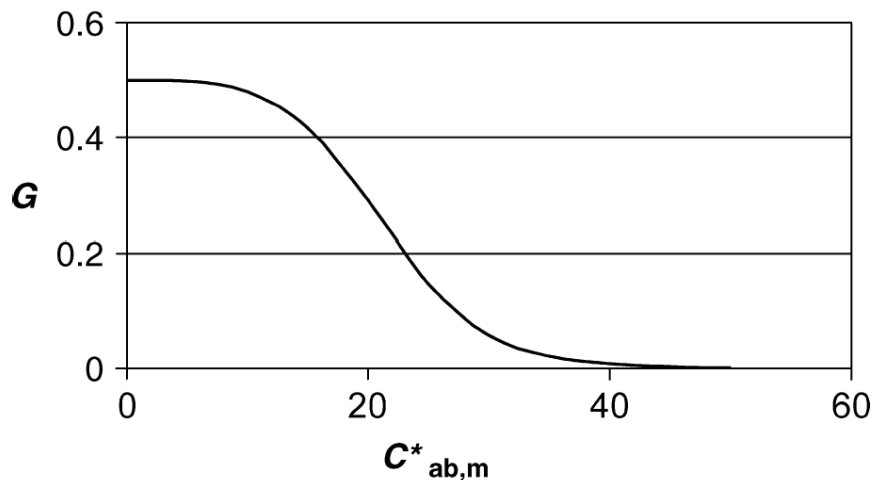
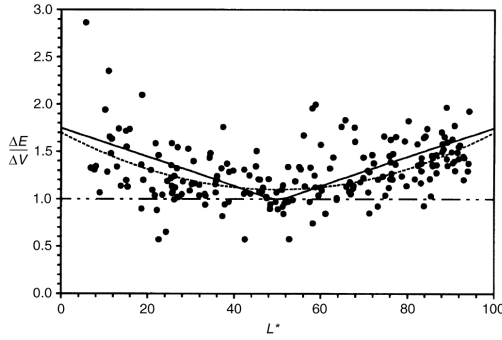


Figure 24: Dependence of  $G$  with increasing chroma (from Schanda 2007)

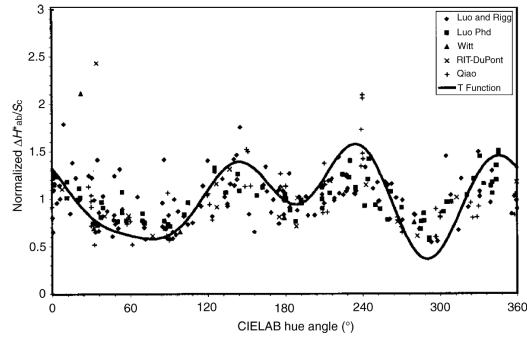
While the CMC color distance formula introduced a weighting function regarding the lightness ( $S_L$ ) the CIE94 neglected this. Due to the so called crispening effect that states that color differences are more visible if the contrast to the background is very small, a weighting function is inevitable regarding the lightness (Figure 25a). Therefore, in low lightness areas the weighting function regarding the lightness difference ( $S_L$ ) had to be extended to be more sensitive to small color differences. The same weighting for  $S_C$  as in the CIE94 color distance formula was used in the CIEDE2000 color distance function. However, by plotting the experimental data of multiple experimental color difference datasets against the CIE  $L^*a^*b^*$  hue angle it Luo and Rigg could implement a curve fit that follows the general distribution of the data (Figure 25b). This "wavy" function was implemented into the CIEDE2000 color distance formula as weighting function  $S_H$  seen in equation 48. Due to the rotation of the ellipses another so called rotation



term was added to the formula leading to the CIEDE2000 color distance formula (Equations 51–53).



(a) Dependence of relative color difference on lightness (from Luo, Cui, et al. 2001)



(b) Normalized CIE L\*a\*b\* hue-difference data against the hue angle (from Luo, Cui, et al. 2001)

The rotation term introduced in the CIEDE2000 color distance formula proves to be very effective in a very particular hue angle around 270°. The new coordinates, parametric factors, weighting function, and the rotation term are merged into the CIEDE2000 formula with equation 40. During the performance testing using the method standardized residual sum of squares (STRESS) (Melgosa et al. 2008; Melgosa 2013; García et al. 2007). Melgosa et al. showed that the CIEDE2000 color distance formula yields the lowest or best results in comparison to the CIE94, CMC( $l : c$ ), and Euclidean color distance (Melgosa 2013).

**CIEDE2000 color distance formula**

$$\Delta E_{00} = \sqrt{\left(\frac{\Delta L'}{k_L S_L}\right)^2 + \left(\frac{\Delta C'}{k_C S_C}\right)^2 + \left(\frac{\Delta H'}{k_H S_H}\right)^2 + R_T \left(\frac{\Delta C'}{k_C S_C}\right) \left(\frac{\Delta H'}{k_H S_H}\right)} \quad (40)$$

- $\Delta L'$  =CIEDE2000 lightness difference
- $\Delta C'$  =CIEDE2000 chroma difference
- $\Delta H'$  =CIEDE2000 hue angle difference
- $S_L, S_C, S_H$ =CIEDE2000 weighting functions
- $k_L, k_C, k_H$  =CIEDE2000 parametric factors

**CIEDE2000 transformation**

$$L^* = L' \quad (41)$$

$$b^* = L' \quad (42)$$

The elongation of the  $a^*$  coordinate to transform the ellipses into spheres.

$$a^* = a^*(1 + G) \quad (43)$$

with:

$$G = 0.5 \left[ 1 - \left( \frac{(C_{ab,m}^*)^7}{\sqrt{(C_{ab,m}^*)^7 + 25^7}} \right) \right] \quad (44)$$

and

$$C_{ab,m}^* = \frac{(C_{ab,2}^* - C_{ab,1}^*)}{2} \quad (45)$$

**CIEDE2000 weighting functions**

$$S_L = \frac{1 + 0.0158(L'_m - 50)^2}{\sqrt{20 + (L'_m - 50)^2}} \text{ with : } L'_m = \frac{(L'_2 + L'_1)}{2} \quad (46)$$

$$S_C = 1 + 0.045C'_m \text{ with : } C'_m = \frac{(C'_2 + C'_1)}{2} \quad (47)$$

$$S_H = 1 + 0.015C'_m T \quad (48)$$

with:

$$T = 1 - 0.17 \cos(h'_m - 30) + 0.24 \cos(2h'_m) + 0.32 \cos(3h'_m + 6) - 0.20 \cos(4h'_m - 63) \quad (49)$$

where:

$$h'_m = \frac{h'_2 + h'_1}{2} \quad (50)$$

**CIEDE2000 rotation function**

$$R_T = -\sin(2\Delta\Theta)R_C \quad (51)$$

with:

$$\Delta\Theta = 30e^{-\left(\frac{(h'_m - 275)}{25}\right)^2} \quad (52)$$

$$R_C = 2 \times \sqrt{\left(\frac{C'_m{}^7}{C'_m{}^7 + 25^7}\right)} \quad (53)$$

**2.5 Human perception of Wine Color****2.5.1 Descriptive Analysis**

Descriptive analysis is a common method to evaluate food or wine by a panel of trained consumers. In this case unstructured linear line scales are used to describe the intensity of the attribute. The fixed endpoints of the scales are the two extremes that can be used to describe the attribute. To mitigate bias, a huge amount of standardization is used. In order to achieve a certain amount of objectivity the panel has to be fairly large (Lawless and Heymann 2010). A common way to use linear line scales is shown in Figure 26. Figure 26a shows the commonly used linear line scale for red wine. The upper scale is used to evaluate lightness and the lower scale is used for the evaluation of the color hue. Lightness has the extremes light and dark, whereas the hue scale has the extremes orange and violet as well as the middle point red. Figure 26b shows the linear line scales commonly used for white wine. It uses the same lightness scale. However, the hue scale is different as it ranges from green to yellow to orange. To transfer the ratings to numbers the distance between a fixed point, which is in this case either from the extremes or from the middle point in both directions is measured.

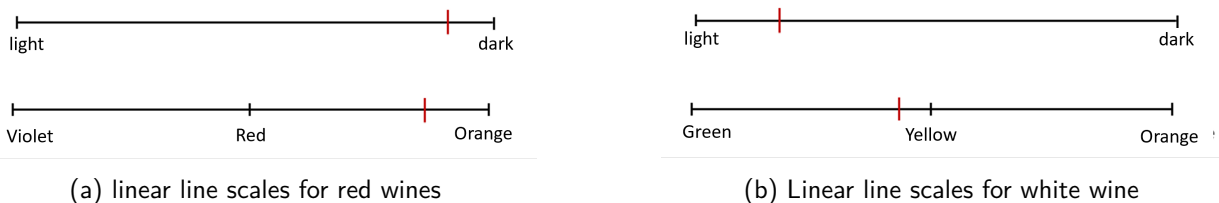


Figure 26: Linear line scales used for the description of red wine and white wine. The first scale evaluates the lightness of a wine. The second scale evaluates the color hue of a wine (modified according to Durner et al. 2010)

For the olfactoric and gustatoric evaluation blacked out glasses are used to mitigate bias coming from the color of the wine. When clear glasses are used, e.g. for color evaluation, it is crucial to fill the glasses with exactly the same volume, so that none of the samples sticks out, by anything other than their appearance. However, the descriptive analysis with respect to illumination and viewing conditions is not as standardized as the olfactoric or gustatoric evaluation (Fairchild 2018). Another problem, that can create a bias regarding wine color is the color evaluation in a standard wine tumbler. The curvature of the glass leads to different depths, different light refraction and Huertas et al. showed, that the Euclidean

distance between the center point and another point in the wine glass can be up to 21 (Huertas et al. 2003).

### 2.5.2 Sensory threshold Analysis

In order to obtain sensory thresholds, e.g. for aroma compounds, in traditional sensory analysis the so called triangle test can be used among others. During a triangle test the panel is presented with three samples simultaneously. Two of the samples have the same constitution, the other differs. Each panelist has to determine either the odd one or the two similar ones. This process is repeated with increasing concentration of the target compound, so that the difference between the odd one and the similar one increases with each cycle.

The underlying null hypothesis is that the probability of making the correct selection, when no difference between the samples are perceived is one in three ( $H_0 : P = \frac{1}{3}$ ). The alternative hypothesis is fulfilled if the probability of making the correct selection is over one in three ( $H_1 : P > \frac{1}{3}$ ). The alternative hypothesis is one-sided and the test is one-tailed, giving a number of six different serving orders (AAB, ABA, BAA, BBA, BAB, ABB). To calculate the sensory threshold, the best estimated threshold according to ASTM E-679-04 can be used. First, the individual sensory threshold for each panelist is calculated by the geometric mean between the first sustainable correct selection and the last incorrect selection. Secondly, the group sensory threshold is calculated from the individual sensory thresholds, also with the geometric mean (Equation 54).

$$\bar{x}_G = \left( \prod_{i=1}^n x_i \right)^{\frac{1}{n}} = \sqrt[n]{x_1 \cdot x_2 \cdot x_3 \cdot \dots \cdot x_n} \quad (54)$$

To use the triangle test for color a minor change has to be applied. First of all, a concentration of a certain pigment may be not enough to make assumptions about the color difference threshold as the color of a wine does not rely on the concentration of one specific color pigment. Therefore the complete wine has to be altered, e.g. by diluting dark red wine with other, lighter wines.

Previously established color difference thresholds for wine used completely different techniques. An often cited study conducted by Martínez et al. used an approach, where 105 pairs of wine were evaluated in comparison to a reference pair with a fixed color difference of  $\Delta E_{ab}^* = 4$ . The panel was asked to rate if the presented color difference is smaller or greater than the reference pair. Using logistic regression the visual perception threshold was evaluated. The 50% acceptance threshold was reached at a Euclidean color difference of 2.82. With a safety factor the visual color threshold was rounded to  $\Delta E_{ab}^* \cong 3$  (Martínez et al. 2001).

## 3 A detailed description of statistical modeling, machine learning, and Artificial Intelligence

In recent years the development of artificial intelligence has reached a point, at which untrained personnel is able to use applications based on AI. However, there are common misconceptions about the nature of different subsets of AI and how it differs from traditional statistics. The emphasis of traditional statistics is to draw conclusions about a population based on a sample set. For this, traditional statistics use hypothesis testing, regression analysis and probability theory. It often involves, manual analysis and interpretation by statisticians. Statistical modeling on the other hand constructs models that represent

the relationship among variables in data. It utilizes traditional statistics as well as other methods such as machine learning algorithms. The goals of machine learning are to develop algorithms that can learn patterns without being explicitly intended to do so. However, machine learning is only a subset of artificial intelligence. AI aims to develop systems that can perform tasks that otherwise would need human intelligence. Those incorporate perception, reasoning, learning, and problem solving.

### 3.1 Logistic Regression

Logistic regression is a special form of linear regression that is used in ML for classification. The relation to linear regression is not intuitively understandable as instead of using continuous data to predict an outcome, logistic regression predicts probabilities in terms of boolean statements, meaning it predicts the probability of an outcome as instances of true and false. The problem that makes the comparison to linear regression so complex is the scales of the data. In linear regression, technically speaking, the values of the y-axis can be  $-\infty$  to  $\infty$ . The probability scale only ranges from 0 to 1 on the y-axis. Because of that a regression using the least-squares approach is not possible. To mitigate the limitation of the probability scale, logistic regression uses the logit function to calculate the log-odds from the probability using the following function (Harrell 2001):

$$\log_{odds} = \ln\left(\frac{p}{1-p}\right) \quad (55)$$

with  $p$  being the probability of an instance. Samples in the training dataset that are in the positive class have the probability  $p = 1$ . By substituting  $p$  it is evident that positive samples are  $\log_{odds}(x^+) = \infty$ , whereas the samples of the negative class are  $\log_{odds}(x^-) = -\infty$ . The transformation of the probability to log-odds transforms the y-axis in a way that is comparable to the y-axis in linear regression. However, it is still not possible to use the least-square approach as the values of the sample are infinity and negative infinity, which means that the residuals that are needed for the least-squares approach are also infinity. Instead, logistic regression fits a line in the log-odds scale and projects the datapoints onto the fit.

Using the logistic function

$$p = \frac{e^{\log_{odds}}}{(1 + e^{\log_{odds}})} \quad (56)$$

the projected log odds can be transferred back into probabilities that lead to the "S"-shaped logistic function which now can be used to classify new data and to calculate the log-likelihood. To obtain the optimal results, the fit with the highest log-likelihood is calculated (Harrell 2001). The log-likelihood is defined as

$$\mathcal{L}(p|n, y) = \sum_{i=1}^n \ln(y_i(p_i = 1)) + \sum_{i=1}^n \ln(1 - y_i(p_i = 0)) \quad (57)$$

where:

$\mathcal{L}(p|n, y)$  = Log-likelihood with respect to  $p$  and  $y$

$y_i$  = y-value on the probability scale after log-odds fit

$p_i$  = Probability of the samples after transferring log-odds back to probability.

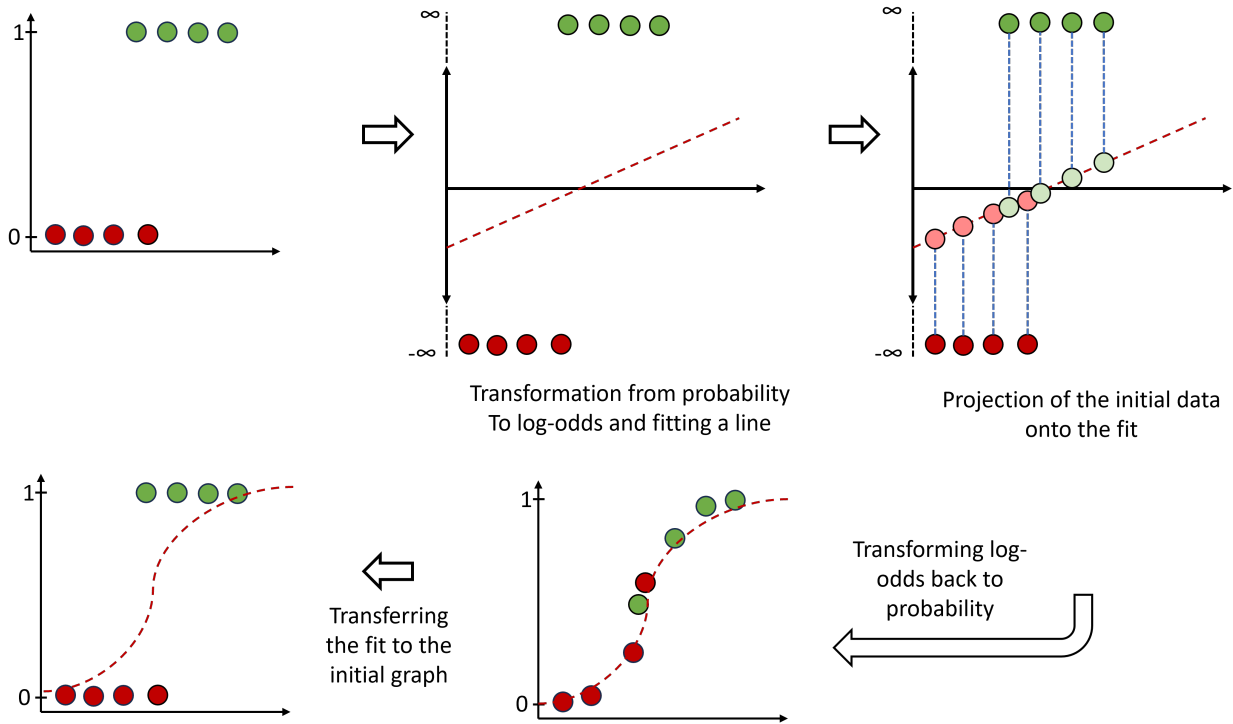


Figure 27: Visual description of the basic ideas of logistic regression (modified according to Harrell 2001)

## 3.2 Support Vector Machine

The precondition of the support vector machine is the existence of  $N$  data points  $(x_1, y_1), (x_2, y_2), \dots, (x_N, y_N)$  with  $x_i \in \mathbb{R}^g$  ( $g \in \mathbb{N}$ ) and  $y_i \in \{-1, 1\}$  for  $i = 1, 2, \dots, N$ . The first component of the vector is called the input data, while the second component of the vector is the class into which the data is categorized. This classification is done by evaluating a function  $f: \mathbb{R}^g \rightarrow \{-1, 1\}$  that correctly classifies the training data. It is imperative to find a function that is  $f(x_i) = y_i$  either  $\forall i = 1, 2, \dots, N$  or for as many  $i$  as possible. It is then possible to classify new data points accordingly (Boser et al. 1992).

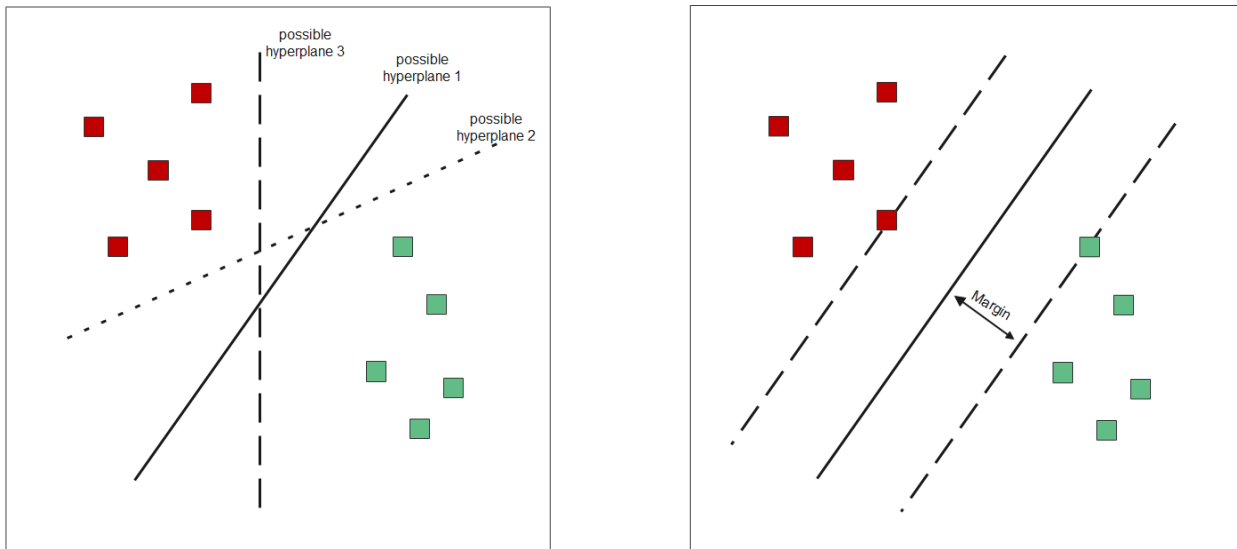
### 3.2.1 Linearly separable data

Figure 28a shows the general problem when searching for a decision boundary. The decision boundary denotes the border between the classes, that can be mathematically described by hyperplanes. Hyperplanes are by definition a geometric concept that is defined as a flat-affine subspace. In an  $n$ -dimensional space, the hyperplanes are  $n - 1$  dimensional. In the context of SVMs all the shown hyperplanes in Figure 28a create a decision boundary that perfectly separates the two classes, the goal is to maximize the general applicability. Therefore, it is necessary to find the hyperplane that has the largest distance between the data points of the two classes (Figure 28b).

The first step is to define a hyperplane as follows:

$$H := \{x \in \mathbb{R}^g \mid \langle \vec{w}, \vec{x} \rangle + b = 0\} \quad (58)$$

Herein  $\vec{w} \in \mathbb{R}^g$  is a vector that is perpendicular to the hyperplane  $H$ , whereas  $b \in \mathbb{R}$  is a constant that regulates the shift of the hyperplane in the direction of  $\vec{w}$ .



(a) Depiction of the problem when searching for optimal decision boundaries

(b) Depiction of the margin using the maximum margin classification

Figure 28: The basic principle of SVM. The most suitable hyperplane has to be identified. With the Hyperplane and the support vectors the margin can be derived.

As mentioned earlier it is necessary to find the largest distance between the two classes and for this the margin is introduced to the problem. By the normalization of the vector  $\vec{w}$  the dot product of  $\vec{w}$  and  $\vec{x}$  is the length of the projection of vector  $\vec{x}$  in the orientation of  $\vec{w}$ . By adding the constant  $b$  the distance of the point  $\vec{x}$  to the hyperplane can be calculated. To define the margin the result of the dot product of  $\vec{w}$  and  $\vec{x}$  added with  $b$  for one class is equal to 1. If the dot product of  $\vec{w}$  and  $\vec{x}$  added with  $b$  is equal to -1, the data point belongs to the other class. This formulation is written as follows:

$$\langle \vec{w}, \vec{x}^+ \rangle + b = 1 \quad (59)$$

$$\langle \vec{w}, \vec{x}^- \rangle + b = -1 \quad (60)$$

$$(61)$$

$$\Rightarrow \langle \vec{w}, (\vec{x}^+ - \vec{x}^-) \rangle = 2 \quad (62)$$

$$\Rightarrow \left\langle \frac{\vec{w}}{\|\vec{w}\|}, (\vec{x}^+ - \vec{x}^-) \right\rangle = \frac{2}{\|\vec{w}\|} \quad (63)$$

If two points of contrary classes that lay directly on the margin are observed it is evident that the margin can be described as  $\frac{1}{\|\vec{w}\|}$ . For mathematical convenience  $\frac{1}{\|\vec{w}\|}$  is transformed to  $\frac{1}{2}\|\vec{w}\|^2$ . Therefore, to maximize the margin the length of  $\vec{w}$  has to be minimized, leading to the following optimization problem:

$$\min \frac{1}{2}\|\vec{w}\|^2 \quad (64)$$

To simplify the expressions written in equation 59 and 60 the factor  $y_i$  is introduced. By multiplication of the respective class the result of the function is always positive. This leads in the constraint of the optimization problem

$$y_i(\langle \vec{w}, \vec{x}_i \rangle + b) \geq 1, \quad \forall i = 1, \dots, N. \quad (65)$$

The result of this is a convex optimization problem, meaning a quadratic target function with a linear

constraint.

$$\min \frac{1}{2} \|\vec{w}\|^2 \quad (66)$$

$$\text{s.t. } y_i(\langle \vec{w}, \vec{x}_i \rangle + b \geq 1), \forall i = 1, \dots, N \quad (67)$$

To solve the optimization problem the Lagrangian multiplier method can be used.

$$\mathcal{L}(x, \alpha) = \nabla f(x) - \alpha \nabla g(x) \quad (68)$$

with:

$f(x)$  = Target function of the optimization problem

$g(x)$  = Constraint of the optimization problem

$\alpha$  = Lagrangian multiplier

This leads to the following expression:

$$\mathcal{L}(\vec{w}, b, \alpha) = \frac{1}{2} \|\vec{w}\|^2 - \sum_{i=1}^N \alpha_i (y_i (\langle \vec{w}, \vec{x}_i \rangle + b) - 1) \quad (69)$$

To solve this the Lagrangian has to be minimized and therefore the  $\alpha$  has to be minimized. For this, the partial differentiation of  $\mathcal{L}$  with respect to the other variables have to be equal to 0:

$$\frac{\partial \mathcal{L}}{\partial \vec{w}} = \vec{w} - \sum_{i=1}^N \alpha_i y_i \vec{x}_i = 0 \quad (70)$$

$$\Rightarrow \vec{w} = \sum_{i=1}^N \alpha_i y_i \vec{x}_i \quad (71)$$

$$\frac{\partial \mathcal{L}}{\partial b} = \sum_{i=1}^N \alpha_i y_i = 0 \quad (72)$$

Substituting the expression above in the primal form of the Lagrangian results in the following solution of the optimization problem:

$$\begin{aligned} \mathcal{L}(\vec{w}, b, \alpha) &= \left[ \frac{1}{2} \left( \sum_{i=1}^N \alpha_i y_i \vec{x}_i \right) \left( \sum_{j=1}^N \alpha_j y_j \vec{x}_j \right) \right] \\ &- \sum_{i=1}^N \alpha_i y_i \vec{x}_i \left( \sum_{j=1}^N \alpha_j y_j \vec{x}_j \right) - \sum_{i=1}^N \alpha_i y_i b + \sum_{i=1}^N \alpha_i \end{aligned} \quad (73)$$

$$\Rightarrow \max \sum_{i=1}^N \alpha_i - \frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N \alpha_i \alpha_j y_i y_j \langle \vec{x}_i, \vec{x}_j \rangle \quad (74)$$

$$\text{s.t. } \sum_{i=1}^N \alpha_i y_i = 0 \text{ and } \alpha_i \geq 0, \forall i = 1, \dots, N \quad (75)$$



According to the Karush-Kuhn-Tucker condition, the following applies to the saddle point:

$$\alpha_i [y_i (\langle \vec{w}, \vec{x}_i \rangle + b) - 1] = 0, \forall i = 1, \dots, N \quad (76)$$

Consequently, it is evident that, while the optimum applies, either  $\alpha_i = 0$  or  $y_i (\langle \vec{w}, \vec{x}_i \rangle + b) = 1$ . Considering the formulation seen in equation 70 it is evident that only the training points with  $\alpha_i > 0$  have an influence of the solution. These vectors that fulfill this expression are called support vectors (Boser et al. 1992).

### 3.2.2 Non-linear separable data

The disadvantage of the hard-margin or maximum margin SVM explained above is that it is only applicable for datasets, where the classes are linearly separable. Figure 29 shows a dataset where the data is only separable in a non-linear fashion. To mitigate this inconvenience the basic idea is that the input vector space is transferred to a higher dimensional feature space ( $\phi$ ) in which the classes of the dataset are linearly separable. Transforming the linear hyperplane from the high-dimensional feature space back to the low-dimensional input vector space the decision boundary appears to be non-linear. In the higher dimensional feature space the dot products are following the form  $\langle \phi(x_i), \phi(x_j) \rangle$ . To limit the computation load a kernel function is used, that maps from  $\mathbb{R}^g \times \mathbb{R}^g$  and that inherits a dot product in  $\mathbb{M} = \mathbb{R}^m$  with  $m \in \mathbb{R} > g$ :

$$k(x_i, x_j) = \langle \phi(x_i), \phi(x_j) \rangle \quad (77)$$

where  $k : \mathbb{R}^g \times \mathbb{R}^g \rightarrow \mathbb{M}$  and is positive semi-definite

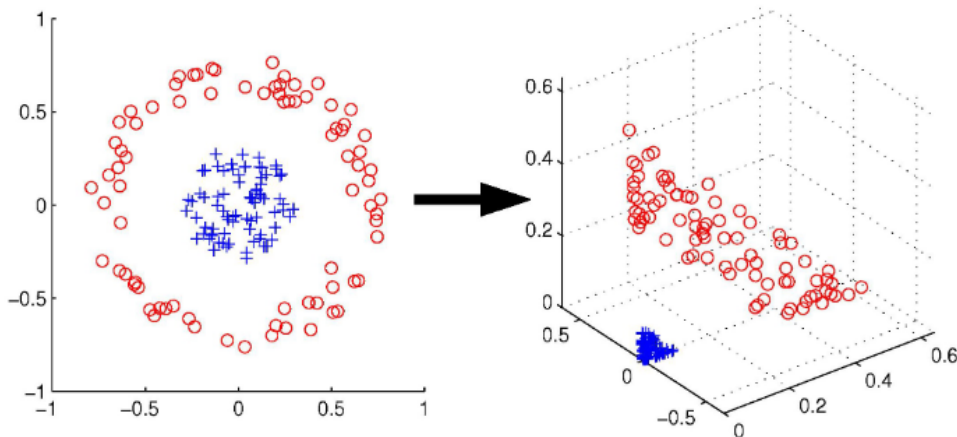


Figure 29: Example for a dataset that is not linear separable. The kernel function transforms the two-dimensional input space in a three-dimensional feature space (from Jordans 2004)

The most commonly used kernel functions in SVM are the aforementioned linear kernel, polynomial kernel, the radial basis function, and the sigmoidal kernel. The kernel function for each kernel are depicted in table 5. The  $\gamma$  parameter as well as the polynomial degree are regularization parameters chosen empirically via hyperparameter tuning (Boser et al. 1992; Bennett and Campbell 2000).

Table 5: Kernel function of the most commonly used kernels.

Kernel Name	Kernel Function
linear	$k(x_i, x_j) = (\gamma \langle \phi(x_i), \phi(x_j) \rangle)$
Polynomial	$k(x_i, x_j) = (\gamma \langle \phi(x_i), \phi(x_j) \rangle + r)^d$
RBF	$k(x_i, x_j) = e^{-\gamma \ \phi(x_i) - \phi(x_j)\ ^2}$
Sigmoidal	$k(x_i, x_j) = \tanh(\gamma \langle \phi(x_i), \phi(x_j) \rangle) + r$

### 3.2.3 Soft margin SVM

The transformation of a non-linearly separable dataset is effective, but computationally intensive. So, a soft margin SVM can be an effective less computationally complex work around. For this purpose the

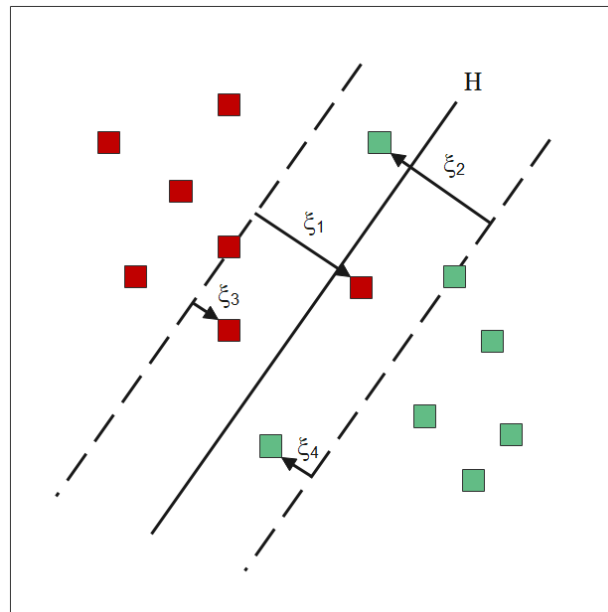


Figure 30: Training dataset with outliers and depiction of the effect of the slack variable  $\xi$ . The slack variable is the distance between the misclassified point and the margin that borders the respective class.

soft margin SVM introduces a slack variable  $\xi \geq 0$  to the optimization problem. If the classification is correct  $\xi_i = 0$ . If the datapoint lays past the decision boundary  $H$ , the classification is considered wrong and here  $\xi_i = 1$  (Figure 30  $\xi_1$  and  $\xi_2$ ). However, if the classification is correct but the position of the datapoint is inside the margin the slack variable is  $0 < \xi_i < 1$  (Figure 30  $\xi_3$  and  $\xi_4$ ).

The mathematical formulation also introduces a regularization parameter  $C$  as a penalty term. The higher  $C$  is set, the higher is the impact of misclassification. The magnitude of this regularization parameter is also determined by hyperparameter tuning

$$\min \frac{1}{2} \|\vec{w}\|^2 + C \sum_{i=1}^N \xi_i \quad (78)$$

$$\text{s.t. } y_i(\langle \vec{w}, \vec{x}_i \rangle + b) \geq 1, \quad \forall i = 1, \dots, N \quad (79)$$

To fulfill the minimization problem the Lagrangian is constructed as mentioned in equation 69:

$$\mathcal{L}(\vec{w}, b, \alpha, \xi_i) = \frac{1}{2} \|\vec{w}\|^2 + C \sum_{i=1}^N \xi_i - \sum_{i=1}^N \alpha_i [y_i(\langle \vec{w}, \vec{x}_i \rangle + b) - 1 + \xi_i] \quad (80)$$

To solve the Lagrangian the partial derivatives are calculated analog to equations 70 – 72:

$$\frac{\partial \mathcal{L}}{\partial \vec{w}} = \vec{w} - \sum_{i=1}^N \alpha_i y_i \vec{x}_i = 0 \Rightarrow \vec{w} = \sum_{i=1}^N \alpha_i y_i \vec{x}_i \quad (81)$$

$$\frac{\partial \mathcal{L}}{\partial b} = \sum_{i=1}^N \alpha_i y_i = 0 \quad (82)$$

$$\frac{\partial \mathcal{L}}{\partial \xi_i} = C - \sum_{i=1}^N \alpha_i = 0 \Rightarrow C = \sum_{i=1}^N \alpha_i = 0 \quad (83)$$

Substituting the differential to the primal form of the Lagrangian leads to the following notation of the optimization problem:

$$\begin{aligned} \mathcal{L}(\vec{w}, b, \alpha, \xi_i) = & \left[ \frac{1}{2} \left( \sum_{i=1}^N \alpha_i y_i \vec{x}_i \right) \left( \sum_{j=1}^N \alpha_j y_j \vec{x}_j \right) + \sum_{i=1}^N \alpha_i \xi_i \right] \\ & - \sum_{i=2}^N \alpha_i \xi_i - \left( \sum_{i=1}^N \alpha_i y_i \vec{x}_i \right) \left( \sum_{j=1}^N \alpha_j y_j \vec{x}_j \right) - \sum_{i=1}^N \alpha_i y_i b + \sum_{i=1}^N \alpha_i \end{aligned} \quad (84)$$

$$\Rightarrow \max \sum_{i=1}^N \alpha_i - \frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N \alpha_i \alpha_j y_i y_j \langle \vec{x}_i, \vec{x}_j \rangle \quad (85)$$

$$\text{s.t } 0 \leq \alpha_i \leq C, \forall i = 1, \dots, N \quad (86)$$

$$\sum_{i=1}^N \alpha_i y_i = 0 \quad (87)$$

Lastly the Karush-Kuhn-Tucker condition is applied to the solution of the optimization problem:

$$\alpha_i [y_i (\langle \vec{w}, \vec{x}_i \rangle) - (1 - \xi_i)] = 0 \quad (88)$$

$$\alpha_i \xi_i = 0 \quad (89)$$

$$y_i (\langle \vec{w}, \vec{x}_i \rangle + b) - (1 - \xi_i) \geq 0, \forall i = 1, \dots, N \quad (90)$$

leading to the conclusion that two different specimen of support vectors exist. Either a support vector is directly on the border of the margin (with  $\xi_i = 0$  and following to that  $0 < \alpha_i < C$ ) or the support vector lays beyond the margin (with  $\xi_i > 0$  and  $\alpha_i = C > 0$ ) (Bennett and Campbell 2000).

To conclude, this chapter shows how support vector machines work and lays the foundation for solving an optimization problem using support vectors. Of course, it is possible to use kernelization with a soft margin as real world data is often very complex and computation effort aside, the kernel functions seen in the last chapter can increase the performance of the soft margin SVM significantly.

### 3.3 Extreme Gradient Boosting

Gradient boosting is a machine learning technique that can be used for regression or classification purposes. It is an ensemble technique, meaning that the predictions combine several base-estimators to produce a strong predictive model. The principles of gradient boosting can be divided into five main ideas. First, the model is initialized with a simple prediction. In many cases the prediction is just the arithmetic mean value of the target variable for regression or the target class for classification. The next idea is that in contrast to bagging techniques (e.g. Random Forest), the base-estimators are not trained independently but sequentially, meaning with every new iteration the new base-estimator is trained to correct the errors made by the last base-estimators. The third idea is called gradient descent optimization. Gradient boosted ensembles contain a loss function and at each iteration the gradient of the loss function with respect to the prediction is calculated. The new base-estimator is trained to reduce the error of the ensemble by minimizing this gradient. XGBoost uses gradient boosting as basic idea, but enhances the optimization that improves performance, speed, and accuracy. However, normal gradient boosting has a crucial disadvantage. Because of the countless iteration the risk of overfitting the model on the training data is very high. To mitigate this, XGBoost uses Lasso regression (L1) and Ridge regression (L2) as regularization terms. Furthermore, XGBoost inherits an intrinsic mechanism to cope with missing data as it is the case with "real world" datasets. It is perfectly fit for large datasets as it uses parallelization to use multiple CPU cores during training, which increases the speed of the algorithm significantly in comparison to other gradient boosting implementation. XGBoost also has a mechanism called tree-pruning and implemented a learning rate mechanism. Tree-pruning can be seen as a control instance to control the size of the built decision trees. The learning rate is a hyperparameter that controls the step size at which the model weights are updated during training. Both hyperparameters are implemented to prevent overfitting. Furthermore, it uses cross validation during the training process (Figure 31). The flexibility of the algorithm leads to customized objective functions so that it can be used for regression and classification (Chen and Guestrin 2016). The main ideas of XGBoost are shown in Figure 31.

#### 3.3.1 Objective function

XGBoost uses the following objective function that needs to be minimized:

$$F(y_i, p_i) = \sum_{i=1}^n L(y_i, p_i) + \sum_{k=1}^K \Omega f(O_v) \quad (91)$$

with:

- $n$  =number of training samples
- $y_i$  =true label of the i-th sample
- $p_i$  =predicted value of the i-th sample
- $K$  =number of base learners (trees)
- $\Omega$  =Regularization Term
- $f(O_v)$ =Function of the output value of k-the base learner

The objective function contains two separate parts. The first part being a loss function, which measures the distance between the predicted values and the true labels. For the regression approach the squared

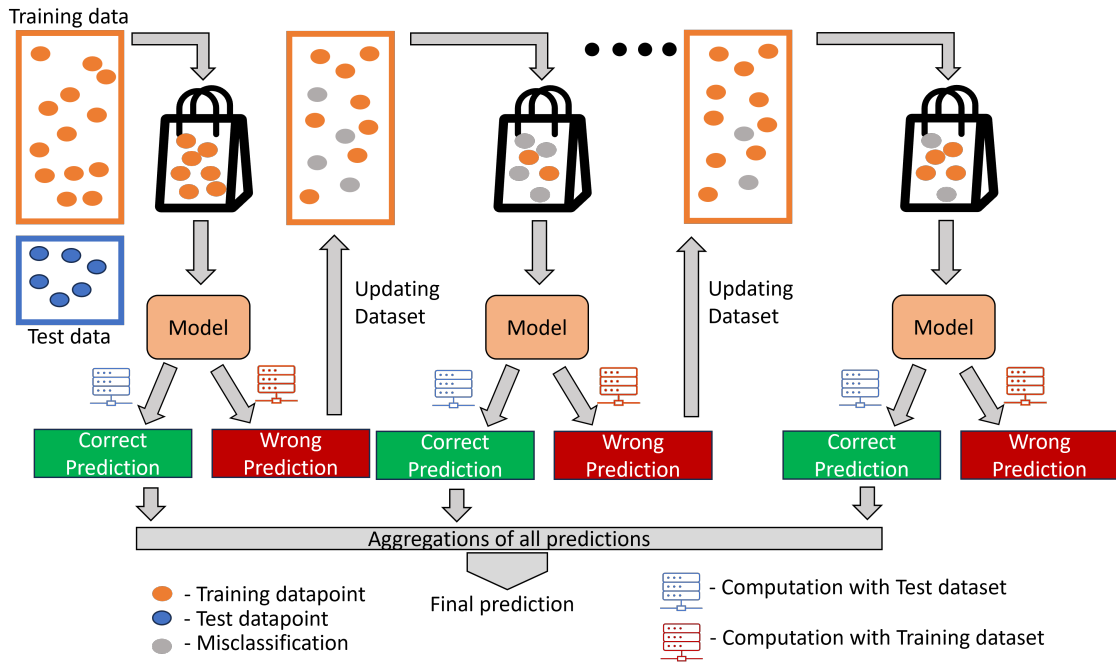


Figure 31: The summary of the main ideas of Extreme Gradient Boosted trees (modified according to Vestner 2021)

loss function is used.

$$L(y_i, p_i) = \frac{1}{2}(y_i - p_i)^2 \quad (92)$$

For binary classification the logistic loss function is used.

$$L(y_i, p_i) = -[y_i \ln(p_i) + (1 - y_i) \ln(1 - p_i)] \quad (93)$$

The second part of the objective function is the regularization term  $\Omega$ . It penalizes the complexity of the model to prevent overfitting. For that XGBoost uses Lasso and Ridge regression, where  $\gamma$  and  $\lambda$  are hyperparameters,  $T$  are the number of terminal nodes and  $O_v$  weight of a leaf. XGBoost combines both regularization terms in their objective function:

$$\Omega f(O_v) = \gamma T + \frac{1}{2} \lambda f(O_v)^2 \quad (94)$$

### 3.3.2 Gradient Boosting

For the gradient boosting the model uses an additive approach. For this, let  $p_i^{(t)}$  be the prediction of the  $i$ -th sample at the  $t$ -th iteration. The prediction of the current base-learner is equal to the prediction of the last base-learner added with the output value of the current base-learner  $p_i^t = p_i^{(t-1)} + O_v$ .

$$F(y_i, p_i)^{(t)} = \sum_{i=1}^n L(y_i, p_i^{(t-1)} + O_v) + \Omega f(O_v) \quad (95)$$

The optimal value or the minimization of the loss function can only be achieved by approximation. To approximate the loss function around the current prediction  $p_i^{(t)}$  XGBoost uses a second degree Taylor series expansion, which is in this case:

$$L(y_i, p_i)^{(t)} \cong L(y_i, p_i^{(t-1)}) + \left[ \frac{\partial L(y_i, p_i^{(t-1)})}{\partial p_i^{(t-1)}} \right] O_v + \frac{1}{2} \left[ \frac{\partial^2 L(y_i, p_i^{(t-1)})}{\partial^2 p_i^{(t-1)}} \right] O_v^2 \quad (96)$$

The first order differential is known as the gradient  $g_i$ , whereas the second order differential is known as a Hessian  $h_i$ , which leads to the to the simplified version of the Taylor approximation:

$$F(y_i, p_i)^{(t)} \cong L(y_i, p_i^{(t-1)}) + g_i O_v + \frac{1}{2} h_i O_v^2 \quad (97)$$

Substituting equation 95 with the approximation in equation 97 leads to the updated objective function:

$$F(y_i, p_1) = \sum_{i=1}^n \left[ L(y_i, p_i^{(t-1)}) + g_i O_v + \frac{1}{2} h_i O_v^2 \right] + \Omega f(O_v) \quad (98)$$

By substituting  $\Omega f(O_v)$  with the equation 94, the complete objective function can be written as:

$$\begin{aligned} F(y_i, p_i)^{(t)} &= \sum_{i=1}^n \left[ L(y_i, p_i^{(t-1)}) + g_i O_v + \frac{1}{2} h_i O_v^2 \right] + \gamma T + \sum_{j=1}^T \frac{1}{2} \lambda O_v^2 \\ &= \sum_{j=1}^T \left[ \sum_{i=1}^n L(y_i, p_i^{(t-1)}) + g_i O_v + \frac{1}{2} \sum_{i=1}^n (h_i + \lambda) O_v^2 \right] + \gamma T \end{aligned} \quad (99)$$

where  $j$  is written as the instance of a leaf and  $T$  is the number of terminal leaves. The loss of the previous base-learner  $(L(y_i, p_i^{(t-1)}))$  is constant for every summation and therefore can be removed, leading to the new expression:

$$\sum_{j=1}^T \left[ \sum_{i=1}^n g_i O_v + \frac{1}{2} \sum_{i=1}^n (h_i + \lambda) O_v^2 \right] + \gamma T \quad (100)$$

The remaining part of the loss function is a quadratic equation and as mentioned before needs to be minimized or in other words the extreme has to be calculated. Therefore, the derivative respective to  $O_v$  has to be equal to zero. With that the prediction of the leafs can be calculated.

$$\begin{aligned} \frac{d}{dO_v} &= \sum_{i=1}^n g_i + \sum_{i=1}^n (h_i) + \lambda O_v = 0 \\ \Rightarrow O_v &= -\frac{\sum_{i=1}^n g_i}{\sum_{i=1}^n (h_i) + \lambda} \end{aligned} \quad (101)$$

The question asked here is how the algorithm calculates the new prediction of the leaf in practice. For regression the output values can be calculated by substituting the loss function with the function displayed in equation 92, leading to the following expression:

$$\begin{aligned}
 L(y_i, p_i) &= \frac{1}{2}(y_i - p_i)^2 \\
 g_i &= \frac{\partial L(y_i, p_i)}{\partial p_i} = -(y_i - p_i) \\
 h_i &= \frac{\partial^2 L(y_i, p_i)}{\partial^2 p_i} = 1 \\
 \Rightarrow O_v &= \frac{\sum_{i=1}^n (y_i - p_i)}{\sum_{i=1}^n 1} = \frac{\text{Sum of residuals}}{\text{Number of residuals}}
 \end{aligned}$$

For classification the negative log-loss function is transferred to the negative log-odds loss function. After that, the derivatives with respect to the log-odds are calculated and the results are transferred back probability:

$$\begin{aligned}
 L(y_i, p_i) &= -[y_i \ln(p_i) + (1 - y_i) \ln(1 - p_i)] \\
 L(y_i, \log_{odds}) &= -y_i \ln(\log_{odds}) + \ln(1 + e^{\log_{odds}}) \\
 g_i &= \frac{\partial L(y_i, \log_{odds})}{\partial \log_{odds}} = -y_i + \frac{e^{\log_{odds}}}{1 + e^{\log_{odds}}} \\
 &\hat{=} -(y_i - p_i) \\
 h_i &= \frac{\partial^2 L(y_i, \log_{odds})}{\partial^2 \log_{odds}} = \frac{e^{\log_{odds}}}{(1 + e^{\log_{odds}})^2} \times \frac{1}{(1 + e^{\log_{odds}})} \\
 &\hat{=} p_i(1 - p_i)
 \end{aligned}$$

Analogous to this, the log-loss function can be used to calculate the output value for classification purposes:

$$O_v = \frac{\sum_{i=1}^n y_i - p_i}{\sum_{i=1}^n p_i(1 - p_i)} = \frac{\text{Sum of residuals}}{\text{Sum of probability} \times (1 - \text{previous probability})} \quad (102)$$

The last question that needs to be answered is, how the algorithm determines how the data is clustered between the leafs. For that, the equation of the output values (Equation 101) has to be inserted into the objective function displayed in equation 99.

$$F(y_i, p_i)^{(t)} = \sum_{j=1}^T \left[ \sum_{i=1}^n (g_i) \left( \frac{\sum_{i=1}^n g_i}{\sum_{i=1}^n (h_i) + \lambda} \right) + \frac{1}{2} \sum_{i=1}^n (h_i) + \lambda \left( \frac{\sum_{i=1}^n g_i}{\sum_{i=1}^n (h_i) + \lambda} \right)^2 \right] + \gamma \quad (103)$$

Through simplification the formula transforms to the corresponding optimal similarity score:

$$F(y_i, p_i)^{(t)} = -\frac{1}{2} \sum_{j=1}^T \left( \underbrace{\frac{\sum_{i=1}^n g_i^2}{\sum_{i=1}^n (h_i) + \lambda}}_{\text{similarity score}} \right) \quad (104)$$

When the algorithm builds a decision tree, the root ( $I$ ) is split up into a left leaf ( $I_L$ ) and a right leaf ( $I_R$ ). For each leaf and the root the similarity score can be calculated. With the following equation the gain ( $G$ ) of a tree can be calculated. The gain can be used as a quality parameter of the tree. The

higher the value the better the tree fit to the data (Chen and Guestrin 2016).

$$G = \frac{1}{2} \left[ \frac{(\sum_{i \in I_L} g_i)^2}{\sum_{i \in I_L} (h_i) + \lambda} + \frac{(\sum_{i \in I_R} g_i)^2}{\sum_{i \in I_R} (h_i) + \lambda} - \frac{(\sum_{i \in I} g_i)^2}{\sum_{i \in I} (h_i) + \lambda} \right] \quad (105)$$

### 3.3.3 Tree pruning and other mechanism to prevent overfitting

Tree pruning can be seen as a safety mechanism to prevent overfitting. The question asked is how XGBoost is able to use regularization to prune the tree. The hyperparameters for this are  $\lambda$  and  $\gamma$ . While  $\lambda$ , being the denominator of the output value lowers the output value and therefore reduces the size of the steps performed by gradient boosting. The hyperparameter  $\gamma$  has only an effect when the tree is already built. By subtracting  $\gamma$  from  $G$  the result of this calculation is the argument whether or not a node and respectively the leaves are deleted. If the result of this calculation is negative the node is deleted, if positive the node is not deleted. The formula can be used for evaluating split candidates. Next to the gradient boosting and tree building XGBoost uses shrinkage or learning rate to prevent overfitting as it was introduced by Friedman et al. 2000. The basic principle is that the newly added weights are multiplied with a factor  $\eta$  after each step of tree boosting to scale the influence of each individual tree. This is common practice as it leaves space to improve the models with future trees.

To summarize, XGBoost uses multiple machine learning principles and enhances them through combination. It is designed for large datasets with a large amount of features and is far more complex than logistic regression and SVM. The principles of XGBoost are compressed into a corresponding python package, which makes it very feasible to use (Chen and Guestrin 2016).



### 3.4 Neural Networks, Multilayer Perceptron Models, and Deep Learning

Neural networks are a powerful tool in machine learning that uses interconnected layers and an specialized training process enabling them to perform a wide range of tasks from classification, regression, image and speech recognition, autonomous driving etc. (Pedregosa F et al. 2011). Neural networks are structured in layers, of which three types exist and each layer consists of numerous neurons. The first layer is the input layer, where the number of neurons corresponds to the number of features in the used dataset. Next to the input layer are the hidden layers. The number of hidden layers correlates with the complexity of the neural network. Hidden layers transform and process the data in intricate ways and pass the results to the next layer, the output layer (Figure 32). The output layer is the neural networks final estimation, where each neuron represent a different class or a continuous value for regression approaches (Starmer 2022; Sonnet 2022; Kruse et al. 2015).

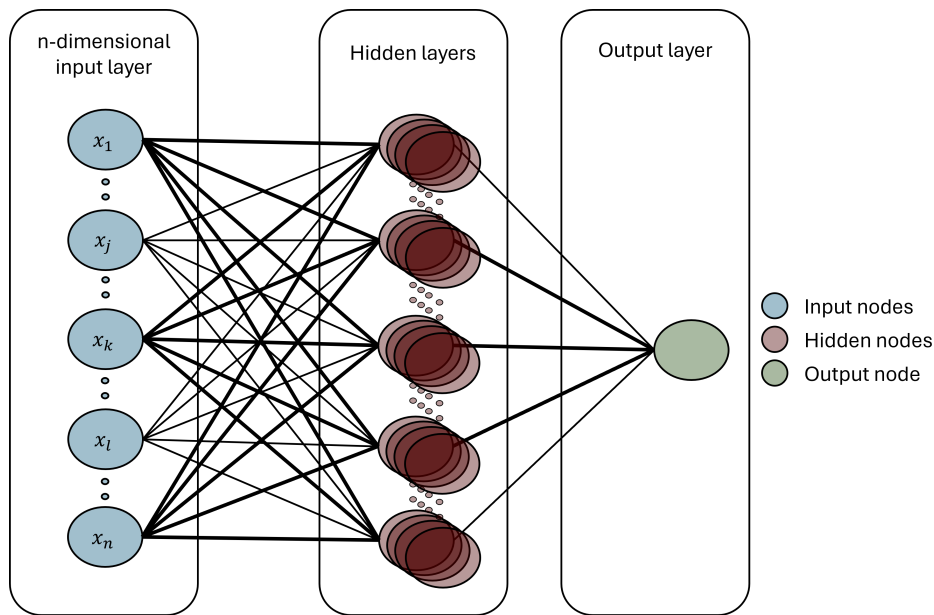


Figure 32: General architecture of a multilayer perceptron neural network (modified according to Bhowmik 2019)

Neurons are fundamental for neural networks as these are the processing nodes. Each node receives data, processes it and passes it forward to the next layer. During processing, the neuron assigns a weight to the data followed by an activation using an activation function (Figure 33). Weights determine the importance of an input, whereas activation functions introduce non-linearity into the model enabling it to model any complex relationship. In order to scale the data a bias  $b_j$  is added as well. During training the neural network adjusts the weights in order to minimize the distance between predicted outcome and actual output, hence the learning. The process starts by so called forward propagation, where the input dataset passes through the layers of the neural network. The error between predicted outcome and actual outcome is calculated by a loss function. In a process called backpropagation the weights are adjusted to minimize the loss function. This process repeats itself until the loss does not change anymore or some other criteria is met (Starmer 2022; Sonnet 2022; Kruse et al. 2015).

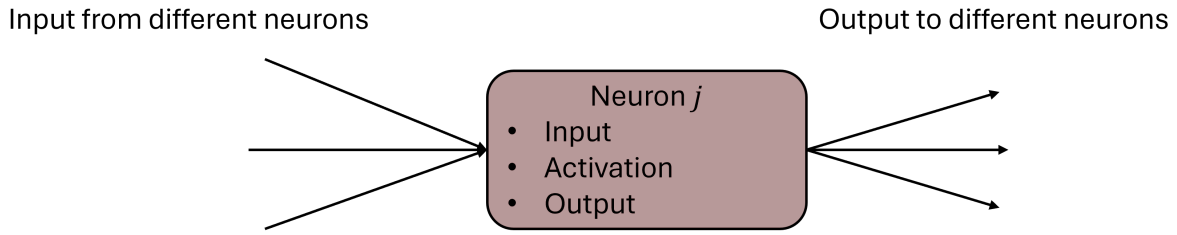


Figure 33: Mode of operation of a single neuron in the neuronal network (modified according to Sonnet 2022).

The output of a neuron  $j$  can be represented as the weighted sum of the outputs of prior established neurons ( $R$ ) scaled with a bias  $b$ :

$$z_j = \sum_{i \in R} w_{ij} x_i + b_j \quad (106)$$

where:

$x_i$  =input to the neuron

$w_{ij}$ =weights associated with the inputs

$b_j$  =bias term

$z_j$  =linear combination of input and weight

The process of computing an output from an input is called forward propagation. The input data is passed through each layer of the neural network, from input to output layer, through the hidden layers, where weights, biases, and activation functions are applied to each neuron. The activation, much like the biological counterpart of the computational neuron, only occurs if a threshold ( $\Theta$ ) is met (Kruse et al. 2015; Sonnet 2022:)

$$f(z_j, \Theta) = \begin{cases} 1 & \text{if } z_j \geq \Theta \\ 0 & \end{cases} \quad (107)$$

The weighted sum of input and weight, scaled by bias represents the x-value in the activation function. By calculation, the magnitude of activation is represented by the corresponding y-value. Next to the stepwise, so called, Y-Heaviside activation, other concept use functions that enable a smooth transition. In this case, a mathematical function is applied to the result of the input to determine the magnitude of activation. Examples for this are the sigmoidal function (Equation 108) (Figure 34). Another function that is often used is the so called Rectifier Linear Unit Function (Equation 109) or ReLU-Function (Figure 34)

$$f(z_j) = \frac{1}{1 + e^{(-z_j)}} \quad (108)$$

The activation with the sigmoidal function follows the "S-shaped" curve, whereas the ReLU function outputs the bigger number, either 0 or  $z_j$ .

$$f(z_j) = \max(0, z_j) \quad (109)$$

Figure 35 shows the calculations during forward propagation. The initial input is multiplied with the attributed weight and added to the attributed bias. The result is corresponding to the x-axis value of

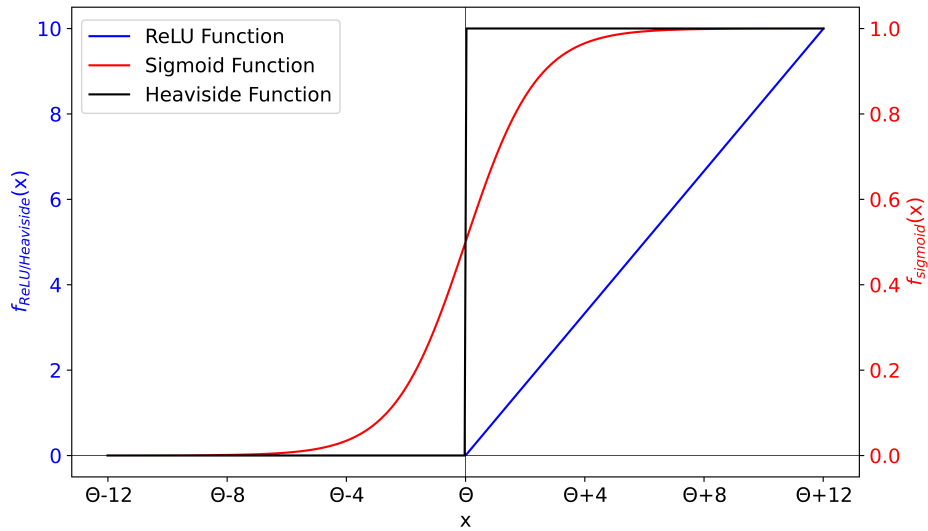


Figure 34: Depiction of the Heaviside, ReLU, and Sigmoidal activation functions of the neuron Kruse et al. 2015; Sonnet 2022.

the activation and the y-axis values are calculated with the used activation function. The output of the neuron is then again multiplied with another weight and added to another bias. This process repeats itself with every added hidden layer. The output value is defined with as the sum of all output of all neurons leading to the output node. It is then possible to scale the output again with a bias or perhaps even activate it once more with an activation function (Starmer 2022).

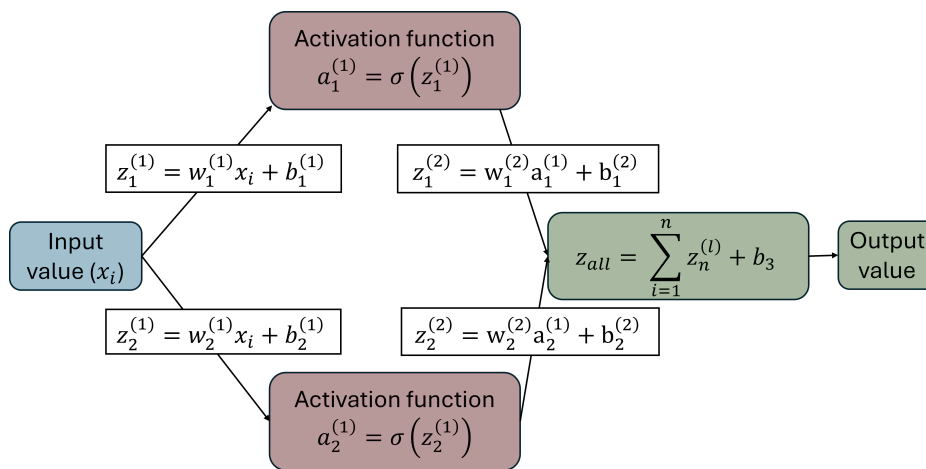


Figure 35: The process of forward propagation in a three layer neural network with one input layer, one hidden layer with two neurons and an output layer (modified according to (Starmer 2022)).

At first weights and biases are chosen randomly and are updated and optimized during the process of backpropagation. Backpropagation relies heavily on the chain rule of analysis as well as gradient descent established in the prior section about XGBoost. After forward propagation the error of the neural network is computed, using for example the residual sum of squares:

$$SSR = \sum_{i=1}^n (\hat{y}_i - z_j)^2 \quad (110)$$

where:

$SSR$  = Sum of Squares Residuals

$\hat{y}_i$  = Observed value

$z_j$  = Predicted value (Output of the neural network)

The optimization of the weights and biases is defined as the partial differentiation of  $SSR$  with respect to the weights and biases.

$$w_{(i+1)} = w_i - \eta \frac{\partial SSR}{\partial w_i} \quad (111)$$

$$b_{(i+1)} = b_i - \eta \frac{\partial SSR}{\partial b_i} \quad (112)$$

where:

$\eta$  = Learning rate

$SSR$  = Sum of Squares Residuals

The link between  $SSR$  and the weights and the biases is established with the function that denotes forward propagation (Equation 106). Therefore, using the chain rule of analysis

$$f(g(x)) = f'(g(x)) \cdot g'(x) \quad (113)$$

and applying it to the partial differentiation seen in Equations 111 and 112 the expression can be written as:

$$\frac{\partial SSR}{\partial w_i} = \frac{\partial SSR}{\partial z_j} \cdot \frac{\partial z_j}{\partial w_i} \quad (114)$$

Substituting the differentiation with the solution, the equation can be rewritten as follows:

$$\frac{\partial SSR}{\partial w_i} = -2 \sum_{i=1}^n (\hat{y}_i - z_j) \cdot x_i \quad (115)$$

Likewise the chain rule can be used to apply gradient descent on the optimization of biases (Starmer 2022):

$$\frac{\partial SSR}{\partial b_i} = -2 \sum_{i=1}^n (\hat{y}_i - z_j) \quad (116)$$

The complete process of forward and backpropagation is called epoch and after updating weights and biases the complete process is repeated until the  $SSR$  are close to 0, or some other criteria is met. The  $SSR$  is just used as an example and other loss functions like mean square error as well as binary cross entropy can be used as a substitute (Chollet et al. 2015; Pedregosa F et al. 2011).

### 3.5 Model Performance Measures

#### 3.5.1 Accuracy, Precision, Recall, F1-Score

The easiest way to measure the performance of a model is the accuracy, by comparing the true predictions to all instances or the complete dataset (Equation 117). The range of the accuracy ranges from 0%, where no prediction was correct, to 100%, where every prediction was correct. The limitation of the accuracy is that it can be misleading in case of imbalances in the dataset. For example, if 95% of the input dataset is the negative class and only 5% of the dataset is the positive class, a model that always predicts negative instances will have high accuracy but poor performance predicting positive instances. Precision measures the correctness of positive predictions (Equation 118). Defined as the ratio of true positive predictions (TP) and the sum of all positive prediction it is a measure that is particularly important, where the cost of a false positive predictions is high. A typical example would be spam mail detection. If an email is falsely classified as spam, the legitimate email would be missed.

Recall measures the performance of the model to classify all relevant cases (Equation 119). It is defined as the ratio of TP and the sum of all positive instances. It is especially relevant, in contrast to precision, when the cost of false negative is high. This is the case in medical diagnostics, where missing a diagnosis could lead to the absence of necessary medical treatment.

The F1-Score is the harmonic mean between precision and recall (Equation 120). It is needed when using imbalanced datasets or the cost of false positive and false negative predictions is similar.

$$Accuracy = \frac{TP + TN}{N + P} \quad (117)$$

$$Precision = \frac{TP}{TP + FP} \quad (118)$$

$$Recall = \frac{TP}{TP + FN} \quad (119)$$

$$F1 = 2 \cdot \frac{Precision \cdot Recall}{Precision + Recall} \quad (120)$$

with:

$TP$  = True positive counts

$TN$  = True negative counts

$FP$  = False Positive counts

$FN$  = False Negative counts

$N$  = All negative instances

$P$  = All positive instances

#### 3.5.2 Receiver Operating Characteristics Curve

Receiver Operating Characteristics (ROC) curves are an essential tool to evaluate the performance of a binary classification model. On the abscissa the false positive rate is depicted, whereas on the ordinate recall is depicted. Classification models outputs a probability score for every instance. By varying said threshold, the cutoff value that decides which instance is considered positive and negative is varied, resulting in different false positive and true positive rates. Each point in a ROC curve represents an FPR/TPR pair at a specific threshold. The diagonal line represents the output of a random guessing classifier. The top left corner (0, 1) connotes with a perfect model, with 100% sensitivity and 0% false positives. The area under the ROC curve (AUC) is calculated and represents the overall ability to

differentiate between positive and negative classes. An  $AUC=1$  is considered a perfect classification, an  $AUC = 0.5$  represents no discriminative power. And if the  $AUC$  falls below 0.5 than random guessing is superior to the model. In conclusion, ROC curves provide in detail understanding of the trade-off between TPR and FPR.

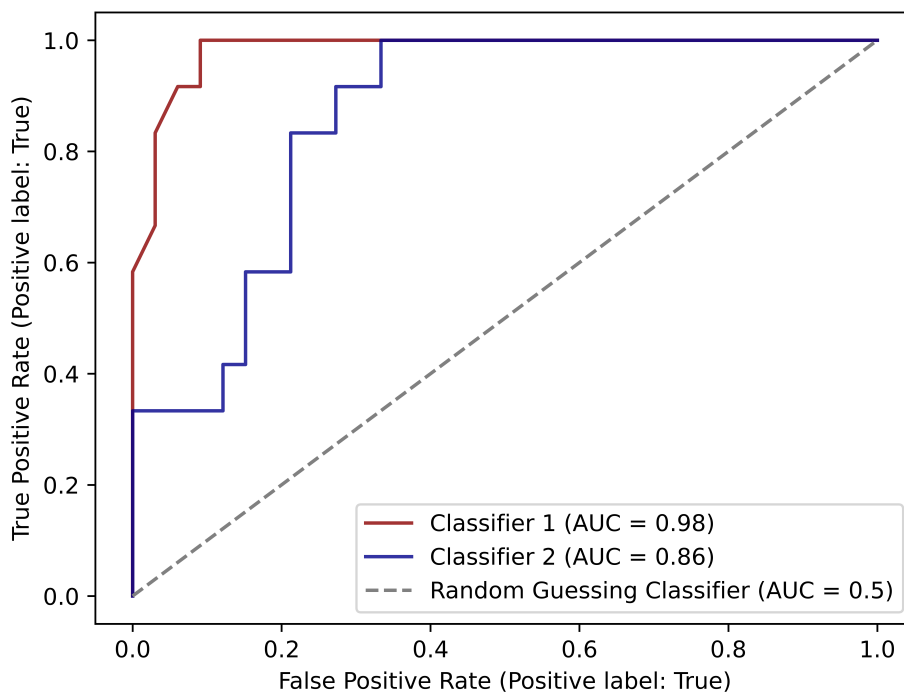


Figure 36: Example ROC curve for two classifiers with different performances. Performance of Classifier 1 is better than Classifier 2. The 45° line is considered a random guessing classifier.

## 4 Scope and aim

The first part of this work aims to improve the understanding of wine color and the link to human perception of wine color. Wine color as a parameter is mostly used for descriptive purposes either of a wine itself or wines in a sample set comparing it to the rest of the sample set. Furthermore, it is common practice to use the color measurement to measure the impact of enological parameter (Durner et al. 2010; Renner et al. 2022). It is also used as a proof of concept method to validate the results of a color evaluation obtained by descriptive analysis and internal quality assessment (Renner et al. 2022). There are two methods recommended by the OIV for the quantification of wine color. The Glories' color measurement that only uses three absorbance values to calculate the parameters color intensity, hue, and *Brilliance* and the CIE L\*a\*b\* color space, which uses the complete transmission spectra to evaluate the color. Regarding the CIE L\*a\*b\* color space, the protocol provided by the OIV is incomplete at best. The OIV recommends a data interval of 5 nm to measure the transmission spectra. Apart from that no requirements are given. It is unclear whether the data interval has an impact on the reproducibility of the CIE L\*a\*b\* measurement, nor is it clear if 5 nm is the best suited data interval. It is also unknown if other photometer parameter like 'scan speed' have an impact on the measurement. Therefore, the first objective was to evaluate if photometer settings like the 'data interval' or the 'scan speed' have an impact on the CIE L\*a\*b\* measurement.

Since Glories' color measurement and the CIE L\*a\*b\* color space are both recommended and widely used, they both have to be comparable in order to correctly describe color. Therefore, the second objective is the comparison of the Glories' color measurement and the CIE L\*a\*b\* color space with each other and the human perception of wine color. Glories color measurement was originally developed for dark red wines, but is now used for rosé, white, and Blanc de noir wines. CIE L\*a\*b\* was developed for color in general. However, it is uncertain if the CIE L\*a\*b\* color space actually represents the human perception of wine color, since the experimental setup leading to CIE L\*a\*b\* and the sensory analysis of wine color are completely different. According to Fairchild, the sensory of color, especially wine color is not well standardized (Fairchild 2018). Therefore, a new sensory method has to be developed. Additionally, the CIE L\*a\*b\* color space has an intrinsic method to compare color by calculating the color distance between two color points. Martinez et al. 2001 introduced an often cited just noticeable difference (JND) of 3 CIE L\*a\*b\* units. However, since then the CIE L\*a\*b\* color, especially the formula calculating the distance between two colors, have been modified and therefore the JND provided by Martinez et al. could not be up-to-date anymore. The OIV recommends the Euclidean color distance, which Martinez et al. used, whereas the CIE recommends the CIEDE2000 color distance formula. The CIEDE2000 color distance formula is not often used in wine analysis and the main reason for this is the absence of reference values for wine. So, the next objective is to re-evaluate the JND under the modernized aspects of the color distance calculation with the goal to provide new references whether a color distance in wine is visible or not.

The second part of the dissertation aims to find new applications of spectrophotometric data or color data in the wine industry asking whether it is possible to extract more information from spectrophotometric data. In other fields of research like medicine, spectrophotometric data in combination with ML algorithms is used for diagnostic purposes (Hsich 2011). Transferring this to wine research a possible new application could be the classification of different wine styles with the use of spectrophotometric data as input data to train an algorithm for wine quality management. In wine research, spectrophotometric data is not yet used to train machine learning algorithms. The focus of wine related machine learning applications focuses

on the prediction is subjective parameters like the prediction of wine quality. One study investigated the predictability of New Zealand Pinot Noir wines using a total of 47 volatile components and other parameters like pH, alcohol, anthocyanin levels. This study was conducted using a real dataset, that was enhanced by an oversampling algorithm, creating a mixture of real and synthetic data (Bhardwaj et al. 2022). Two other studies tried to predict the quality of red wine using the Machine Learning Repository of the University College London (UCL). Different algorithms like Naive Bayes, Random Forest, SVM, and MLP of which the best performance yielded the MLP model (Kumar et al. 2020; Shaw et al. 2020). However, it is not known, how spectrophotometric data performs and how color data can be utilized as input data in wine research.

The evaluated classification algorithms include logistic regression as the least complex algorithm. Increasing the complexity, SVM is also tested during these studies, investigating the performance of different kernel functions and other hyperparameter settings. Increasing the complexity again, XGBoost and Neural networks are evaluated for their use in wine color analysis. Due to their complexity, transformed methods like Glories' color measurement or CIE  $L^*a^*b^*$  cannot be used with these algorithms. Therefore, principle component analysis (PCA) can be used to reduce the dimensionality of transmission/absorption spectra. One target is to develop an ML application to classify commercial Blanc de noir compared to rosé and white wines. Blanc de noir wines are subject to legal regulations and the process of approving a Blanc de noir wine is regulated by visual color determination, a completely subjective parameter (Bundesrat Drucksache 175/21 2021). To classify the wines according to their spectrophotometric properties would objectify this process. Another problem regarding academic efforts in machine learning is that it usually ends with published results. To apply the results of these studies, the user still needs to acquire knowledge in different coding languages, which is not user-friendly. Therefore, to make the classification model accessible to wine chemists and winemakers, a browser application will be developed. This includes a dashboard that compares the spectrophotometric data of unknown wine samples to a given dataset of commercial wines. Another more difficult target is the classification of red wine. The goal of the associated work is to determine whether or not it is possible to differentiate barrel-aged red wines from non-barrel aged red wines. Due to the barrel, the oxygen uptake is higher than in a regular steel tank (INOX), and oxidation is linked to browning reactions in wine (Waterhouse 2002). It is well known, that the color of a wine is directly linked to the expectations of a wine (Ballester et al. 2009; Morrot et al. 2001; Nguyen and Durner 2023; Brochet and Dubourdieu 2001; Wang and Spence 2019). Therefore, to classify barrel-aged wine correctly would help the winemaker to assure the quality of the product, as they could assess if their barrel-aged wines match the color of other barrel-aged wines and therefore match the expectation of the customer. Another objective of this is to outline the limits of spectrophotometric data as input for classification via machine learning.



## 5 Cumulative part of the dissertation

### 5.1 Specification and Simplification of Analytical Methods to Determine Wine Color

The protocols recommended by the OIV to obtain CIE  $L^*a^*b^*$  coordinates are rather vague. Besides a 5 nm data interval set on the photometer to measure the transmission spectra, no other parameters are recommended. Furthermore it is unclear if detector systems and photometer settings have an effect on the CIE  $L^*a^*b^*$  calculation. In this study, the impact of data interval (0.5 nm, 1 nm, and 5 nm) and scan speed (100 nm/min and 1000 nm/min) is evaluated under the parameters distinguishability and reproducibility. The experiments were carried out on two photometers. One photometer used a Rowland off-circle monochromator and a silicon photodiode as detector, whereas the other one used a Czerny-Turner monochromator with a PMT as detector. It was shown that the scan speed and data interval have no impact on the distinguishability between the wines, but an impact on the reproducibility between the replication. In both photometers, the scan speed of 1000 nm/min combined with a data interval of 1 nm provided the most reproducible measurement.

Furthermore, since the CIE  $L^*a^*b^*$  color space and the Glories method are both recommended by the OIV, their comparability was investigated. In dark red wines both methods are comparable and correlate well with each other. However, the results deviate from each other in lighter wines with an  $L^* > 20$  and white wine. Therefore, the methods cannot be used interchangeably.

The third objective of the paper, of which the work was performed by the Chair of Bioprocess engineering was to simplify the color measurement, implementing the CIE  $L^*a^*b^*$  color space into a portable analysis system. This was done, by comparing different interpolation methods to predict a transmission spectrum from eight different transmission values. The predicted spectrum was used to calculate the CIE  $L^*a^*b^*$  coordinates. Cubic splines showed the lowest difference between measured and predicted CIE  $L^*a^*b^*$  coordinates.

## Specification and Simplification of Analytical Methods to Determine Wine Color

Authors:

Marcel Hensel<sup>1\*</sup>, Sarah Di Nonno<sup>2\*</sup>, Yannick Mayer<sup>2</sup>, Marina Scheiermann<sup>1</sup>, Jörg Fahrer<sup>3</sup>, Dominik Durner<sup>1</sup>, Roland Ulber<sup>2</sup>

Affiliations:

<sup>1</sup> Weincampus Neustadt, Dienstleistungszentrum ländlicher Raum Rheinpfalz, 67435 Neustadt an der Weinstraße, Germany

<sup>2</sup> Chair of Bioprocess Engineering, University of Kaiserslautern-Landau, 67663 Kaiserslautern Germany

<sup>3</sup> Division of Food Chemistry and Toxicology, Department of Chemistry, University of Kaiserslautern-Landau, 67663 Kaiserslautern, Germany

**Disclaimer:**

**\* These authors contributed equally to this work. Work presented in chapters 2.4 and 3.3 were exclusively performed by the Chair of Bioprocess Engineering, RPTU (Sarah Di Nonno and Yannick Mayer). Chapters 2.1, 2.2, 2.3, 3.1, and 3.2 were exclusively performed by the DLR Rheinpfalz (Marcel Hensel and master thesis Marina Scheiermann). For more information, refer to the author contribution section.**

Citation:

Hensel, M.; Di Nonno, S.; Mayer, Y.; Scheiermann, M.; Fahrer, J.; Durner, D.; Ulber, R. Specification and Simplification of Analytical Methods to Determine Wine Color. *Processes* 2022, 10, 2707. <https://doi.org/10.3390/pr10122707>

Copyright: ©2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license.

## Article

# Specification and Simplification of Analytical Methods to Determine Wine Color

Marcel Hensel <sup>1,†</sup> , Sarah Di Nonno <sup>2,†</sup> , Yannick Mayer <sup>2</sup>, Marina Scheiermann <sup>1</sup>, Jörg Fahrer <sup>3</sup> , Dominik Durner <sup>1</sup> and Roland Ulber <sup>2,\*</sup> 

<sup>1</sup> Weincampus Neustadt, Dienstleistungszentrum Ländlicher Raum Rheinpfalz, 67435 Neustadt an der Weinstraße, Germany

<sup>2</sup> Chair of Bioprocess Engineering, Technical University of Kaiserslautern, 67663 Kaiserslautern, Germany

<sup>3</sup> Division of Food Chemistry and Toxicology, Department of Chemistry, Technical University of Kaiserslautern, 67663 Kaiserslautern, Germany

\* Correspondence: ulber@mv.uni-kl.de

† These authors contributed equally to this work.

**Abstract:** The color of wine is an important quality parameter essential for the first impression of consumers. The International Organization of Vine and Wine (OIV) recommends two methods to describe wine color: color calculation according to Glories and the determination of coordinates in the CIE L\*a\*b\* color space. The measurement of wine color is often not feasible for winemakers because the required instrumentation is expensive and bulky. In this study, the influence of photometer settings on the calculated color was investigated based on 14 wines. Furthermore, the CIE L\*a\*b\* and Glories system were compared using 56 red and 56 white wines. Photometer settings were found to influence the reproducibility of color determination. In addition, CIE L\*a\*b\* system do not correlate in all wines with the Glories system and Glories probably provides less information about wine color. Using interpolation, CIE L\*a\*b\* coordinates were calculated from single wavelength measurements taken by a small-sized and inexpensive portable analysis system, which could be used by winemakers in the future.

**Keywords:** wine color; CIE L\*a\*b\*; Glories, OIV; Photometry; portable analysis system; Lagrange interpolation; cubic splines; Sprague interpolation



check for updates

**Citation:** Hensel, M.; Di Nonno, S.; Mayer, Y.; Scheiermann, M.; Fahrer, J.; Durner, D.; Ulber, R. Specification and Simplification of Analytical Methods to Determine Wine Color. *Processes* **2022**, *10*, 2707. <https://doi.org/10.3390/pr10122707>

Academic Editor: Cristina Soares

Received: 7 November 2022

Accepted: 4 December 2022

Published: 15 December 2022

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Color has a significant influence on the perceived quality of wine, altering the perception of aroma, taste and mouthfeel [1]. White wine dyed red was described using the olfactory attributes typical for red wine in contrast to the undyed white wine, which was perceived as having the olfactory attributes typical for white wine [2]. Given its influence on consumer perception, wineries aim to manage the wine color already during the production process to match consumer expectations with their final products. Objective methods and reliable devices are necessary to measure wine color during the winemaking process.

According to Grassmann's laws, color can be described using three independent parameters [3]. These can be a primary color, the color intensity and the white intensity, or three primary valances, e.g., red, green, and blue [3]. There are two established methods to describe the color of wine. The first is the Glories method, which uses absorbance values at wavelengths of 420 nm, 520 nm and 620 nm, leading to a yellow, red and blue color impression for the observer [4,5]. However, variance of the human eye and the influence of surrounding light on color perception are not considered in the Glories calculations. Because of this, the Commission Internationale de L'Eclairage (CIE) defined the CIE standard system, which is a color space based on the color coordinates of X, Y and Z. The calculation of these coordinates is more complex, requiring a complete transmission spectrum in the

visible range (380–780 nm). Furthermore, a standard observer [6] and a standard illuminant [7] are needed for calculation, which include the influence of the sensitivity of the human eye for different colors and the surrounding light on color perception. The standard observer is based on experimental work [8–11], where the spectral sensitivity of the eye for the three primary valances red, green and blue was investigated. The standard illuminant defines power distribution of standardized light sources, for example natural daylight on a sunny day [7]. The CIE XYZ system allows the objective and reproducible determinations of color, but colors in this system are not visually equidistant [12], meaning the Euclidean distance ( $\Delta E$ ) between two colors in this space does not correlate with the perceived color difference [13]. The CIE  $L^*a^*b^*$  color space aims to achieve visual equidistance [12]. The International Organization of Vine and Wine (OIV) recommends the Glories and the CIE  $L^*a^*b^*$  system for color determination in wine [14,15].

Regarding the CIE  $L^*a^*b^*$  system, the OIV recommends to record transmission spectra in 5 nm steps, in the following mentioned as data intervals. However, it is unclear whether a data interval of 5 nm is optimal. While the CIE  $L^*a^*b^*$  method, as described by the OIV, is popular in wine research, other versions of this method are used. Publications throughout the years have used data intervals of 1 nm, 2 nm or 10 nm for transmission measurements [16–18]. To evaluate this, an Analysis of Variance (ANOVA) and Bartlett's test was used. No other requirements for technical parameters such as scan speed are given. It is not clear whether these photometer settings have an influence on the resolution, i.e., the ability to distinguish samples, and the reproducibility of the measurement.

Methods to assess wine color require expensive equipment as well as technical personnel to operate and maintain it in a laboratory environment. Therefore, a cheaper and easier, yet reliable, method to measure and calculate the color of wine is needed. In recent years, interest among wine producers in portable photometers has grown. Two basic modes of operation are available. The first mode of operation includes a broadband light source (e.g., white light LEDs, tungsten) and light dispersing elements such as prisms [19,20] and diffraction gratings [21], which are often combined with other optical elements, including lenses [20,22–35] or light guides [19]. This first mode of operation allows the capture of a complete transmission spectrum, though the optical elements required are expensive and must be aligned accurately. The second mode of operation uses a light source at a specific wavelength, such as a light emitting diode (LED) [36–39]. With this light source, an inexpensive light sensor, like a photodiode [37], -transistor [38] or -resistor [39] can be used. This second mode of operation is less costly, but such a photometer cannot be used to record a complete transmission spectrum, which is why it cannot be used for the CIE  $L^*a^*b^*$  color calculation, as recommended by the OIV without further data processing. To calculate the CIE  $L^*a^*b^*$  coordinates from single wavelength measurements based on empirical methods are based on outdated standards, for example another standard observer or illuminant. Furthermore, these methods are not applicable in the same extent for every wine.

The overall objectives of this study are to specify technical parameters for reproducible wine color determination, to compare different methods for wine color determination, and to implement the CIE  $L^*a^*b^*$  coordinate determination without expensive and bulky equipment. The first goal was to examine the influence of the photometer 'data interval', which must be set to record transmission spectra for the calculation of the CIE  $L^*a^*b^*$  coordinates. Furthermore, the influence of the photometer setting 'scan speed' was investigated. The second goal was to determine whether the Glories and CIE  $L^*a^*b^*$  systems are correlated, since both are commonly used and recommended. Older publications, which correlated the Glories method with the CIE  $L^*a^*b^*$  color space, used the CIE 1931 2°-standard observer, the standard illuminant C or different pathways of cuvettes [17,40]. The current standard is the CIE 1964 10°-standard observer in combination with the standard illuminant D65, referred to a 10 mm pathway cuvette [12]. The third goal of this study was to develop a calculation method to determine CIE  $L^*a^*b^*$  coordinates based on single wavelength transmission measurements instead of using a complete transmission spectrum. Approaches to derive CIE  $L^*a^*b^*$  coordinates from these measurements were investigated either by

direct correlation or by interpolation. For direct correlation, empirical methods after Ayala et al. [41] and Hardy et al. [42] were performed. Furthermore, different interpolation methods, namely Cubic splines and Lagrange and Sprague interpolation were investigated". To reach the best approximation to the method according to the OIV, the calculated values were compared with data from the established laboratory method. The calculation methods were transferred to an inexpensive portable analysis system, which should allow portable on-site measurements.

## 2. Materials and Methods

### 2.1. Wines

The sample set consisted of 56 commercial red wines (Table S1) and 56 commercial white wines (Table S2). The red wines were produced from eight grape varieties, grown in eight countries during the vintages 2012 to 2021. The white wines were from six varieties, seven countries, and vintages from 2013 to 2021. Prices were between 5 and 20 €, a range considered average for wine on an international level [43,44]. The wines were selected to consider major wine producing countries [45], important grape varieties [46], and vintages that are currently available in the commercial wine trade, with the goal of representing the diversity of wine color.

### 2.2. Specification of Photometer Settings to Obtain CIE L\*a\*b\* Coordinates

To specify technical parameters for wine color measurement, spectra from 300 to 900 nm of seven Merlot wines and seven Chardonnay wines were recorded with the data intervals of 0.5, 1, and 5 nm in combination with scan speeds of 100 and 1000 nm/min. Further technical parameters are given in Table S3. The red wines were measured in a 1 mm flow cuvette and the white wines in a 10 mm cuvette. To comply with the Lambert Beer law and with the recommendation of the CIE and OIV [7], the red wine spectra were corrected to a 10 mm pathway before CIE L\*a\*b\* calculation. Triplicate measurements were conducted on a double beam photometer (V-730, JASCO, Tokyo, Japan). The CIE L\*a\*b\* coordinates were extracted from the spectra between 380 and 780 nm in combination with the 10°-standard observer and D65 standard illuminant according to the CIE [7]. Experiments were repeated on another double beam photometer with a different monochromator and a different detector system (Varian Cary 100, Agilent, Santa Clara, CA, USA) but with comparable settings to investigate the influence of different photometers on the color calculation of wine.

The individual CIE L\*a\*b\* coordinates of the seven wines of one grape variety were averaged, yielding the mean CIE L\*a\*b\* coordinates for the grape variety (Equation (1)).

$$\bar{L}^* = \frac{\sum L_i^*}{n}, \bar{a}^* = \frac{\sum a_i^*}{n}, \bar{b}^* = \frac{\sum b_i^*}{n} \quad (1)$$

For each wine,  $\Delta E$  was then calculated from its individual CIE L\*a\*b\* coordinates and the mean CIE L\*a\*b\* coordinates for the variety (Equation (2)).

$$\Delta E = \sqrt{(\bar{L}^* - L_i^*)^2 + (\bar{a}^* - a_i^*)^2 + (\bar{b}^* - b_i^*)^2} \quad (2)$$

To assess the influence of data interval and scan speed on the resolution, which is here the ability to distinguish wines by color, and to assess the reproducibility of measurements, the sum of squared deviations (SSD) was calculated for every repetition separately (Equation (3)). Subsequently, the mean SSD ( $n = 3$ ) was calculated reflecting the ability to distinguish wines by color and the standard deviation reflects the reproducibility of the measurements (Equation (3)).

$$SSD(\bar{\Delta E}) = \sum (\Delta E_i - \bar{\Delta E})^2 \quad (3)$$

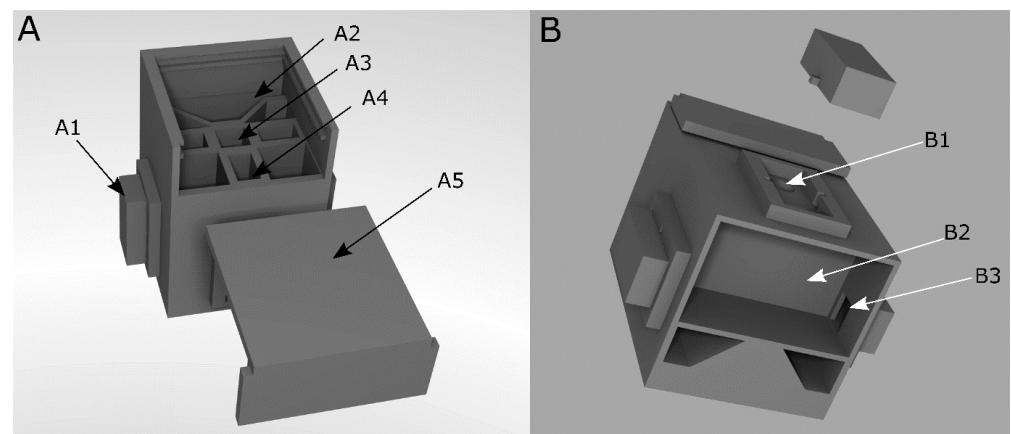
### 2.3. Comparison between CIE $L^*a^*b^*$ Color Space and Glories Color Measurement

Spectra from 300 to 900 nm of all 112 wines were recorded on a double beam photometer (V-730, JASCO, Tokyo, Japan) using a data interval of 1 nm in combination with a scan speed of 1000 nm/min. CIE  $L^*a^*b^*$  coordinates were calculated, as described earlier. The 56 red wines were divided into two lots discriminated by their  $L^*$ -values as light red wines ( $L^* > 20$ ,  $n = 22$ ) and dark red wines ( $L^* < 20$ ,  $n = 34$ ). The Glories absorbance values at wavelengths 420 nm, 520 nm and 620 nm were extracted from the same spectra used to calculate CIE  $L^*a^*b^*$  coordinates. Data from these two systems was correlated using Spearman coefficients, and the comparison between the CIE  $L^*a^*b^*$  color space and the Glories color measurement was visualized using heatmaps.

### 2.4. Portable Analysis System

#### 2.4.1. Measurement Chamber

A portable analysis system, as described earlier [47], was used for transmission measurements to simplify the wine color determination. The measurement chamber (Figure 1) was 3D-printed in black acrylonitrile butadiene styrene (ABSplus™, Stratasys GmbH, Frankfurt am Main, Germany). To avoid using optical filters or dispersing elements, LEDs of single wavelengths were integrated as the light source. Light was measured using a phototransistor. The measurement principle is based on the discharge of a capacitor by the photocurrent of the phototransistor, and the discharge time is correlated with light intensity and thus with the transmission of a sample. The in-detail description of the circuit was described earlier [47].



**Figure 1.** CAD model of the measurement chamber of the portable analysis system. (A): Front view. (B): Bottom view. A1: Lid to cover inserted LEDs and phototransistor, A2: Socket for SMD-LED board. A3: Place for a 10 mm-cuvette, A4: light path from LED to phototransistor, A5: Cover for the measuring chamber, B1: Place for LED or phototransistor B2: Bottom compartment for storage of a microcontroller and the circuit board, and B3: Path for cables to connect inserted LED and phototransistor [47].

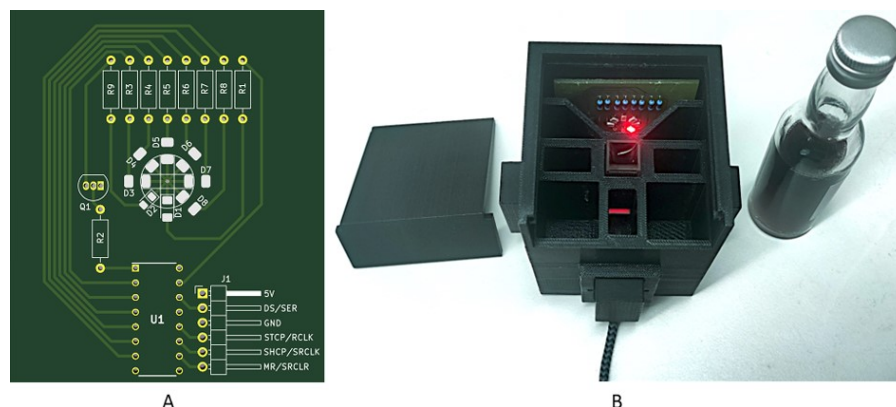
The LEDs and the phototransistor are placed in LED inserts by LED mounting clips (Figure 1(B1)). The circuit board is mounted in the bottom compartment of the measuring chamber (Figure 1(B2)). To perform transmission measurements, the LED is positioned opposite the light sensor. The sample is inserted into a cuvette in the middle of the measuring chamber between the light source and sensor. After the components and the sample has been inserted, the chamber is closed to avoid interference from ambient light (Figure 1(A1,A5)).

#### 2.4.2. Data Recording

For interpolation of the transmission spectra of wine, LEDs with the required wavelengths were integrated into the system. An LED board with all required surface-mounted-



device-LEDs (SMD-LEDs) was developed to avoid the installation and removal of LEDs during the measurement. The LEDs were controlled by the microcontroller via a shift register (Figure 2).



**Figure 2.** (A) Developed LED board for color measurement with the portable analysis system. Eight SMD-LEDs with the needed wavelengths (Table 1) were placed in a circle (D1–D4). The required series resistors were connected (R1–R9). The LEDs were controlled by a shift register (U1). The board was developed in KiCAD (KiCAD version 5.1.9, Jean Pierre Charras, France). The LED board can be placed into the portable analysis system (B).

For 10 white wines and 10 red wines, measurement data obtained using the portable analysis system was correlated with corresponding transmission values from the laboratory photometer. This is important because of the varying sensitivity of the phototransistor at different wavelengths [48] and with this, the different correlations for the measured values with the portable analysis system and the laboratory photometer for different wavelengths. Subsequently, transmission values of the remaining white and red wines were calculated. Red wines were measured in a 4 mm glass cuvette, and white wines were measured in a 10 mm glass cuvette. Afterwards, the spectra of wine were interpolated, and CIE  $L^*a^*b^*$  coordinates were calculated according to the OIV.

#### 2.4.3. Data Processing of Single Wavelength Measurements to Retrieve CIE $L^*a^*b^*$ Coordinates

In one approach, transmission values at single wavelengths were used directly in the empirical formulae to predict the CIE  $L^*a^*b^*$  coordinates. In another approach, the transmission spectra of wine samples were interpolated based on transmission values at single wavelengths. Completed spectra were then used to calculate CIE  $L^*a^*b^*$  coordinates. To determine the accuracy of each method,  $\Delta E$  was calculated between the predicted CIE  $L^*a^*b^*$  coordinates and those calculated according to the OIV (Equation (4)).

$$\Delta E = \sqrt{(L^*_{measured} - L^*_{predicted})^2 + (a^*_{measured} - a^*_{predicted})^2 + (b^*_{measured} - a^*_{predicted})^2} \quad (4)$$

These investigations were performed for 56 red and 56 white wines. Spectra of all 112 wines were recorded on a double beam photometer (V-730, JASCO, Tokyo, Japan) using a data interval of 1 nm in combination with a scan speed of 1000 nm/min.

#### Empirical Formulae to Calculate Wine Color

Two empirical formulae from Hardy et al. [42] and Ayala et al. [41] were compared. Both methods allow the calculation of CIE  $L^*a^*b^*$  coordinates from transmission values at single wavelengths. Hardy et al. prescribed the balanced ordinates method, according to which the CIE XYZ color valances are calculated by four transmission values at 445 nm,

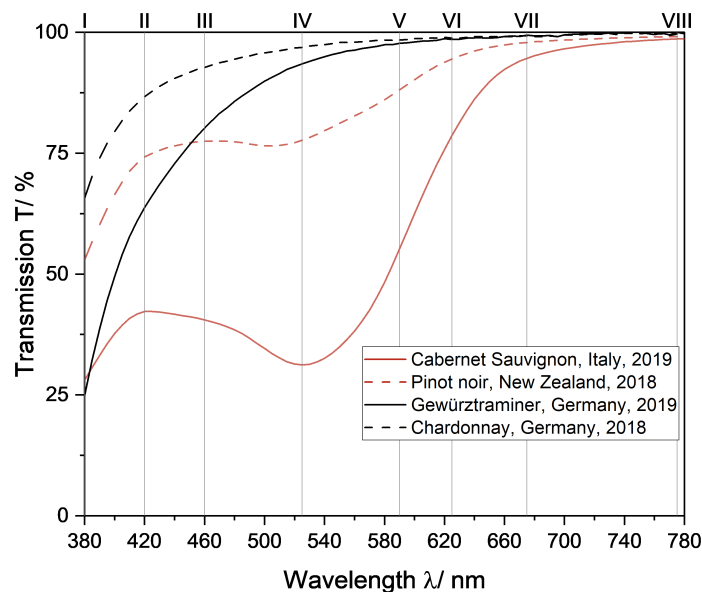
495 nm, 550 nm and 625 nm. These calculated color coordinates are related to the standard illuminant C and the 2°-standard observer as reference points. The method published by Ayala et al. is based on characteristic vector analysis, where CIE XYZ color valences are calculated from three transmission measurements at 440 nm, 530 nm and 600 nm. These color coordinates are related to the D65 standard illuminant, the 10°-standard observer, and a 2 mm path length. Both methods calculate the CIE XYZ color valences, which can be used to determine the CIE L\*a\*b coordinates [12].

#### Interpolation Methods

Three different interpolation methods were compared to generate complete transmission spectra subsequently used to calculate CIE L\*a\*b\* coordinates. Interpolation according to Lagrange [49] delivers a polynomial with a degree of  $n - 1$ , where  $n$  is the number of supporting points. With the interpolation according to Sprague [50], polynomials of the fifth degree are calculated to describe the course of a function between two supporting points. Consequently,  $n - 1$  polynomials are required to describe the function between  $n$  supporting points. For the calculation of the coefficients, data from outside the data set is needed, thus extrapolation of data is necessary. The interpolation with cubic splines [51] yields  $n - 1$  polynomials of third grade. Like the interpolation according to Sprague, the function between two supporting points is described by this polynomial. For interpolation with cubic splines, two assumptions must be made due to double overdetermination of the equation system. In this case, the first derivation at the upper and lower bounds of reconstructed spectra were set to one for 380 nm and zero for 780 nm.

#### Selection of Supporting Points for Interpolation

Eight supporting points were chosen at extreme points, inflection points, and other points of interest in the transmission spectra of red and white wine (Figure 3) for interpolation.



**Figure 3.** Transmission spectra of two red wines and two white wines to illustrate the choice of supporting points for color determination based on interpolation of transmission spectra. The position of supporting points is marked by numbers (I—VIII, Table 1)

The transmission spectra from red wine show a local maximum in the range of 400–440 nm, an inflection point in the range of 440–510 nm, a local minimum in the range of 510–560 nm, an inflection point in the range of 560–620 nm and a plateau in the range of 620–650 nm. To ensure the spectra were interpolated correctly between these ranges, supporting points were selected in these regions. Therefore, considering the commercial



availability of LEDs, the transmission of wine samples was measured at the wavelengths 420 nm, 525 nm, 560 nm and 625 nm (Table 1 II, IV, V and VI). Because the distance between supporting point VI and the end of the spectrum is quite large and would likely cause errors, an additional supporting point was added in this range for greater accuracy of interpolated spectra (Table 1, VII). Lastly, supporting points at the upper and lower bounds of the transmission spectra were chosen, since the slope here varied from wine to wine (Figure 3 and Table 1, I and VIII). No additional LED was chosen between VII and VIII because the transmission in this region does not change from stable without any further points of interest.

**Table 1.** Supporting points chosen in transmission spectra for color determination of red and white wine based on interpolation methods. Supporting points were chosen based on transmission spectra of red wine (Figure 3(I–VIII)).

Number	Region	Wavelength Region	Wavelength of Commercially Available LED
I	Lower bound ultraviolet region	380 nm	380 nm
II	Local maximum	400–440 nm	420 nm
III	Gap filler <sup>1</sup>	440–510 nm	460 nm
IV	Local minimum	510–560 nm	525 nm
V	Gap filler <sup>1</sup>	560–620 nm	590 nm
VI	Strong increasing slope	620–650 nm	625 nm
VII	Gap filler <sup>1</sup>	650–775 nm	675 nm
VIII	Edge infrared region	780 nm	775 nm

<sup>1</sup> Gap fillers to bridge areas in transmission spectra.

### 2.5. Statistical Analysis

Calculations using the empirical methods and Lagrange interpolation were conducted in Excel (Microsoft, Redmond, WA, USA). Cubic splines and Sprague interpolation were implemented in MATLAB (R2020b, MathWorks, USA). To evaluate the normal distribution of data, the Shapiro-Wilk test was conducted in Origin (Pro) (2020b, OriginLab, USA), where  $H_0$  states that the data follows a normal distribution. To verify significance of the results, an ANOVA, where the  $H_0$  is accepted if there are no differences between the mean values of the observed factor, was conducted in combination with Tukey's HSD post hoc test. Here, the  $H_0$  states that two observed means are equal. Both tests were also conducted in Origin(Pro) (2020b, OriginLab, USA). XLSTAT (2020, ADDINSOFT, France) was used for Bartlett's test to verify the significance between variances.  $H_0$  is accepted, if the variances are equal. Correlation matrices depicting Spearman coefficients were programmed in Python using the *pandas* [52], *numpy* [53], *matplotlib* [54] and *seaborn* [55] libraries.

## 3. Results and Discussion

### 3.1. Influence of Photometer Settings on the Ability to Distinguish Wines and to Obtain Reproducible CIE $L^*a^*b^*$ Coordinates

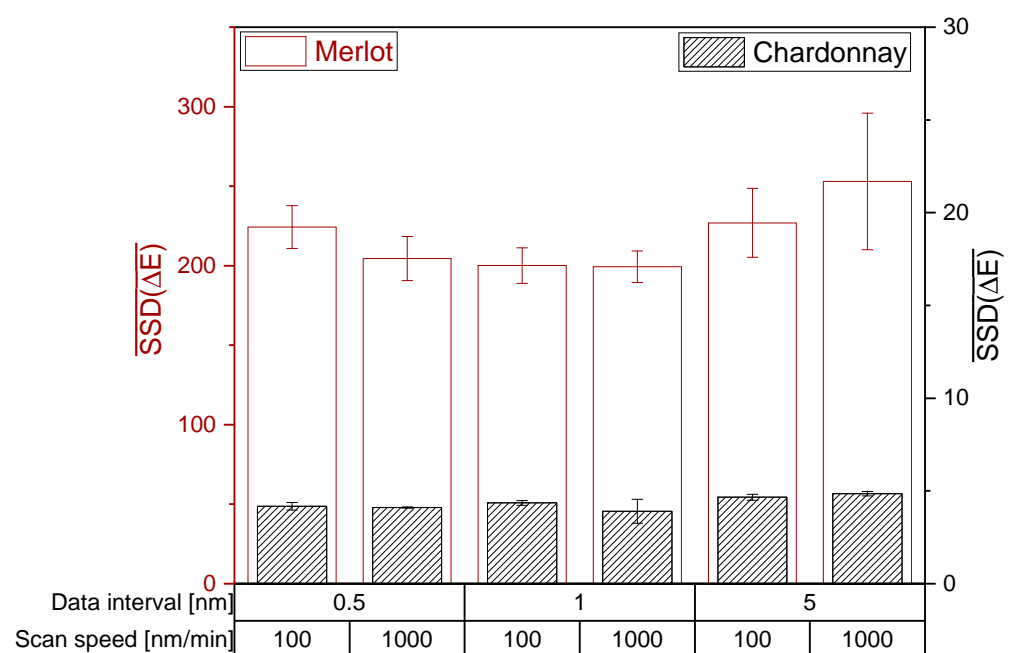
For both photometers,  $H_0$  of the ANOVA and Tukey HSD test is accepted and therefore the setting of the data interval and scan speed did not significantly affect the mean SSD ( $\Delta E$ ). Regarding Bartlett's test,  $H_0$  is rejected and the photometer settings have a significant impact on the variances of SSD( $\Delta E$ ). Hence, the investigated photometer settings did not affect the ability to distinguish wines by the CIE  $L^*a^*b^*$  coordinates of red and white wines (Figures 4 and 5). However, a high data interval of 5 nm in red wine resulted in the lowest reproducibility. Accordingly, data intervals lower than 5 nm are recommended for the CIE  $L^*a^*b^*$  measurement of red wines.

The investigation of data interval and scan speed on different photometers revealed similar mean SSD( $\Delta E$ ) for red wines, but not for white wines. The different mean SSD( $\Delta E$ ) between the photometers suggest that different devices have different sensitivities in their detector systems in the limit of detection area, as demonstrated for white wine with unsaturated color. For red wines, the mean sum of squared deviations showed only

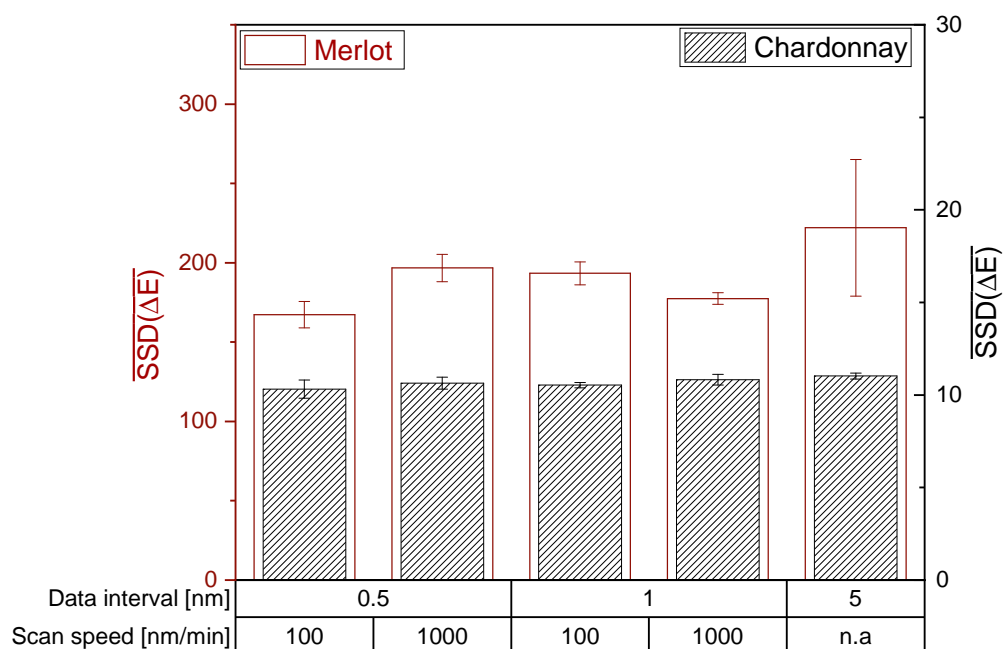
differences for an extremely low data interval of 0.5 nm. Depending on the machine and its detector system, an extremely low data interval could lead to increased noise, consequently reducing the ability to distinguish the wine color. Therefore, a data interval higher than 0.5 nm is recommended.

For red wines, the mean sum of squared deviations showed only differences for an extremely low data interval of 0.5 nm. Depending on the machine and its detector system, an extremely low data interval could lead to increased noise, consequently reducing the ability to distinguish the wine color. Therefore, a data interval higher than 0.5 nm is recommended.

As for the scan speed, no differences between 100 and 1000 nm/min could be observed. Overall, data recording with 1 nm data interval and the faster scan speed of 1000 nm/min is suggested to obtain the best results for CIE L\*a\*b\* coordinates.



**Figure 4.** Influence of the photometer settings data interval and scan speed on the CIE L\*a\*b\* coordinates on the JASCO V-730 double beam photometer. The mean sum of squared deviations are shown with error bars (SD; n = 3;  $\alpha$  = 0.05).



**Figure 5.** Influence of the photometer settings data interval and scan speed on the CIE  $L^*a^*b^*$  coordinates on the Varian Cary 100 double beam photometer. The mean sum of squared deviation is shown with error bars (SD;  $n = 3$ ;  $\alpha = 0.05$ ).

### 3.2. Correlation between CIE $L^*a^*b^*$ and Glories Method

After determining the specific photometer settings that yield reproducible CIE  $L^*a^*b^*$  coordinates, the spectra of 56 red wines and 56 white wines were recorded to examine whether the CIE  $L^*a^*b^*$  coordinates are comparable to Glories absorbance values. A good comparability between the two methods would prefer the Glories method over the CIE  $L^*a^*b^*$  color space, since the Glories method does not require a complete transmission spectrum. The measured CIE  $L^*a^*b^*$  coordinates and Glories parameters were correlated in a correlation matrix displaying the Spearman coefficient. For dark red wine, the CIE  $L^*a^*b^*$  coordinates and Glories absorbance values show a strong negative correlation (Figure 6A(I)). Additionally, the  $L^*a^*b^*$  coordinates correlate positively with itself (Figure 6A(II)) as well as the Glories absorbance values at wavelengths 420 nm, 520 nm and 620 nm (Figure 6A(III)).

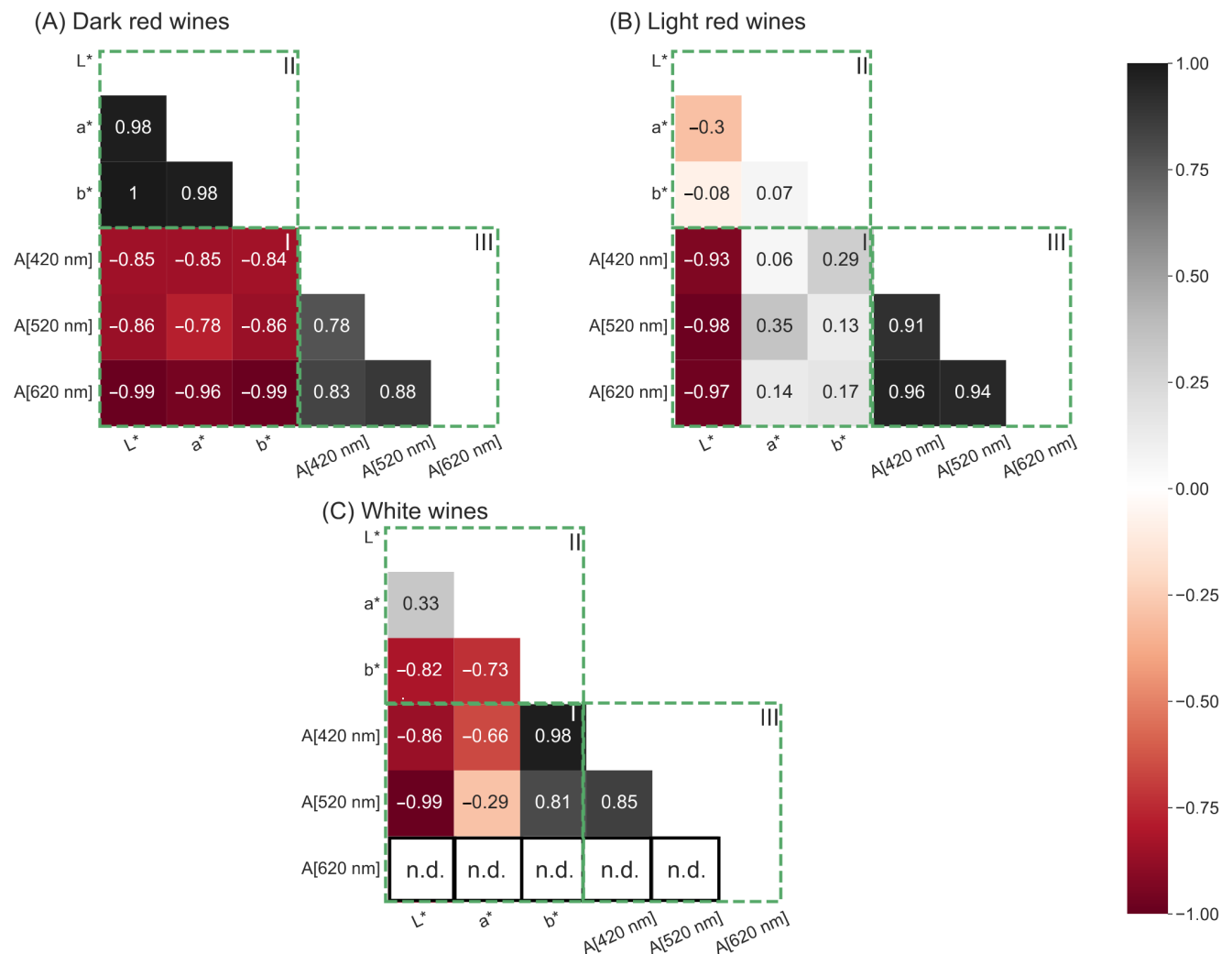
A different observation was made for light red wine (Figure 6B). The coefficients in Figure 6B(I) are very low, indicating a weak correlation between the CIE  $L^*a^*b^*$  coordinates and the Glories absorbance values. Furthermore, a weak correlation between the  $L^*$ ,  $a^*$ , and  $b^*$  coordinates was observed (Figure 6B(II)), while the Glories absorbance values demonstrate a strong positive correlation with each other (Figure 6B(III)).

Visualizing the Spearman coefficients for white wine, a strong negative correlation between the  $L^*$  and  $b^*$  coordinates and the Glories absorbance values was observed (Figure 6C(I)). However, the  $a^*$  coordinate showed a weak correlation with all three Glories absorbance values (Figure 6C(I)). In white wine, the  $L^*$ , and the  $b^*$  coordinate correlated positively with each other (Figure 6C(II)). Furthermore, the  $a^*$ , and the  $b^*$  coordinate correlated positively with each other, but not the  $L^*$  and  $a^*$  coordinate (Figure 6C(II)). The Glories absorbance values, correlated well with each other (Figure 6C(III)). However, in white wine the absorbance at 620 nm was under the photometric accuracy and therefore could not be used in the correlation (Figure 6C(III)).

The observations demonstrate that the Glories and the CIE  $L^*a^*b^*$  method can be used interchangeably for dark red wine, but not for white wine and light red wine. The Glories absorbance values in dark red, light red, and white wines correlated strongly among themselves, indicating a violation of Grassmann's first law, which states that three independent parameters are needed to completely describe color. In dark red wines, the

CIE  $L^*a^*b^*$  coordinates correlate with each other as well, but since this is not the case in white wines and light red wines, it could be argued that the CIE  $L^*a^*b^*$  coordinates are more independent than the Glories absorbance values. Therefore, the CIE  $L^*a^*b^*$  color space is better suited for use in wine than color measurement, according to Glories.

Additionally, the Glories color measurement violates Grassmann's first law in white wine because, without absorbance values at 620 nm, the remaining two absorbances at 420 and 520 nm cannot sufficiently describe color.

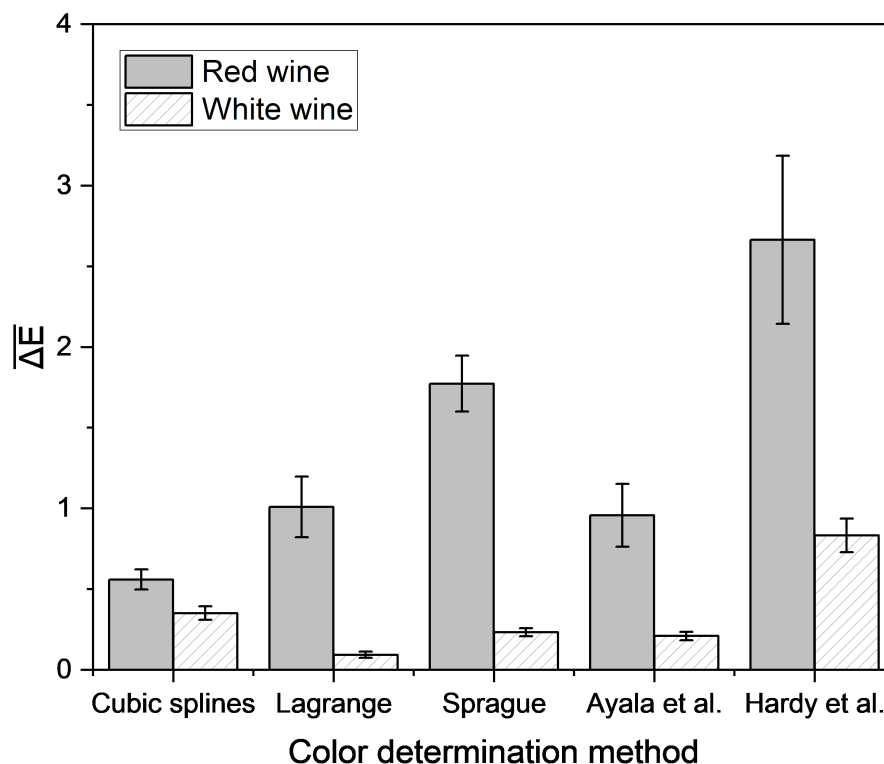


**Figure 6.** Correlation matrices depicting the Spearman coefficients for dark red wine ((A),  $n = 34$ ), light red wine ((B),  $n = 22$ ), and white wine ((C),  $n = 56$ ). Evaluated were the correlations between CIE  $L^*a^*b^*$  coordinates and Glories absorbance values at wavelengths 420 nm, 520 nm and 620 nm (A[420 nm], A[520 nm], A[620 nm]) (I), between  $L^*$ ,  $a^*$ , and  $b^*$  coordinates (II), and between Glories absorbance values (III).

### 3.3. Simplification of Wine Color Measurement

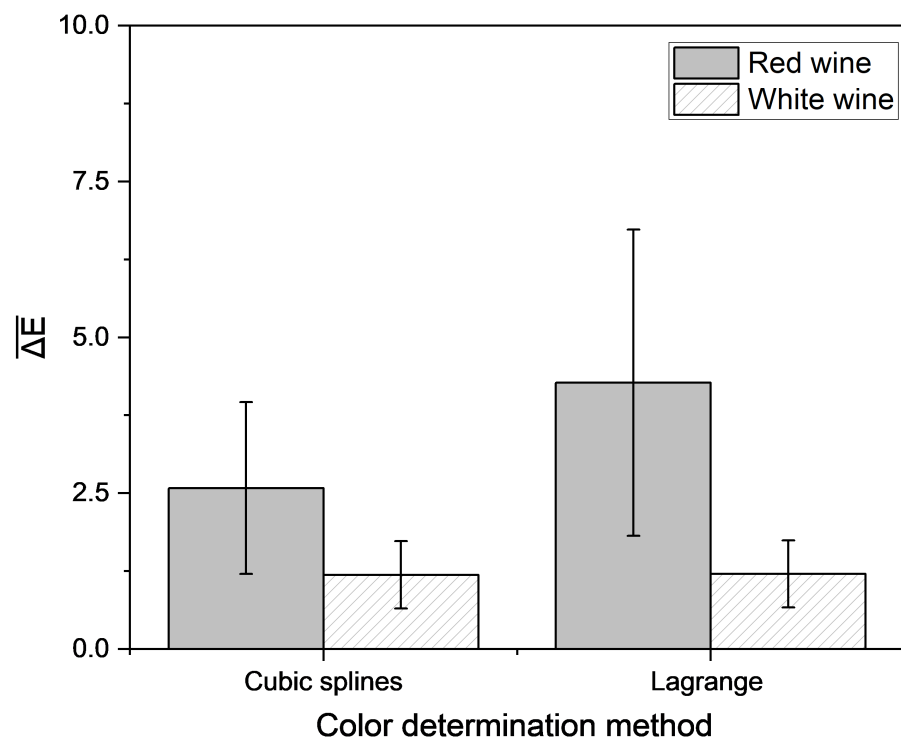
Different empirical and interpolation methods were investigated to determine the color of wine using single wavelength transmission values measured using a laboratory photometer. For each wine, CIE  $L^*a^*b^*$  coordinates were calculated from complete spectra and from spectra derived from single wavelength transmission values via empirical and interpolation methods. The accuracy of the derived CIE  $L^*a^*b^*$  coordinates was determined

by their  $\Delta E$  in relation to the CIE L\*a\*b\* coordinates calculated from the complete spectra (Figure 7).



**Figure 7.** Mean Euclidean color distance in the CIE L\*a\*b\* color space based on different empirical methods and the reconstructed transmission spectra, determined by different interpolation approaches compared with the color determination method according to OIV [51]. Transmission data were measured with a laboratory photometer. Mean and 95% confidence interval of 56 red and 56 white wines are shown.

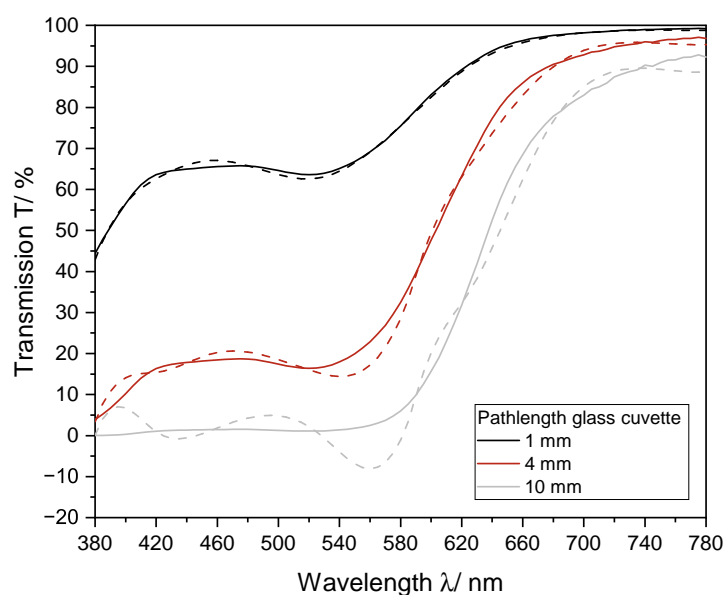
For red wine, the lowest color difference between measured and predicted CIE L\*a\*b\* coordinates was achieved using the interpolation via cubic splines. The method of Hardy et al. yielded the highest color difference for both red and white wine. Regardless of the method, the color difference for white wine was lower than for red wine. For white wine, using the method recommended by the OIV, the lowest deviation was attained by the interpolation of spectra according to Lagrange. In general, a color difference under 3 is not perceivable to the observer [56], but there are different data for the threshold value for visible color, ranging from 1 to 14 [57]. This suggests that differences between results obtained using these color determination methods and those from the official method of the OIV likely cannot be perceived by the observer, except for the method according to Hardy et al. for red wines. Therefore, the methods are suitable for the determination of wine color. Because the interpolation with cubic splines and interpolation according to Lagrange demonstrated the best results for red and white wine, these methods were integrated into the portable analysis system. The difference in color measured using this system and the color calculated using the OIV method was determined (Figure 8).



**Figure 8.** Mean Euclidean color distance in the CIE L\*a\*b\* color space based on different empirical methods and the reconstructed transmission spectra, determined by different interpolation approaches compared with the color determination method according to OIV [51]. Transmission data were measured with a portable photometer and compared to data measured with the laboratory photometer. Mean and 95% confidence interval of 46 red and 46 white wines are shown.

The Euclidean color distance determined by the portable analysis system was higher as compared to that determined by the laboratory photometer alone. To establish the interpolation methods, the required transmission values were extracted directly from the measured spectra. Therefore, the same data set was used for all color determination methods. For implementation into the portable analysis system, the measured data had to be correlated with the transmission at the laboratory photometer in an extra step to be suited for color determination. In contrast to the laboratory photometer, a 4 mm cuvette was used for the color calculation with the portable analysis system. It could be shown that the use of a cuvette with a longer pathlength leads to a greater deviation of the interpolated spectra from the spectra, recorded with the laboratory device (Figure 9). Therefore, an adapter for usage of 1 mm cuvette should be integrated.

The lowest Euclidean color distance, calculated with the portable analysis system from the color measured with the laboratory photometer for red wine resulted from interpolation with cubic splines, in agreement with the previously shown investigations. There is a total of 72% of the 46 red wines below  $\Delta E$  of three. For white wine, all methods resulted in similar color differences, which were below a value of three. These findings indicate that wine color can be determined using a portable low-cost photometer by interpolation of transmission spectra, for the majority of the sample set closely approximating results from the laboratory photometer.



**Figure 9.** Influence of pathlength of the cuvette on the interpolation of the transmission spectrum via cubic splines for a red wine. Measured was a Pinot Noir (2019) from Germany. The measured spectrum (solid line) and the interpolated spectrum (dashed line) for different pathlengths of glass cuvettes are shown. The spectrum was recorded with a 1 mm, 4 mm, and 10 mm glass cuvette.

#### 4. Conclusions

This study demonstrated the significant influence of the photometer settings data interval and scan speed on the reproducibility of measurements. The CIE  $L^*a^*b^*$  coordinates were found to correlate with the Glories absorbance values in dark red wine ( $L^* < 20$ ), but not in white wine and light red wine ( $L^* > 20$ ). However, Glories absorbance values correlated with themselves in all three wine categories, which indicates a dependency between parameters in violation of Grassmann's first law. The CIE  $L^*a^*b^*$  color space is potentially more reliable for color calculation in wine, but further research is required to investigate which of the two systems correlate better with the human color perception. The calculation of CIE  $L^*a^*b^*$  coordinates for wine color based on single wavelength measurements instead of complete transmission spectra is possible. The best results using a laboratory photometer were achieved using the cubic splines (for red wine) and Lagrange (for white wine) methods of interpolation. Implementation of the portable analysis system resulted in a higher color difference between measured and predicted CIE  $L^*a^*b^*$  coordinates. The determination of white wine color was more successful than in red wine. Furthermore, color determination could be improved using an adapter for a 1 mm cuvette in the portable analysis system. With this investigations, color inexpensive determination of wine is possible. The portable analysis system can be built with low-cost components for about 100€. In comparison to that, laboratory photometers are available from 1.000 to more than 10.000€ This color calculation is highly specialized on wine and cannot be performed with other matrices due to different maxima and minima throughout the spectra. Here, other supporting points are needed. Furthermore, color of sparkling wine or turbid samples cannot be determined due to light scattering. These samples have to be degassed or filtered. Therefore, these investigations could be focused on in future studies.

**Supplementary Materials:** The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/pr10122707/s1>, Table S1: Variety, Origin, and Vintage of the used red wines, Table S2: Variety, Origin, and Vintage of the used white wines, Table S3: Photometer comparison regarding sipper and photometer configuration, Figure S1: Python code for correlation matrices, Figure S2: MATLAB code for interpolation with cubic splines, and Figure S3: MATLAB code for interpolation according to Sprague.

**Author Contributions:** Conceptualization, M.H., S.D.N., D.D. and R.U.; methodology, M.H., Y.M., M.S. and S.D.N.; software, M.H., Y.M. and S.D.N.; validation, M.H. and S.D.N.; formal analysis, M.H. and S.D.N.; investigation, M.H., Y.M., M.S. and S.D.N.; resources, M.H. and S.D.N.; data curation, M.H. and S.D.N.; writing—original draft preparation, M.H. and S.D.N.; writing—review and editing, D.D., J.F. and R.U.; visualization M.H. and S.D.N.; supervision, D.D. and R.U.; project administration, M.H., S.D.N., D.D. and R.U.; funding acquisition, D.D. and R.U. All authors have read and agreed to the published version of the manuscript.

**Funding:** This work was financially supported by the AiF, Arbeitsgemeinschaft industrieller Forschungsvereinigungen “Otto von Guericke” e.V. (Project Number 20964).

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** Not applicable.

**Acknowledgments:** The authors would like to thank the working group of Scharfenberger-Schmeer for their ideas and help as a project partner in our team. Furthermore, we would like to thank Thi Nguyen for assistance with the English language and his ideas and help during the finalization of this manuscript.

**Conflicts of Interest:** The authors declare no conflict of interest.

### Abbreviations

The following abbreviations are used in this manuscript:

A	Absorbance
ANOVA	Analysis of Variance
CIE	Commission Internationale de L’Eclairage
$\Delta E$	Euclidean color distance
LED	Light Emitting Diodes
SMD-LED	Surface-mounted-device LED
SSD	Sum of squared deviations
OIV	Organisation Internationale de la Vigne et du Vin

### References

1. Peynaud, E.; Blouin, J. *El Gusto del Vino: El gran Libro de la Degustación*, 2a. ed., reimpr. ed.; Mundi-Prensa: Madrid, Spain, 2002.
2. Morrot, G.; Brochet, F.; Dubourdieu, D. The color of odors. *Brain Lang.* **2001**, *79*, 309–320. [[CrossRef](#)] [[PubMed](#)]
3. Grassmann. XXXVII. On the theory of compound colours. *Lond. Edinb. Dublin Philos. Mag. J. Sci.* **1854**, *7*, 254–264. [[CrossRef](#)]
4. Glories, Y. La couleur des vins rouges. Ire partie: Les équilibres des anthocyanes et des tanins. *OENO ONE* **2016**, *18*, 195. [[CrossRef](#)]
5. Ribéreau-Gayon, P. *Handbook of Enology: Volume 2: The Chemistry of Wine Stabilization and Treatments*, 2nd ed.; John Wiley: Chichester/West Sussex, England, 2006.
6. *ISO/CIE 11664-1:2019; Colorimetry—Part 1: CIE Standard Colorimetric Observers*. DIN e.V.: Berlin, Germany, 2019.
7. *DIN EN ISO/CIE 11664-2:2019; Colorimetry—Part 2: CIE Standard Illuminants*. DIN e.V.: Berlin, Germany, 2011.
8. Wright, W.D. A re-determination of the trichromatic coefficients of the spectral colours. *Trans. Opt. Soc.* **1929**, *30*, 141–164. [[CrossRef](#)]
9. Wright, W.D. A re-determination of the mixture curves of the spectrum. *Trans. Opt. Soc.* **1930**, *31*, 201–218. [[CrossRef](#)]
10. Guild, J. The colorimetric properties of the spectrum. *Philos. Trans. R. Soc. Lond. Ser. A Contain. Pap. Math. Phys. Character* **1931**, *230*, 149–187. [[CrossRef](#)]
11. Stiles, W.S.; Burch, J.M. N.P.L. Colour-matching Investigation: Final Report (1958). *Opt. Acta Int. J. Opt.* **1959**, *6*, 1–26. [[CrossRef](#)]
12. *DIN EN ISO/CIE 11664-4:2019; Colorimetry—Part 4: CIE 1976 L\*a\*b\* Colour Space*. DIN e.V.: Berlin, Germany, 2020.
13. *DIN EN ISO/CIE 11664-6:2019; Colorimetry—Part 6: CIEDE2000 Colour-Difference Formula*. DIN e.V.: Berlin, Germany, 2016.



14. *Compendium of International Analysis of Methods-Method OIV-MA-AS2-11: Determination of Chromatic Characteristics According to CIELab*; International Organisation of Wine and Vine: Paris, France, 2006; ISBN 978-2-85038-033-4.
15. *Compendium of International Analysis of Methods-OIV Chromatic Characteristics-Method OIV-MA-AS2-07B: Determination of Chromatic Characteristics According to CIELab*; International Organisation of Wine and Vine: Paris, France, 2006; ISBN 978-2-85038-033-4.
16. Pérez-Magariño, S.; González-Sanjosé, M.L. Application of absorbance values used in wineries for estimating CIELAB parameters in red wines. *Food Chem.* **2003**, *81*, 301–306. [[CrossRef](#)]
17. García-Marino, M.; Escudero-Gilete, M.L.; Heredia, F.J.; Escribano-Bailón, M.T.; Rivas-Gonzalo, J.C. Color-copigmentation study by tristimulus colorimetry (CIELAB) in red wines obtained from Tempranillo and Graciano varieties. *Food Res. Int.* **2013**, *51*, 123–131. [[CrossRef](#)]
18. Zhang, X.K.; Lan, Y.B.; Huang, Y.; Zhao, X.; Duan, C.Q. Targeted metabolomics of anthocyanin derivatives during prolonged wine aging: Evolution, color contribution and aging prediction. *Food Chem.* **2021**, *339*, 127795. [[CrossRef](#)]
19. Wang, L.J.; Naudé, N.; Chang, Y.C.; Crivaro, A.; Kamoun, M.; Wang, P.; Li, L. An ultra-low-cost smartphone octochannel spectrometer for mobile health diagnostics. *J. Biophotonics* **2018**, *11*, e201700382. [[CrossRef](#)] [[PubMed](#)]
20. Dutta, S.; Sarma, D.; Patel, A.; Nath, P. Dye-Assisted pH Sensing Using a Smartphone. *IEEE Photonics Technol. Lett.* **2015**, *27*, 2363–2366. [[CrossRef](#)]
21. Ju, Y.G. Fabrication of a low-cost and high-resolution papercraft smartphone spectrometer. *Phys. Educ.* **2020**, *55*, 035005. [[CrossRef](#)]
22. Kong, W.; Kuang, D.; Wen, Y.; Zhao, M.; Huang, J.; Yang, C. Solution Classification With Portable Smartphone-Based Spectrometer System Under Variant Shooting Conditions by Using Convolutional Neural Network. *IEEE Sens. J.* **2020**, *20*, 8789–8796. [[CrossRef](#)]
23. Jian, D.; Wang, B.; Huang, H.; Meng, X.; Liu, C.; Xue, L.; Liu, F.; Wang, S. Sunlight based handheld smartphone spectrometer. *Biosens. Bioelectron.* **2019**, *143*, 111632. [[CrossRef](#)]
24. Zhang, C.; Cheng, G.; Edwards, P.; Zhou, M.D.; Zheng, S.; Liu, Z. G-Fresnel smartphone spectrometer. *Lab Chip* **2016**, *16*, 246–250. [[CrossRef](#)]
25. Edwards, P.; Zhang, C.; Zhang, B.; Hong, X.; Nagarajan, V.K.; Yu, B.; Liu, Z. Smartphone based optical spectrometer for diffusive reflectance spectroscopic measurement of hemoglobin. *Sci. Rep.* **2017**, *7*, 12224. [[CrossRef](#)]
26. Maity, M.; Gantait, K.; Mukherjee, A.; Chatterjee, J. Visible Spectrum-based Classification of Malaria Blood Samples on Handheld Spectrometer. In Proceedings of the 2019 IEEE International Instrumentation and Measurement Technology Conference (I2MTC), Auckland, New Zealand, 20–23 May 2019; IEEE: New York, NY, USA, 2019; pp. 1–5. [[CrossRef](#)]
27. Wilkes, T.C.; Pering, T.D.; McGonigle, A.J.S.; Willmott, J.R.; Bryant, R.; Smalley, A.L.; Mims, F.M.; Parisi, A.V.; England, R.A. The PiSpec: A Low-Cost, 3D-Printed Spectrometer for Measuring Volcanic SO<sub>2</sub> Emission Rates. *Front. Earth Sci.* **2019**, *7*. [[CrossRef](#)]
28. Wilkes, T.C.; McGonigle, A.J.S.; Willmott, J.R.; Pering, T.D.; Cook, J.M. Low-cost 3D printed 1 nm resolution smartphone sensor-based spectrometer: Instrument design and application in ultraviolet spectroscopy. *Opt. Lett.* **2017**, *42*, 4323–4326. [[CrossRef](#)]
29. Albert, D.R.; Todt, M.A.; Davis, H.F. A Low-Cost Quantitative Absorption Spectrophotometer. *J. Chem. Educ.* **2012**, *89*, 1432–1435. [[CrossRef](#)]
30. Long, K.D.; Woodburn, E.V.; Le, H.M.; Shah, U.K.; Lumetta, S.S.; Cunningham, B.T. Multimode smartphone biosensing: The transmission, reflection, and intensity spectral (TRI)-analyzer. *Lab Chip* **2017**, *17*, 3246–3257. [[CrossRef](#)] [[PubMed](#)]
31. Zhao, R.a.; Shen, T.; Lang, T.; Cao, B. Visible smartphone spectrometer based on the transmission grating. In Proceedings of the 2017 16th International Conference on Optical Communications and Networks (ICOON), Wuzhen, China, 7–10 August 2017; pp. 1–3. [[CrossRef](#)]
32. Plaipichit, S.; Wicharn, S.; Buranasiri, P. Spectroscopy system using digital camera as two dimensional detectors for undergraduate student laboratory. *Mater. Today Proc.* **2018**, *5*, 11114–11122. [[CrossRef](#)]
33. Ding, H.; Chen, C.; Qi, S.; Han, C.; Yue, C. Smartphone-based spectrometer with high spectral accuracy for mHealth application. *Sens. Actuators A Phys.* **2018**, *274*, 94–100. [[CrossRef](#)]
34. Woodburn, E.V.; Long, K.D.; Cunningham, B.T. Analysis of Paper-Based Colorimetric Assays with a Smartphone Spectrometer. *IEEE Sens. J.* **2019**, *19*, 508–514. [[CrossRef](#)] [[PubMed](#)]
35. Gallegos, D.; Long, K.D.; Yu, H.; Clark, P.P.; Lin, Y.; George, S.; Nath, P.; Cunningham, B.T. Label-free biodetection using a smartphone. *Lab Chip* **2013**, *13*, 2124–2132. [[CrossRef](#)]
36. Liu, L.; Bi, H. Utilising Smartphone Light Sensors to Measure Egg White Ovalbumin Concentration in Eggs Collected from Yinchuan City, China. *J. Chem.* **2020**, *2020*, 1–8. [[CrossRef](#)]
37. Cesar Souza Machado, C.; Da Silveira Petrucci, J.F.; Silva, S.G. An IoT optical sensor for photometric determination of oxalate in infusions. *Microchem. J.* **2021**, *168*, 106466. [[CrossRef](#)]
38. Khoshmaram, L.; Mohammadi, M.; Nazemi Babadi, A. A portable low-cost fluorimeter based on LEDs and a smart phone. *Microchem. J.* **2021**, *171*, 106773. [[CrossRef](#)]
39. Valenzuela, I.C.; Cruz, F.R.G. (Eds.). *Opto-Electric Characterization of pH Test Strip Based on Optical Absorbance Using Tri-Chromatic LED and Phototransistor*; IEEE: New York, NY, USA, 2015.
40. Negueruela, A.I.; Echávarri, J.F.; Pérez, M.M. A Study of Correlation Between Enological Colorimetric Indexes and CIE Colorimetric Parameters in Red Wines. *Am. J. Enol. Vitic.* **1995**, 353–356.

41. Ayala, F.; Echávarri, J.F.; Negueruela, A.I. A New Simplified Method for Measuring the Color of Wines. *Am. J. Enol. Vitic.* **1997**, *48*, 359–369.
42. Hardy, A.C. *Handbook of Colorimetry*, 4. auflage ed.; M.I.T. Press: Cambridge, MA, USA, 1966.
43. U.S. Bureau of Labor Statistics. *CPI Average Price Data, U.S. City Average (AP)*; U.S. Bureau of Labor Statistics: Washington, DC, USA, 2022.
44. Australian Government. *Priority 1: Increasing Demand and the Premium Paid for All Australian Wine*; Australian Government: Canberra, Australia, 2020.
45. OIV. *The State of Vitiviniculture*; OIV: Paris, France, 2020.
46. Anderson, K.; Nelgen, S. Database of Regional, National and Global Winegrape Bearing Areas by Variety, 1960 to 2016. Available online: <https://economics.adelaide.edu.au/wine-economics/databases#database-of-regional-national-and-global-winegrape-bearing-areas-by-variety-1960-to-2016> (accessed on 2 September 2020).
47. Di Nonno, S.; Ulber, R. Portuino-A Novel Portable Low-Cost Arduino-Based Photo- and Fluorimeter. *Sensors* **2022**, *22*, 7916. [[CrossRef](#)] [[PubMed](#)]
48. Mouser electronics Inc. Phototransistor SFH300. Available online: [https://www.mouser.de/ProductDetail/OSRAM-Opto-Semiconductors/SFH-300-FA-3-4/?qs=K5ta8V%252BWhZU3h4yIF8ddQ==&gclid=EAIaIQobChMI5Zn5r\\_Hc7gIVleh3Ch2TngduEAAYASAAEgJ1w\\_D\\_BwE](https://www.mouser.de/ProductDetail/OSRAM-Opto-Semiconductors/SFH-300-FA-3-4/?qs=K5ta8V%252BWhZU3h4yIF8ddQ==&gclid=EAIaIQobChMI5Zn5r_Hc7gIVleh3Ch2TngduEAAYASAAEgJ1w_D_BwE) (accessed on 13 December 2022).
49. Dahmen, W.; Reusken, A. *Numerik für Ingenieure und Naturwissenschaftler*, 2008th ed.; Springer: Berlin/Heidelberg, Germany, 2008.
50. de Kerf, J.L. The interpolation method of Sprague-Karup. *J. Comput. Appl. Math.* **1975**, *1*, 101–110. [[CrossRef](#)]
51. *Colorimetry*, 3rd ed.; CIE Technical Report; Commission Internationale de l’Eclairage: Vienna, Austria, 2004; Volume 15.
52. Reback, J.; McKinney, W.; Jbrockmendel; den Bossche, J.V.; Augspurger, T.; Cloud, P.; Gfyoung; Sinhrks; Klein, A.; Roeschke, M.; et al. pandas-dev/pandas: Pandas 1.0.3, 2020. Available online: <https://zenodo.org/record/3715232#.Y5k4TX1BxPY> (accessed on 5 November 2022).
53. Harris, C.R.; Millman, K.J.; van der Walt, S.J.; Gommers, R.; Virtanen, P.; Cournapeau, D.; Wieser, E.; Taylor, J.; Berg, S.; Smith, N.J.; et al. Array programming with NumPy. *Nature* **2020**, *585*, 357–362. [[CrossRef](#)] [[PubMed](#)]
54. Hunter, J.D. Matplotlib: A 2D graphics environment. *Comput. Sci. Eng.* **2007**, *9*, 90–95. . [[CrossRef](#)]
55. Waskom, M. seaborn: Statistical data visualization. *J. Open Source Softw.* **2021**, *6*, 3021. [[CrossRef](#)]
56. Martínez, J.A.; Melgosa, M.; Pérez, M.M.; Hita, E.; Negueruela, A.I. Note. Visual and Instrumental Color Evaluation in Red Wines. *Food Sci. Technol. Int.* **2001**, *7*, 439–444. [[CrossRef](#)]
57. HEREDIA, F.J.; GUZMÁN-CHOZAS, M. The color of wine: A historical perspective, II. trichromatic methods. *J. Food Qual.* **1993**, *16*, 439–449. [[CrossRef](#)]

## Supplementary Data

### S1.1. Used Red wine

*Table S1: Variety, Origin, and Vintage of the used red wines*

<b>Number</b>	<b>Variety</b>	<b>Country</b>	<b>Vintage</b>
<b>1</b>	<b>Cabernet Sauvignon</b>	France	2018
<b>2</b>	<b>Cabernet Sauvignon</b>	Italy	2018
<b>3</b>	<b>Cabernet Sauvignon</b>	USA	2014
<b>4</b>	<b>Cabernet Sauvignon</b>	Germany	2018
<b>5</b>	<b>Cabernet Sauvignon</b>	Australia	2016
<b>6</b>	<b>Cabernet Sauvignon</b>	Spain	2016
<b>7</b>	<b>Cabernet Sauvignon</b>	Italy	2019
<b>8</b>	<b>Merlot</b>	Italy	2018
<b>9</b>	<b>Merlot</b>	France	2018
<b>10</b>	<b>Merlot</b>	USA	2018
<b>11</b>	<b>Merlot</b>	France	2016
<b>12</b>	<b>Merlot</b>	New Zealand	2018
<b>13</b>	<b>Merlot</b>	France	2019
<b>14</b>	<b>Merlot</b>	USA	2017
<b>15</b>	<b>Pinot Noir</b>	France	2019
<b>16</b>	<b>Pinot Noir</b>	Germany	2019
<b>17</b>	<b>Pinot Noir</b>	Germany	2013
<b>18</b>	<b>Pinot Noir</b>	Germany	2017
<b>19</b>	<b>Pinot Noir</b>	Germany	2018

20	<b>Pinot Noir</b>	New Zealand	2018
21	<b>Pinot Noir</b>	New Zealand	2019
22	<b>Primitivo</b>	Italy	2020
23	<b>Primitivo</b>	Italy	2017
24	<b>Primitivo</b>	Italy	2020
25	<b>Primitivo</b>	Italy	2019
26	<b>Primitivo</b>	Italy	2017
27	<b>Primitivo</b>	USA	2018
28	<b>Primitivo</b>	USA	2017
29	<b>Syrah</b>	Australia	2018
30	<b>Syrah</b>	Italy	2019
31	<b>Syrah</b>	USA	2017
32	<b>Syrah</b>	France	2019
33	<b>Syrah</b>	New Zealand	2019
34	<b>Syrah</b>	France	2020
35	<b>Syrah</b>	France	2019
36	<b>Tempranillo</b>	Spain	2015
37	<b>Tempranillo</b>	Spain	2015
38	<b>Tempranillo</b>	Spain	2018
39	<b>Tempranillo</b>	Spain	2019
40	<b>Tempranillo</b>	Spain	2017
41	<b>Tempranillo</b>	Spain	2018
42	<b>Tempranillo</b>	Spain	2012
43	<b>Vernatsch</b>	Germany	2018

44	Vernatsch	Germany	2018
45	Vernatsch	Italy	2019
46	Vernatsch	Italy	2019
47	Vernatsch	Italy	2019
48	Vernatsch	Italy	2019
49	Vernatsch	Germany	2018
50	Lemberger	Austria	2018
51	Lemberger	Germany	2018
52	Lemberger	Austria	2017
53	Lemberger	Germany	2019
54	Lemberger	Germany	2016
55	Lemberger	Germany	2019
56	Lemberger	Germany	2019

## S1.2. Used White wines

Table S2: Variety, Origin, and Vintage of the used white wines

Number	Variety	Country	Vintage
1	Chardonnay	France	2018
2	Chardonnay	Italy	2020
3	Chardonnay	Germany	2018
4	Chardonnay	Spain	2019
5	Chardonnay	Germany	2016

<b>6</b>	<b>Chardonnay</b>	Australia	2019
<b>7</b>	<b>Chardonnay</b>	USA	2018
<b>8</b>	<b>Riesling</b>	Germany	2015
<b>9</b>	<b>Riesling</b>	France	2017
<b>10</b>	<b>Riesling</b>	USA	2013
<b>11</b>	<b>Riesling</b>	Germany	2018
<b>12</b>	<b>Riesling</b>	Germany	2019
<b>13</b>	<b>Riesling</b>	Germany	2020
<b>14</b>	<b>Riesling</b>	Germany	2020
<b>15</b>	<b>Riesling</b>	Germany	2019
<b>16</b>	<b>Riesling</b>	Germany	2018
<b>17</b>	<b>Riesling</b>	Germany	2019
<b>18</b>	<b>Riesling</b>	Germany	2019
<b>19</b>	<b>Riesling</b>	Germany	2018
<b>20</b>	<b>Riesling</b>	Germany	2019
<b>21</b>	<b>Riesling</b>	Germany	
<b>22</b>	<b>Pinot gris</b>	Italy	2019
<b>23</b>	<b>Pinot gris</b>	Italy	2020

24	<b>Pinot gris</b>	France	2017
25	<b>Pinot gris</b>	New Zealand	2018
26	<b>Pinot gris</b>	Germany	2019
27	<b>Pinot gris</b>	Germany	2019
28	<b>Pinot gris</b>	Germany	2019
29	<b>Sauvignon blanc</b>	France	2019
30	<b>Sauvignon blanc</b>	Germany	2019
31	<b>Sauvignon blanc</b>	Italy	2019
32	<b>Sauvignon blanc</b>	USA	2019
33	<b>Sauvignon blanc</b>	New Zealand	2020
34	<b>Sauvignon blanc</b>	France	2018
35	<b>Sauvignon blanc</b>	Germany	2018
36	<b>Pinot blanc</b>	Germany	2019
37	<b>Pinot blanc</b>	Germany	2020
38	<b>Pinot blanc</b>	Germany	2018
39	<b>Pinot blanc</b>	Germany	2017
40	<b>Pinot blanc</b>	Germany	2020
41	<b>Pinot blanc</b>	Italy	2019

42	Pinot blanc	France	2017
43	Gewürztraminer	Germany	2016
44	Gewürztraminer	Germany	2018
45	Gewürztraminer	France	2019
46	Gewürztraminer	Germany	2019
47	Gewürztraminer	France	2019
48	Gewürztraminer	Spain	2019
49	Gewürztraminer	Germany	2019
50	Gewürztraminer	Italy	2019
51	Gewürztraminer	Italy	2019
52	Gewürztraminer	Spain	2018
53	Gewürztraminer	Germany	2019
54	Gewürztraminer	France	2018
55	Gewürztraminer	Germany	2019
56	Gewürztraminer	Germany	2020

### S1.3. Photometer settings

Table S3: Photometer comparison regarding sipper and photometer configuration

Parameter	JASCO V-730	Varian Cary 100
Sipper configuration	1 mm flow cuvette	



Suction time	3 sec	8 sec
Flow time	0 sec	0 sec
Wait time	1 sec	5 sec
Drain time	5 sec	8 sec
<b>Photometer configuration</b>		
Photometric mode	Absorption	Absorption
Bandwidth	1 nm	2 nm
Response time	0.96 sec	0.03-0.6 sec (depends on scan speed)
Correction mode	Baseline	Baseline
Light source > 340 nm	Tungsten lamp	Tungsten lamp
Light source < 340 nm	Deuterium lamp	Deuterium lamp
Filter exchange	Stepwise	n. A.
Monochromator system	Single monochromator in Rowland off-circle arrangement	Czerny-Turner monochromator
Detector system	Si-Photodiode (S1337)	Photomultiplier tube (R928)

Figure S1. Python Code for correlation matrices

```

#import packages
import numpy as np # numeric calculations
import pandas as pd #excel data import
import matplotlib.pyplot as plt #plotting base
import seaborn as sns #statistical plotting
import warnings
warnings.filterwarnings('ignore')
%matplotlib inline

#import excel file
df = pd.read_excel(r'Excel-file directory')
#show columns and row length
df.shape
# DataFrame optimization
df.columns = df.columns.str.strip() # delete spaces in column names
df.info()# check properties (non-null count)
df.isnull().sum() #check properties (null-count)
data= df.dropna() # drop rows without values
data.isnull().sum() # re-check null count
data = data.rename(columns={'420 nm': 'A[420 nm]', '520 nm': 'A[520 nm]', '620
nm': 'A[620 nm]'}) #rename columns
round(data.describe(),2) # describe the data count, mean, std 25%,50%,75%
quantils
#Calculate and plot Heatmap
sns.set(font_scale=1.4)#set overall font size
plt.figure(figsize=(12,8.27)) #set figure size
correlations = data.corr(method='spearman') # calculate spearman
correlation
mask = np.zeros_like(correlations) # mask the upper half of the correlation
matrix that shows the same as the down half
mask[np.triu_indices_from(mask)] = True
sns.set_style("whitegrid", {'axes.grid' : False}) # Hide Grid
#plot heatmap
sns.heatmap(round(correlations,2), cmap='RdGy', annot=True,
            annot_kws={"size": 18}, vmin=-1, vmax=1,square=True,mask=mask)
plt.yticks(rotation=0) #rotation of y ticks
plt.xticks(rotation=15) #roaction of x ticks
#save plot as vector graphic
plt.savefig(r"save directory", format='svg')

```

**Figure S2. MATLAB Code for interpolation with cubic splines**

```
% Cubic splines
% Enter wavelengths
x = [
380
420
460
525
590
625
675
775
];

% Enter transmission values
y = [
t_380
t_420
t_460
t_525
t_590
t_625
t_675
t_775
];

a = (y(2)-y(1))./(x(2)-x(1));
b = (y(8)-y(7))./(x(8)-x(7));

yy = spline(x,[1; y; 0]);
```

**Figure S3. MATLAB code for interpolation according to Sprague**

```

%%Sprague

% Enter wavelengths
x = [
380
420
460
525
590
625
675
775
];

% Enter transmission values
y = [
t_380
t_420
t_460
t_525
t_590
t_625
t_675
t_775
];

a = (884*y(1)-1960*y(2)+3033*y(3)-2648*y(4)+1080*y(5)-180*y(6))/209;
b = (508*y(1)-540*y(2)+488*y(3)-367*y(4)+144*y(5)-24*y(6))/209;
c = (-24*y(3)+144*y(4)-367*y(5)+488*y(6)-540*y(7)+508*y(8))/209;
d = (-180*y(3)+1080*y(4)-2648*y(5)+3033*y(6)-1960*y(7)+884*y(8))/209;

w = [a; b; y; c; d];

b = size(y);

e = zeros(b(1)-1,6);

%%

for i=3:9
a0 = w(i);
a1 = (w(i-2)-8*w(i-1)+8*w(i+1)-w(i+2))/12;
a2 = (-w(i-2)+16*w(i-1)-30*w(i)+16*w(i+1)-w(i+2))/24;
a3 = (-9*w(i-2)+39*w(i-1)-70*w(i)+66*w(i+1)-33*w(i+2)+7*w(i+3))/24;
a4 = (13*w(i-2)-64*w(i-1)+126*w(i)-124*w(i+1)+61*w(i+2)-12*w(i+3))/24;
a5 = (-5*w(i-2)+25*w(i-1)-50*w(i)+50*w(i+1)-25*w(i+2)+5*w(i+3))/24;

e(i-2,1)= a0;
e(i-2,2)= a1;
e(i-2,3)= a2;
e(i-2,4)= a3;
e(i-2,5)= a4;
e(i-2,6)= a5;
end

```

## 5.2 New insights into Wine Color Analysis: A Comparison of Analytical Methods to Sensory Perception of Red and White Varietal Wines

The results presented in section 5.1 indicate that the CIE  $L^*a^*b^*$  color space and the Glories color measurement system cannot be used interchangeably. To determine which method is most used for the quantification of the wine color, both methods are compared to the human perception of the wine color. For this, the same wines used for the correlation presented in section 5.1 were evaluated by a panel of 24 trained individuals. Linear line scales for lightness and hue were used for the descriptive analysis. To evaluate the performance of each approach, the ability of each method to differentiate between different grape varieties were used.

In red wine, both the  $L^*$  coordinate and the Glories' CI parameter were able to depict the human perception of the lightness of the wines. Regarding the hue, the CIE  $L^*a^*b^*$  color space was able to depict the human perception of the hue more closely than the hue parameters of the Glories color measurement ( $T$  and  $dA(\%)$ ). However, counterintuitive to the definition, the  $a^*$  coordinate, which indicates the redness of a wine, was not able to differentiate between the grape varieties well. A better resolution could be achieved by the  $b^*$  coordinate, which indicates the yellowness of wine, leading to the conclusion that humans search for differences apart from the red color. Although better, the CIE  $L^*a^*b^*$  color space and the perceived hue deviate from each other in the light red color area. Similar effects could be observed using white wines where both the  $L^*$  coordinate and Glories' CI parameter were able to differentiate the grape varieties according to the human perception of the lightness. Apart from that, the  $b^*$  coordinate matched the human perception of the color hue closest, outperforming the hue parameter of Glories' color measurement system.

The second objective was to re-evaluate the JND with the CIEDE2000 formula and compare it with the Euclidean color distance using a triangle test. The previously established JND of 3 could be reproduced in dark red wine using the Euclidean color distance. However, the JND differs throughout the color space. Differences in the light yellow area were differentiated the easiest, followed by dark red wines and light red wines. Future studies should use CIE  $L^*a^*b^*$  instead of Glories' color measurement and use the new JND using the CIEDE2000 formula.

## **New insights into Wine Color Analysis: A Comparison of Analytical Methods to Sensory Perception of Red and White Varietal Wines**

Authors:

Marcel Hensel<sup>1</sup>, Marina Scheiermann<sup>1</sup>, Jörg Fahrer<sup>3</sup>, Dominik Durner<sup>1</sup>, Affiliations:

<sup>1</sup> Weincampus Neustadt, Dienstleistungszentrum ländlicher Raum Rheinpfalz, 67435 Neustadt an der Weinstraße, Germany

<sup>2</sup> Division of Food Chemistry and Toxicology, Department of Chemistry, University of Kaiserslautern-Landau, 67663 Kaiserslautern, Germany

### **Disclaimer:**

**The presented work was exclusively performed by the DLR Rheinpfalz (Marcel Hensel and master thesis Marina Scheiermann). For more information, refer to the author contribution section.**

Copyright: Reprinted with permission from M. Hensel, M. Scheiermann, J. Fahrer, D. Durner, *New insights into wine color analysis: A comparison of analytical methods to sensory perception for red and white varietal wines*, *Journal of agricultural and food chemistry* (2023). Copyright ©2023 American Chemical Society.

# New Insights into Wine Color Analysis: A Comparison of Analytical Methods to Sensory Perception for Red and White Varietal Wines

Marcel Hensel, Marina Scheiermann, Jörg Fahrner, and Dominik Durner\*



Cite This: <https://doi.org/10.1021/acs.jafc.3c01284>



Read Online

ACCESS |



Metrics & More



Article Recommendations



Supporting Information

**ABSTRACT:** The Glories method and CIE  $L^*a^*b^*$  color space are poorly correlated in the light red high-saturation color area. Non-uniformities in the CIE  $L^*a^*b^*$  color space have led to the CIEDE2000 formula, while the Euclidean color distance is still used in wine research. In this study, 112 white and red wines were examined to compare the Glories method, CIE  $L^*a^*b^*$ , and human perception using monovarietal wines from different grape varieties. The objective of this study was to investigate which of the two methods and which parameter of each method better aligned with human perception. The visual color threshold was re-evaluated with the CIEDE2000 formula and triangle testing. CIE  $L^*a^*b^*$  more closely matched the human perception, elevating the use of CIE  $L^*a^*b^*$  over the use of the Glories method. Visual color thresholds were better expressed with CIEDE2000 but still varied depending upon the color area in the CIE  $L^*a^*b^*$  color space.

**KEYWORDS:** CIE  $L^*a^*b^*$ , Glories, colorimetry, color sensory, wine color, CIEDE2000, Euclidean color distance

## INTRODUCTION

Color is the first sensory impression to be perceived when wine is poured into the glass, leading to consumer expectations regarding aroma, taste, and mouthfeel.<sup>1,2</sup> One study showed that Italian Novello wines with an intense color were associated with higher quality.<sup>3</sup> Another study on rosé wines from Provence confirmed that wine color, among other parameters, influences perceived typicality.<sup>4</sup> A study on Spanish Fondillón wines demonstrated that color is an important driver for customers to buy wines.<sup>5</sup> Besides the influence of color on wine quality and typicality, Blanc de Noir wines in Germany are tied to specific regulations regarding color.<sup>6</sup> Such official regulations and many publications investigating the influence of wine color on wine quality ratings, consumer acceptance, and purchase decisions underline the great importance of wine color. The subjective nature of human perception, however, underlines the need of an objective and reliable parameter that reflects the human perception.

Two methods to characterize wine color are recommended by the International Organisation of Wine and Vine (OIV). The first uses the equidistant  $L^*a^*b^*$  color space according to the Commission Internationale de l'Eclairage (CIE). The  $L^*a^*b^*$  parameters are Cartesian coordinates, which can be transformed to the polar coordinates  $L^*$ ,  $C^*$ , and  $h^0$ . In these color spaces, equally distant colors should represent perceived color differences of equal size.<sup>7,8</sup>

The second method was developed by Glories and uses absorbance values at 420, 520, and 620 nm. These primary values are used to calculate secondary parameters: color intensity (CI) for lightness, brilliance [dA (%)] for the amount of red color in a wine, and hue ( $T$ ) for the color evolution in the orange/brown color area.<sup>9,10</sup>

Studies have shown the use of anthocyanin profiles to differentiate between wines from different grape varieties.<sup>11</sup> This approach is costly and not easily interpretable, because consumers commonly differentiate grape varieties by wine color. Rather than analysis of anthocyanins, polymeric pigments, and other polyphenols, analysis of wine color is a feasible tool for classification that more closely reflects consumer perception.

Both Glories' method and CIE color spaces are widely used to differentiate between wines within one variety,<sup>12–14</sup> but to more completely understand the capabilities of both methods, a comparison using different varieties is needed. Somers conducted a study comparing Cabernet Sauvignon and Shiraz wines using the Glories method and found a significant difference between the varieties using CI but not dA and  $T$ .<sup>15</sup>

Furthermore, previous studies have demonstrated that Glories' method and the CIE  $L^*a^*b^*$  color space cannot be used interchangeably for light red wines and white wines.<sup>16</sup> The Glories method was originally developed for dark red wines but is now also used for rosé, white, and Blanc de Noir wines.<sup>17–19</sup> Therefore, it is unclear whether the Glories color parameters reflect human perception of color outside of dark red wines. On the other hand, CIE color spaces should represent any aspect of perceived color, but it is uncertain whether it is able to do so in wine, given the experimental conditions, in which the CIE color spaces designed are completely different from those used during the sensory

**Special Issue:** Highlights of the In Vino Analytica Scientia Conference 2022

**Received:** February 28, 2023

**Revised:** May 9, 2023

**Accepted:** May 9, 2023

evaluation of wine color. During the development of the CIE color spaces, panelists were asked to match a given hue via additive color mixing of the three primary colors red, green, and blue,<sup>20–23</sup> whereas linear line scales are used in sensory to describe color.<sup>24–28</sup> It has yet to be determined which of the two methods, CIE or Glories, more closely aligns with human perception of wine color. It was also not yet investigated which of the parameters do align with the perceived color.

Besides that, several methods exist for the calculation of the distance between two colors in the CIE  $L^*a^*b^*$  color space.<sup>29,30</sup> The simplest method, currently recommended by the OIV, is the calculation of the CIE76 Euclidean color distance ( $\Delta E_{ab}^*$ ).<sup>7</sup> However, the CIE has refined their definition of color distance over the years as a result of perceptual non-uniformities within the CIE  $L^*a^*b^*$  color space. Newer color distance formulas proposed by the CIE include the CIE94 metric and the CIEDE2000 ( $\Delta E_{00}$ ) metric.<sup>31</sup> Besides the CIE, other societies, such as the Color Measurement Committee (CMC) of the Society of Dyers and Colorists (SDC) have defined their own color distance formulas.<sup>31</sup> Many studies used color distance formulas to express a color threshold. For illustration and printing technology,  $\Delta E_{ab}^* = 2.3$  was proposed as a color threshold to measure the performance of printers and monitors.<sup>32</sup> In the field of medical displays used for diagnosis,  $\Delta E_{00} = 1$  is used as a color threshold.<sup>33</sup> In the field of viticulture and enology, an oft-cited study using Spanish red wines proposed a color threshold of  $\Delta E_{ab}^* = 3$  for wine.<sup>34</sup> To investigate the performance of the different color distance formula in equidistant color spaces, a study was conducted using the standardized residual sum of squares (STRESS) technique to calculate the difference between computed and visually perceived color differences.<sup>35,36</sup> The STRESS value for commonly used  $\Delta E_{ab}^*$  indicated that it was less accurate than both  $\Delta E_{00}$  and CIE94. Accordingly, the CIE currently recommends the CIEDE2000 color distance formula.<sup>37</sup> Nevertheless, the CIE  $L^*a^*b^*$  color space in combination with  $\Delta E_{ab}^*$  were still recommended for wine color by the OIV.<sup>38</sup>

In the present study, the applicability of the Glories method and the CIE  $L^*a^*b^*$  color space are evaluated by a comparison approach using wines of different varieties. The objective was to investigate which method better aligned with the visual color perception. Furthermore, the alignment of each parameter of the two methods was investigated. Additionally, the smallest visible color difference was re-evaluated as a result of improved knowledge and new suggestions by the CIE regarding color distance formulas.

## MATERIALS AND METHODS

**Samples.** *Samples for the Varietal Comparison of the Different Color Measurements.* A total of 112 commercial red and white wines were purchased and analyzed in July 2021. Seven wines each of eight red varieties (Cabernet Sauvignon, Lemberger, Merlot, Pinot Noir, Primitivo, Shiraz, Tempranillo, and Vernatsch) from the vintages 2012–2020 were analyzed. The wines originated from Australia, Austria, France, Germany, Italy, New Zealand, Spain, and the U.S.A. Seven wines each of four white varieties (Chardonnay, Pinot Blanc, Pinot Gris, and Sauvignon Blanc) and 14 wines each of two white varieties (Gewürztraminer and Riesling) from the vintages 2013–2020 were examined. Gewürztraminer and Riesling were represented by seven dry wines and seven sweet wines each. All white wines originated from the same countries as the red wines (except for Austria).

The vintages for each monovarietal red wine were chosen to cover the typical time ranges of consumption and to reflect the market availability. As a result, Tempranillo vintages ranged from 2012 to 2019, Pinot Noir vintages ranged from 2013 to 2019, Cabernet Sauvignon vintages ranged from 2014 to 2019, and Vernatsch vintages ranged from 2018 to 2019. For white wines, Riesling vintages ranged from 2015 to 2020, Chardonnay and Gewürztraminer vintages ranged from 2016 to 2020, and Sauvignon Blanc vintages ranged from 2019 to 2020.

As a proof of concept, a two-way analysis of variance (ANOVA) using the CIE  $L^*a^*b^*$  data of all wines was conducted assessing the impacts of the variety and the vintage. The different vintage ranges of different varieties had a negligible impact compared to the cross-cultural effects.

The list of all wines, including the information about variety, vintage, origin, and producer, is provided in Tables S1 and S2 of the Supporting Information.

**Samples for the Determination of the Visible Color Threshold.** A 2018 Cabernet Sauvignon from Staatsweingut mit Johannitergut (Neustadt an der Weinstraße, Germany) and a 2017 Riesling from Weingut Bürklin-Wolf (Wachenheim, Germany) were used to determine the smallest visible color difference in July 2021.

**Photometric Measurements and CIE  $L^*a^*b^*$  Calculations.** The spectra for CIE  $L^*a^*b^*$  and the absorbance values at 420, 520, and 620 nm were recorded with a V-730 double-beam photometer (JASCO, Japan) in July 2021. The spectra were taken between 300 and 900 nm with a data interval of 1 nm and a scan speed of 1000 nm/min, because it results in the highest reproducibility.<sup>16</sup> Red wines were measured in a 1 mm quartz flow cuvette (Hellma, Germany), and white wines were measured in a 10 mm polystyrene cuvette (Brand, Germany). To comply with the OIV recommendation, the red wine spectra were corrected to a 10 mm cuvette.<sup>38</sup> The CIE  $L^*a^*b^*$  coordinates were calculated with the 10° standard observer and the D65 standard illuminant between 360 and 780 nm.

**Color Sensory. Standardization of the Color Evaluation.** Illumination and viewing conditions for the sensory analysis of wine are, despite the relevance of color impression for odor and taste, not particularly well-defined or standardized.<sup>39</sup> Therefore, a method was developed that reduces most of the influencing external stimuli. For this purpose, a sensory box (720 × 700 × 720 mm) was lined with black fabric, so that as much scattered light as possible is absorbed (Figure S1 of the Supporting Information). The samples were placed on a white sheet of paper. Before the sensory evaluation, the panelists were positioned by means of markings on the box in a way that the viewing angle corresponded to a (0°; 45° x) geometry according to DIN 50337-7, where the angle between illuminant is 0° and the reflected light is detected at a 45° angle by the observer.<sup>40</sup> The illuminant was placed 30 cm from the sample, as recommended by the manufacturer of the D65 illuminant. For the evaluation, the room was completely darkened. The wines were pipetted in 10 mm polystyrene cuvettes to furthermore standardize the light pathway. The samples were covered with nitrogen and sealed to prevent oxidation and alcohol evaporation, and the liquid meniscus was covered with black tape to ensure that the different refractions of the liquid meniscus have no influence on the outcome.

**Color Assessment Panel.** The panel was chosen demographically and consisted of 23 trained panelists, of whom 12 were female and 11 were male, because color perception is dependent upon the gender.<sup>38</sup> The panel was categorized in different age groups ranging from 20 to 30, from 31 to 40, from 41 to 50, and from 51 to 60, because age also plays a role in color perception.<sup>41</sup> The judges did a color vision test based on the Farnsworth–Munsell 100 hue test,<sup>42</sup> a training of the visible color threshold evaluation and the actual color threshold evaluation. For the descriptive color analysis, the panel was trained in one session for dark red wine, light red wine, and white wine with the blends used for the color threshold determination.

**Descriptive Color Analysis.** Two scales were used, one for lightness and one for hue. The lightness scale ranged from light (0) to dark (100). For red wines, the hue scale ranged from violet (−50) to red (0) to orange (50), whereas for white wine, the hue scale ranged from



green (0) to yellow (50) to orange (100). The performance of the sensory evaluation, the Glories color measurement, and the CIE  $L^*a^*b^*$  color space were measured by the ability to distinguish between the different varieties. The results of the sensory evaluation were used as a reference point to evaluate the accuracy of the analytical methods.

**Determination of the Visible Color Threshold.** The determination of the smallest visible color difference was performed according to DIN EN ISO 4210:2021 via a triangle test. Blends of Cabernet Sauvignon and Riesling, in different dilutions, were used to create light yellow, light red, and dark red solutions (Table 1). This type of classification is needed, because dark red wines and light red wines yielded different results comparing CIE  $L^*a^*b^*$  and Glories' color measurement system.<sup>16</sup> For the dark red blends, undiluted Cabernet Sauvignon was used as a base wine and diluted with Riesling. For the

base wine of the light red blends, the Cabernet Sauvignon was diluted by a factor of 8 with Riesling, and then Cabernet Sauvignon was used to create the dilutions. For the light yellow blends, the undiluted Riesling wine was used as a base wine and diluted with Cabernet Sauvignon. The Euclidean color distance and the CIEDE2000 color difference formula between the dilutions and the base wine were calculated according to the literature (eqs 1 and 2).<sup>29,30</sup>

$$\Delta E_{ab} = \sqrt{(L_1^* - L_2^*)^2 + (a_1^* - a_2^*)^2 + (b_1^* - b_2^*)^2} \quad (1)$$

with  $L^*$ ,  $a^*$ , and  $b^*$  = CIE  $L^*a^*b^*$  coordinates (CIE76)

$$\Delta E_{00} = \sqrt{\left(\frac{\Delta L'}{k_L S_L}\right)^2 + \left(\frac{\Delta C'}{k_C S_C}\right)^2 + \left(\frac{\Delta h'}{k_h S_h}\right)^2 + R_T \left(\frac{\Delta C'}{k_C S_C}\right) \left(\frac{\Delta h'}{k_h S_h}\right)} \quad (2)$$

with  $\Delta L'$  = lightness difference (CIEDE2000),  $\Delta C'$  = chroma difference (CIEDE2000),  $\Delta h'$  = hue angle difference (CIEDE2000),  $k_L$ ,  $k_C$ , and  $k_h$  = parametric factors for lightness, chroma, and hue angle,  $S_L$ ,  $S_C$ , and  $S_h$  = weighting functions for lightness, chroma, and hue angle, and  $R_T$  = rotation function (the CIEDE2000 transformation is shown in the Supporting Information).

For the threshold analysis, 3 mL of the blend was displayed in photometer cuvettes prepared as mentioned before. To calculate the visible color difference, the best estimated threshold according to ASTM E-679-04 was used.

**Table 1. Dilutions To Produce Blends of Cabernet Sauvignon and Riesling Yielding Samples for the Smallest Visible Color Difference Testing in the Classes: Light Yellow Color Area, Light Red Color Area, and Dark Red Color Area<sup>a</sup>**

dilution	$L^*$	$a^*$	$b^*$	$\Delta E_{ab}^*$	$\Delta E_{00}$
Dark Red Blends <sup>b</sup>					
reference red wine <sup>c</sup>	5.36	32.13	9.24		
1:80	5.22	31.62	8.99	0.58	0.25
1:40	5.58	32.79	9.62	0.79	0.34
1:20	6.09	34.18	10.48	2.50	1.08
1:13.3	6.47	35.06	11.13	3.66	1.59
1:10	6.83	35.82	11.76	4.70	2.06
1:6.7	8.40	38.34	14.43	8.64	3.87
1:5.7	9.07	39.22	15.57	10.20	4.60
1:5	9.98	40.46	17.12	12.35	5.60
1:4	10.81	41.29	18.48	14.10	6.45
1:3.3	12.38	43.20	21.06	17.65	8.06
Light Red Blends <sup>d</sup>					
reference red wine blend (1:8) <sup>b</sup>	64.01	38.45	12.00		
1:7.97	65.94	39.11	11.52	2.09	1.65
1:7.94	66.13	39.71	11.69	2.49	1.84
1:7.91	66.76	40.65	11.47	3.57	2.46
1:7.84	67.20	41.25	11.30	4.09	2.74
1:7.7	66.75	42.28	11.25	4.77	2.77
1:7.27	65.47	43.59	11.74	6.43	2.78
1:6.67	63.63	46.56	12.77	8.16	2.85
1:5.71	58.33	50.82	14.85	13.85	6.44
1:5	54.03	54.15	17.07	19.28	10.30
1:4.4	50.14	56.64	19.32	24.02	14.10
Light Yellow Blends <sup>d</sup>					
reference white wine <sup>e</sup>	98.89	-1.08	6.14		
1:5.000	98.83	-0.98	6.22	0.13	0.15
1:2.500	98.69	-0.87	6.25	0.30	0.32
1:1.666	98.58	-0.82	6.24	0.41	0.41
1:1.250	98.47	-0.71	6.27	0.57	0.58
1:1.000	98.49	-0.65	6.30	0.61	0.66
1:666	98.12	-0.38	6.39	1.07	1.10
1:500	98.06	-0.22	6.34	1.21	1.32
1:400	97.85	-0.03	6.43	1.51	1.61
1:333	97.89	0.16	6.31	1.60	1.84
1:250	97.32	0.63	6.43	2.33	2.59

<sup>a</sup>After dilutions were prepared, samples have been characterized by CIE  $L^*a^*b^*$  to calculate  $\Delta E_{ab}^*$  and  $\Delta E_{00}$ . <sup>b</sup>Diluted with Riesling.

<sup>c</sup>Undiluted Cabernet Sauvignon. <sup>d</sup>Diluted with Cabernet Sauvignon.

<sup>e</sup>Undiluted Riesling.

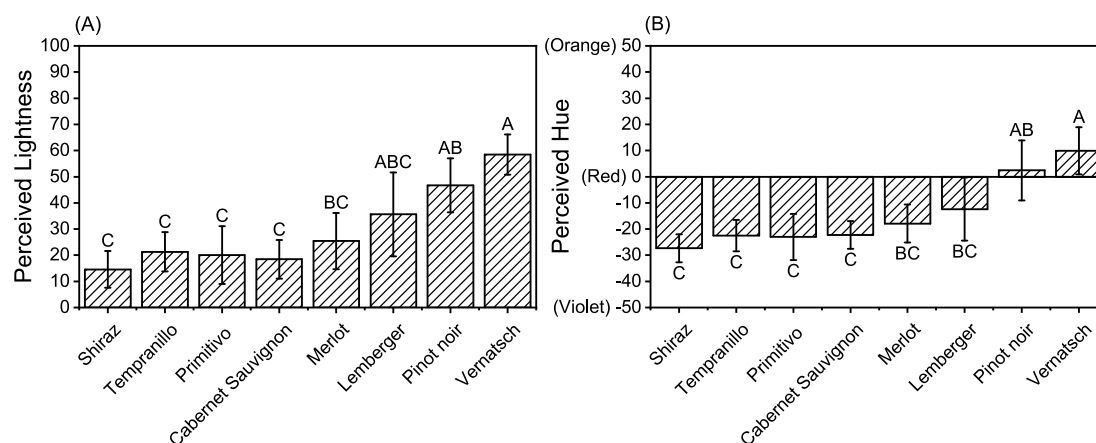
## SOFTWARE AND DATA ANALYSIS

For the sensory data, a panel analysis was performed in Excel (Microsoft, Redmond, WA, U.S.A.) with XLSTAT (2020, ADDINSOFT, France). Normality testing via the Shapiro–Wilks test, homoscedasticity testing via Levene's test, ANOVA, and Tukey's honest significant difference (HSD) post hoc test were performed with Origin(Pro) (2020b, OriginLab, Northampton, MA, U.S.A.).

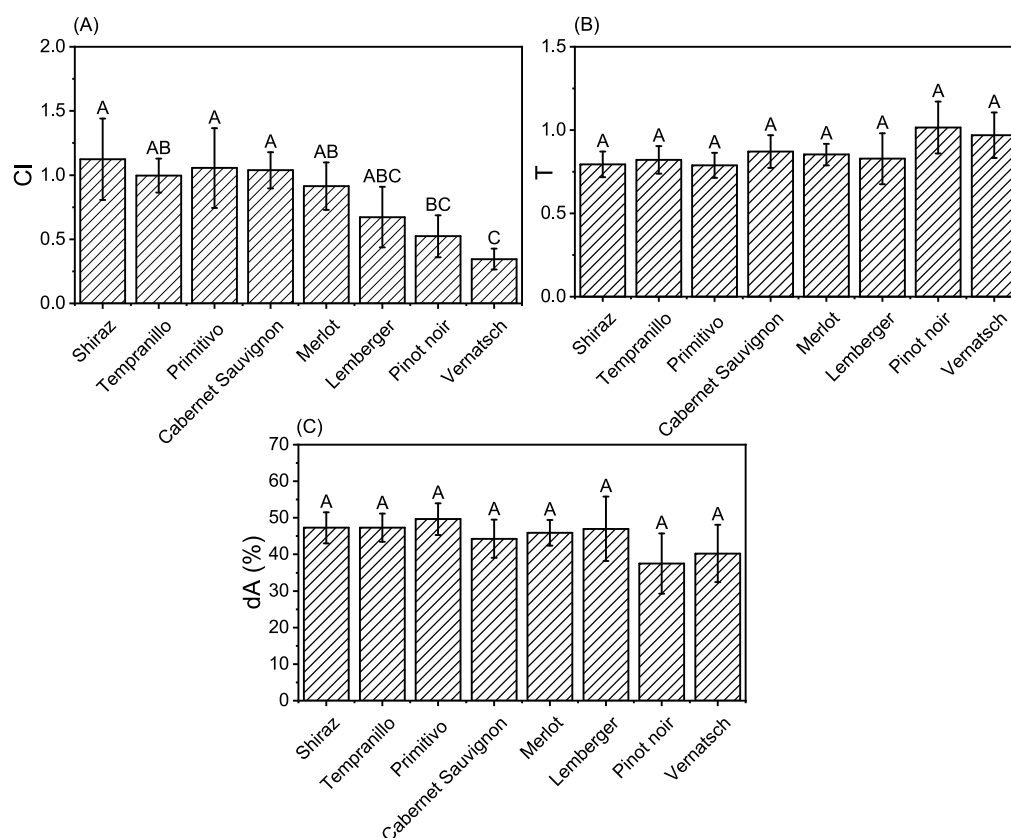
## RESULTS AND DISCUSSION

**Wine Color Determination by Visual Perception and Analytical Methods. Red Wine. Visual Perception.** Lightness ratings distinguished light red wines (Vernatsch and Pinot Noir) from dark red wines (Shiraz, Tempranillo, Primitivo, and Cabernet Sauvignon) (Figure 1A). The same discrimination could be observed with the hue: Vernatsch and Pinot Noir were rated the most orange, while Shiraz, Tempranillo, Primitivo, and Cabernet Sauvignon were rated the most violet (Figure 1B). Merlot and Lemberger wines were found in between the light and dark groups for both lightness and hue. Intervarietal differentiation based on perceived lightness was possible between Merlot and Vernatsch, while perceived hue could be used to differentiate between Vernatsch from Merlot and Lemberger. The results of perceived hue revealed a lower variance within the varieties compared to perceived lightness, suggesting that perceived hue is higher in discriminative power. There is currently no standard approach to describe wine color, and studies used sometimes one,<sup>24,26</sup> two,<sup>43,44</sup> or even more<sup>25</sup> scales for different aspects of color. Results here indicate that more than one color scale should be used to fully describe the color of red wines by means of sensory evaluation.

**Glories' Color Measurement.** Glories' color intensity CI, indicating lightness, allowed for discrimination between some varieties (Figure 2A), similar to sensory ratings of perceived lightness (Figure 1A). It allowed for the differentiation of Vernatsch from dark red wines Shiraz, Primitivo, Tempranillo, and Cabernet Sauvignon. However, it was not possible to distinguish Tempranillo from Pinot Noir using CI (Figure 2A)



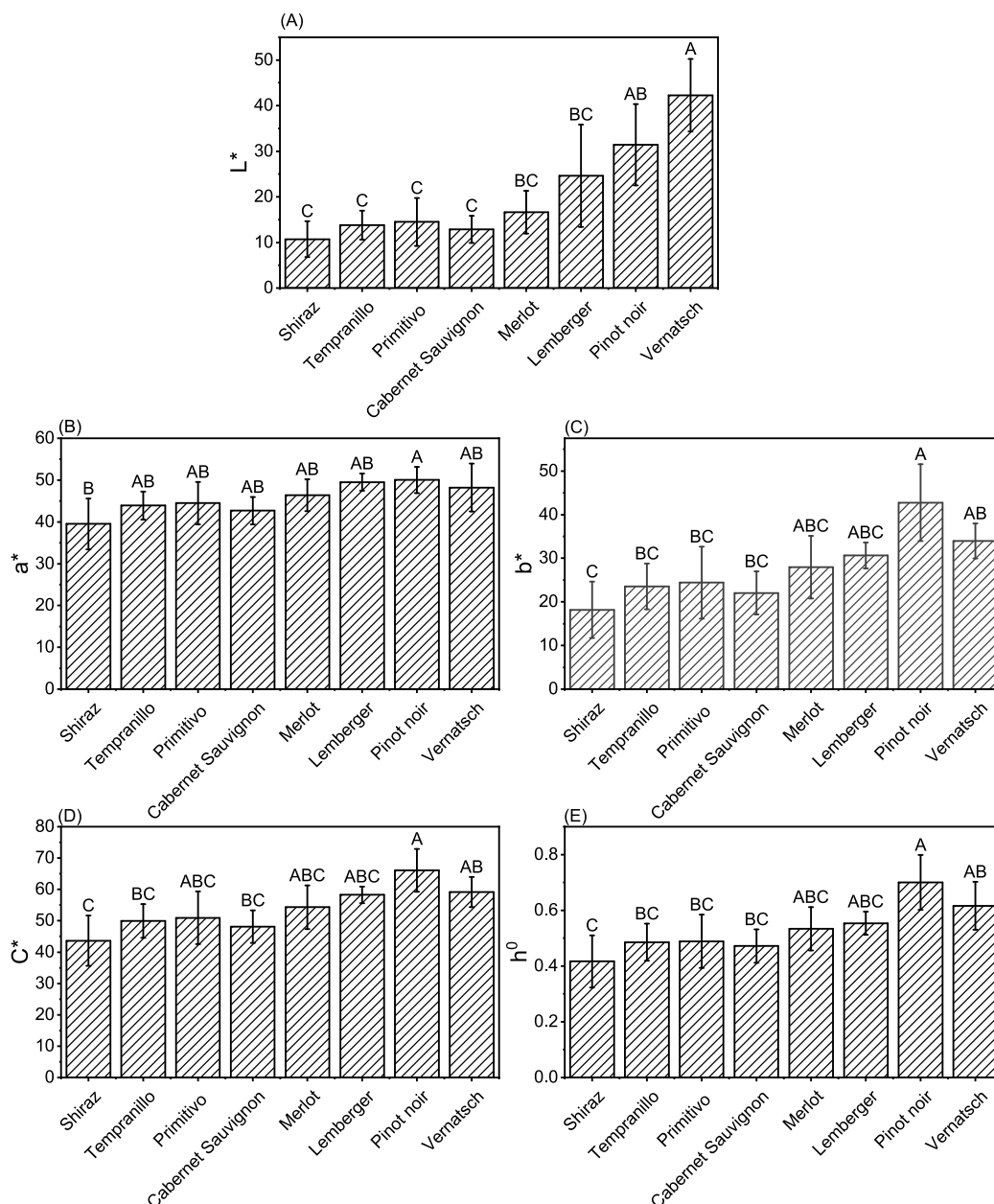
**Figure 1.** Sensory ratings for (A) lightness and (B) hue for eight different red grape varieties, averaged for seven wines per grape variety. Error bars show the standard deviation of the seven wines of each grape variety. Letters show significance groups determined by ANOVA ( $p < 0.001$ ) and Tukey's HSD post hoc test ( $p < 0.001$ ).



**Figure 2.** Glories' parameters (A) color intensity CI, (B) hue  $T$ , and (C) brilliance  $dA$  (%) for eight different red grape varieties, averaged for seven wines per grape variety. Error bars show the standard deviation of the seven wines of each grape variety. Letters show significance groups determined by ANOVA ( $p < 0.001$ ) and Tukey's HSD post hoc test ( $p < 0.001$ ).

as it was with perceived lightness (Figure 1A). Glories'  $T$  (Figure 2B) and  $dA$  (%) (Figure 2C) parameters failed to differentiate varieties. This is not in agreement with findings from sensory evaluation, where perceived hue could be used to distinguish between some varieties (Figure 1B). Therefore, it is not recommended to use Glories' hue parameters to differentiate between varieties. While Glories' hue parameters may be used for differentiation between single wines, it is not particularly suited for intervarietal comparisons. An explanation could be the different scales used in  $T$  and the perceived

color. Studies of Somers used Cabernet Sauvignon and Shiraz wines. The findings revealed an 8-fold difference between the two varieties, whereas the hue parameters  $T$  and  $dA$  (%) only differed 2.4-fold, suggesting a higher potential of CI to differentiate between varieties, while the potential of  $T$  and  $dA$  (%) to differentiate between varieties is lower.<sup>15</sup> The findings of the presented study extend the findings of Somers using more varieties, concluding that the postulated potential of CI is high enough to depict significant differences between varieties, while this is not the case for  $T$  and  $dA$  (%).



**Figure 3.** CIE color space coordinates (A)  $L^*$ , (B)  $a^*$ , (C)  $b^*$ , (D)  $C^*$ , and (E)  $h^\circ$  for eight different red grape varieties, averaged for seven wines per grape variety. Error bars show the standard deviation of the seven wines of each grape variety. Letters show significance groups determined by ANOVA ( $p < 0.001$ ) and Tukey's HSD post hoc test ( $p < 0.001$ ).

**CIE Color Spaces.** The  $L^*$  coordinate was found to differentiate better than both the human eye and Glories' CI, because  $L^*$  differentiated Lemberger from Vernatsch in addition to the varieties that could be differentiated in prior sections (Figure 3A).

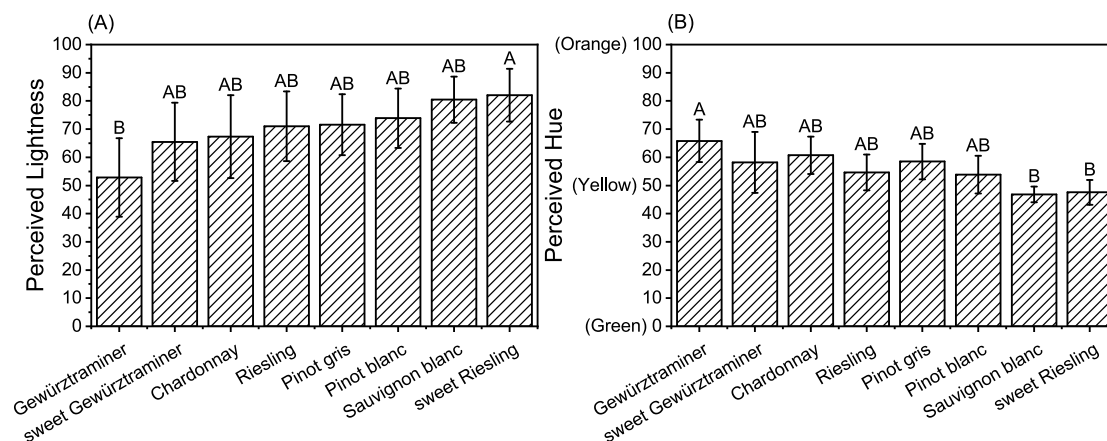
Like Glories' hue parameters, the  $a^*$  coordinate is not well-suited to differentiate between varieties because it could only differentiate between Pinot Noir and Shiraz. These findings show that the human eye differentiates the color of red wine not by their redness but by their color difference apart from the red color.

However, the  $b^*$  coordinate was better suited to distinguish between grape varieties, differentiating Pinot Noir from Tempranillo, Primitivo, Cabernet Sauvignon, and Shiraz as well as Vernatsch from Shiraz.

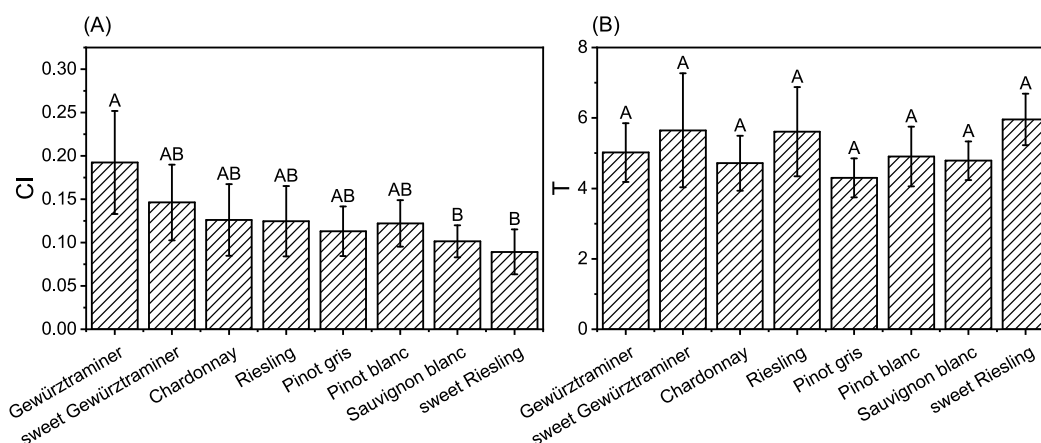
The conversion from Cartesian coordinates  $a^*$  and  $b^*$  to polar coordinates  $C^*$  and  $h^\circ$  resulted in a slightly different intervarietal differentiation. With  $C^*$ , it was possible to distinguish Pinot Noir from Shiraz, Tempranillo, and Cabernet Sauvignon as well as Vernatsch from Shiraz.  $h^\circ$  allowed for the same intervarietal differentiation as the  $b^*$  coordinate.

The four hue-related CIE coordinates  $b^*$ ,  $C^*$ , and  $h^\circ$  yielded similar results to the perceived hue. Therefore, it is possible to use CIE  $L^*a^*b^*$  or CIE  $L^*C^*h^\circ$  for intervarietal comparison, reflecting the ability of the human eye. These findings show that, counterintuitive to the definition of the CIE parameters, the  $a^*$  coordinate does not align very well with the perceived redness.

**White Wine. Visual Perception.** In contrast to the red wines (Figure 1A), the white wines could barely be distinguished by



**Figure 4.** Sensory ratings for (A) lightness and (B) hue for eight different white grape varieties/wine styles, averaged for seven wines per grape variety/wine style. Error bars show the standard deviation of the seven wines of each grape variety. Letters show significance groups determined by ANOVA ( $p < 0.001$ ) and Tukey's HSD post hoc test ( $p < 0.001$ ).



**Figure 5.** Glories' parameters (A) color intensity CI and (B) color hue  $T$  for eight different white grape varieties/wine styles, averaged for seven wines per grape variety/wine style. Error bars show the standard deviation of the seven wines of each grape variety. Letters show significance groups determined by ANOVA ( $p < 0.001$ ) and Tukey's HSD post hoc test ( $p < 0.001$ ).

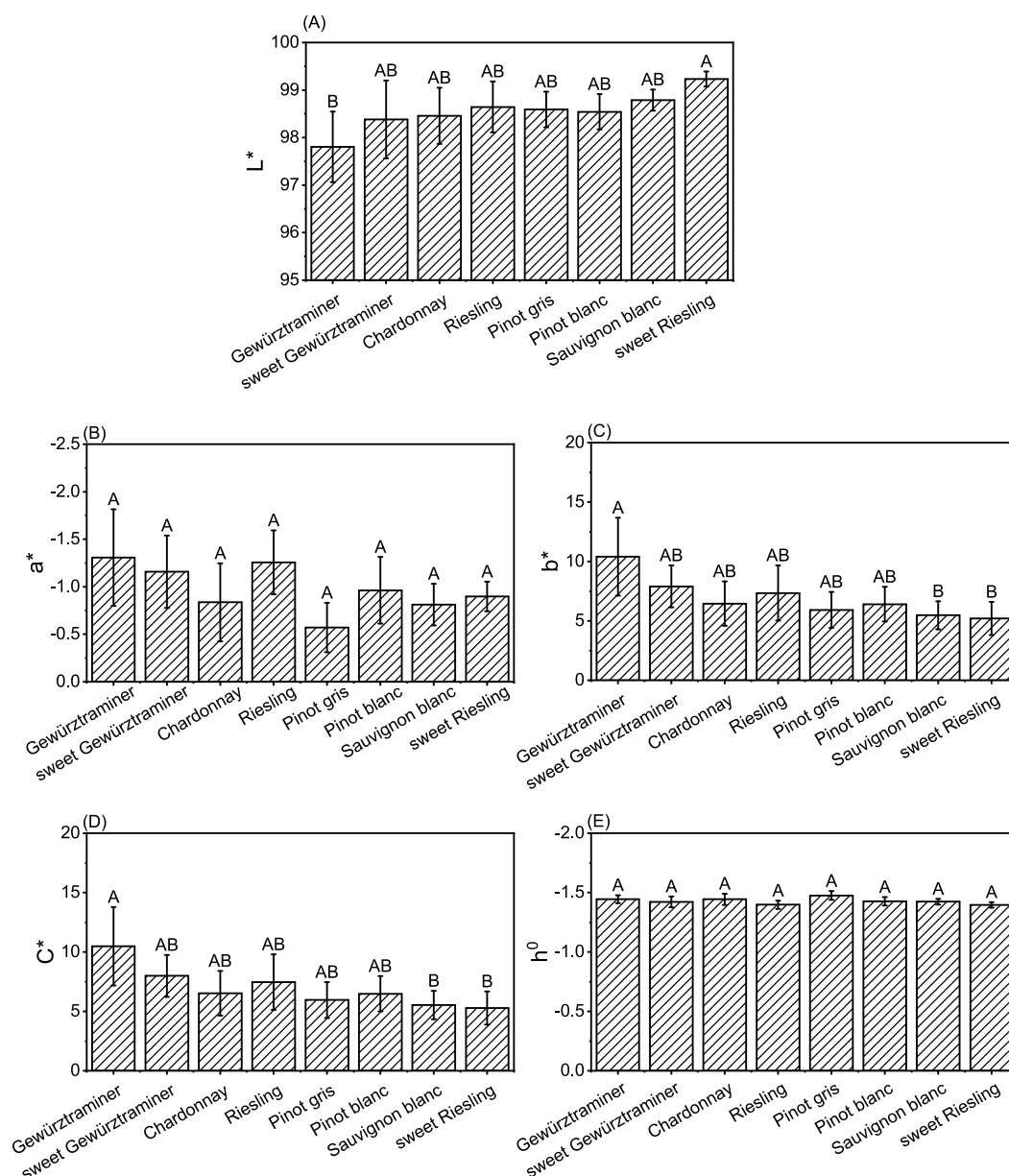
panelists in terms of lightness (Figure 4A), with only dry Gewürztraminer being significantly different from the sweet Riesling.

With regard to the perceived hue (Figure 4B), the sweet Riesling and dry Sauvignon blanc wines were found to differ significantly from the dry Gewürztraminer wines, adding one more significantly different variety in comparison to perceived lightness (Figure 4A). Overall, results from the sensory evaluation of white wine color agreed with results from the evaluation of the red wines, indicating that perceived hue shows a greater intervarietal variation and that perceived lightness shows a greater intravarietal variation. In general, more than one scale should be used to characterize white and red wine colors.

**Glories.** Figure 5A shows that it is possible to differentiate between white varieties with the Glories parameter CI. Gewürztraminer was the darkest variety, and sweet Riesling was the lightest variety. There are similarities to perceived lightness (Figure 5A). Additionally, with Glories' CI, it was possible to distinguish Gewürztraminer from Sauvignon Blanc. However, Glories'  $T$  has no significant power to distinguish between white wine varieties (Figure 4B), consistent with the results from the analysis of the red wines. Like in red wines, these findings do not discredit the use of the Glories method to

compare white wines, but rather they show the limitations for the varieties chosen for the current study of Glories' color measurement.

**CIE Color Spaces.** In Figure 6A, the  $L^*$  coordinate determined that Gewürztraminer wines were the darkest and sweet Riesling wines were the lightest, in agreement with lightness ratings from sensory evaluation. The  $a^*$  coordinate (Figure 6B) showed no intervarietal differentiation, also in agreement with the results from sensory evaluation. All wines, despite having negative  $a^*$  coordinates, were perceived as yellow or orange but not green. However, the  $b^*$  coordinate (Figure 6C) showed a similar intervarietal differentiation to Glories' CI, allowing for dry Gewürztraminer wines to be differentiated from sweet Riesling and Sauvignon Blanc, as was also the case with the perceived hue (Figure 4B). The  $C^*$  coordinate (Figure 6D) depicted differences in the perceived hue in the same way as the  $b^*$  coordinate. The  $h^0$  coordinate (Figure 6E) showed no potential to differentiate the varieties/wine styles, as was the case for red wines. Unlike in red wines, CIE  $L^*C^*h^0$  did not outperform CIE  $L^*a^*b^*$ . This can be explained by the small differences between the wines and the extremely small range between the  $a^*$  values, throughout all varieties.



**Figure 6.** CIE color space coordinates (A)  $L^*$ , (B)  $a^*$ , (C)  $b^*$ , (D)  $C^*$ , and (E)  $h^\circ$  for eight different white grape varieties/wine styles, averaged for seven wines per grape variety/wine style. Error bars show the standard deviation of the seven wines of each grape variety. Letters show significance groups determined by ANOVA ( $p < 0.001$ ) and Tukey's HSD post hoc test ( $p < 0.001$ ).

#### Determination of the Visible Color Threshold.

The smallest visible color difference for wine in different color areas, examined using the triangle test and wine blends, is summarized in Table 2. Values for the smallest visible color difference expressed as  $\Delta E_{ab}^*$  are generally higher than those expressed as  $\Delta E_{00}$ . Regardless of the color distance formula,

**Table 2.** Smallest Visible Color Difference in Different Color Areas Expressed as  $\Delta E_{ab}^*$  and  $\Delta E_{00}$

color area	$\Delta E_{ab}^*$	$\Delta E_{00}$
dark red	3.1	1.4
light red	8.1	3.4
light yellow	0.60	0.64
$\Delta E_{\max} - \Delta E_{\min}$	7.5	2.8

the smallest visible color difference is highest in the light red color area and lowest in the light yellow color area.

The smallest visible color difference, expressed as  $\Delta E_{ab}^*$ , in the dark red area is agreement with the smallest visible color difference postulated by Martinez et al.,<sup>34</sup> who also used dark red wines. It appears that no studies have been published regarding the smallest visible color difference of white wines. The difference in  $\Delta E_{ab}^*$  among the light yellow, light red, and dark red color areas may be explained by the tolerance ellipses investigated by Luo et al. that led to the development of the CIEDE2000 color difference formula. The tolerance ellipses describe a mathematical area within the CIE color space within which a color difference is not visible. These tolerance ellipses become smaller as they become closer to the point of neutrality in the CIE color space ( $a^* = 0$ ,  $b^* = 0$ ).<sup>45</sup> White wines are closest to this neutral point, followed by dark red wines, and



then light red wines. This explains the high variance among the smallest visible color differences calculated as Euclidean color distances. Calculation using the CIEDE2000 color difference formula to retrieve  $\Delta E_{00}$  led to a convergence of the smallest visible color differences for dark red, light red, and white wines: the range of the smallest visible color differences decreased 2.7-fold. The optimal result would be the same smallest visible color difference independent from the color area, but according to the studies by Melgosa et al., the STRESS value of the CIEDE2000 color difference formula is not zero, indicating a discrepancy between visual perception and computed color differences.<sup>36</sup>

In conclusion, this study established new insights into wine color analysis by evaluating light and dark red wines as well as white wines photometrically with the Glories color measurement and the CIE color spaces and visually with descriptive sensory analysis. The study extends knowledge from previous publications using a wide range of commercial wines from dark red to light red color as well as from yellow to orange hue. CIE  $L^*a^*b^*$  is said to be more accurate than Glories, but Glories' color measurement is easier to understand.<sup>9</sup> However, this study showed that not every CIE coordinate aligns with the perceived color. Vice versa, not every Glories parameter is unable to describe the perceived color. For red wines, Glories' CI and  $L^*$  were both able to represent the perceived lightness. The Glories parameters  $T$  and  $dA$  (%) were not able to do the same regarding the perceived hue. However, this was also the case for the  $a^*$  coordinate. The results show that the closest approximation to the perceived color hue could be achieved by the  $b^*$  coordinate, even for red wines, followed by  $C^*$  and  $h^0$  coordinates. The CIE  $L^*a^*b^*$  color space in total had a higher alignment with the perceived color for red wines; however, there are some discrepancies in the high saturated light red color area. There, the CIE  $L^*a^*b^*$  color space has problems to represent the perceived color. There are similarities between red and white wines, because Glories' CI and  $L^*$  could both represent the perceived lightness. Also,  $b^*$  was the CIE coordinate that approximated the perceived hue best. However, the  $h^0$  coordinate had no resolution for white wine as it had for red wine.

Furthermore, this study re-evaluated the visible color threshold for wine color using the newest standard CIEDE2000. The results show an improvement regarding the issue of uniformity. However, the CIE  $L^*a^*b^*$  color space in combination with the CIEDE2000 formula is still not completely equidistant, and further studies should investigate the smallest visible color threshold throughout the complete color space. Until further work has been carried out, it is advised to use the CIEDE2000 color distance formula and the presented visible color thresholds as guidelines for future studies.

## ■ ASSOCIATED CONTENT

### SI Supporting Information

The Supporting Information is available free of charge at <https://pubs.acs.org/doi/10.1021/acs.jafc.3c01284>.

Additional experimental details and material, including a three-dimensional (3D) model of the experimental setup (PDF)

## ■ AUTHOR INFORMATION

### Corresponding Author

**Dominik Durner** – Weincampus Neustadt,  
Dienstleistungszentrum Ländlicher Raum Rheinpfalz, 67435  
Neustadt an der Weinstraße, Germany;  
Email: [dominik.durner@hs-kl.de](mailto:dominik.durner@hs-kl.de)

### Authors

**Marcel Hensel** – Weincampus Neustadt,  
Dienstleistungszentrum Ländlicher Raum Rheinpfalz, 67435  
Neustadt an der Weinstraße, Germany; [orcid.org/0000-0003-4360-5328](https://orcid.org/0000-0003-4360-5328)

**Marina Scheiermann** – Weincampus Neustadt,  
Dienstleistungszentrum Ländlicher Raum Rheinpfalz, 67435  
Neustadt an der Weinstraße, Germany

**Jörg Fahrer** – Division of Food Chemistry and Toxicology,  
Department of Chemistry, RPTU Kaiserslautern-Landau,  
67663 Kaiserslautern, Germany

Complete contact information is available at:  
<https://pubs.acs.org/10.1021/acs.jafc.3c01284>

### Author Contributions

conceptualization, Marcel Hensel, Marina Scheiermann, and Dominik Durner; methodology, Marcel Hensel and Marina Scheiermann; software, Marcel Hensel; validation, Marcel Hensel; formal analysis, Marcel Hensel; investigation, Marcel Hensel and Marina Scheiermann; resources, Marcel Hensel; data curation, Marcel Hensel, writing—original draft preparation, Marcel Hensel; writing—review and editing, Jörg Fahrer and Dominik Durner; visualization Marcel Hensel; supervision, Jörg Fahrer and Dominik Durner; project administration, Marcel Hensel and Dominik Durner; and funding acquisition, Dominik Durner. All authors have read and agreed to the published version of the manuscript.

### Funding

This work was financially supported by the Arbeitsgemeinschaft industrieller Forschungsvereinigungen (AiF) "Otto von Guericke" e.V. (Project 20964).

### Notes

The authors declare no competing financial interest.

## ■ ACKNOWLEDGMENTS

The authors thank the working group of Scharfenberger-Schmeer as well as Roland Ulber, Sarah Di Nonno, and Yannick May from RPTU Kaiserslautern-Landau for their ideas and help as a project partner in our team. Furthermore, the authors thank Thi Nguyen for assistance with the English language and his ideas and help during the finalization of this manuscript.

## ■ ABBREVIATIONS USED

$a^*$ , red–green parameter within CIE  $L^*a^*b^*$ ; ANOVA, analysis of variance;  $b^*$ , yellow–blue parameter within CIE  $L^*a^*b^*$ ;  $C^*$ , chroma parameter within CIE  $L^*C^*h^0$ ; CIE, Commission Internationale de l'Éclairage; CMC, Color Measurement Committee;  $dA$  (%), brilliance parameter within the Glories method;  $h^0$ , hue angle within CIE  $L^*C^*h^0$ ; HSD, honest significant difference;  $L^*$ , lightness parameter within CIE  $L^*a^*b^*$  and CIE  $L^*C^*h^0$ ; OIV, Organisation Internationale de la Vin et du Vin; SDC, Society of Dyers and Colorists; STRESS, standardized residual sum of squares;  $T$ , hue

parameter within Glories;  $\Delta E_{00}$ , CIEDE2000 color distance;  $\Delta E_{ab}^*$ , Euclidean color distance

## REFERENCES

- Morrot, G.; Brochet, F.; Dubourdieu, D. The color of odors. *Brain Lang.* **2001**, *79*, 309–320.
- Peynaud, E.; Blouin, J. *El Gusto del Vino: El Gran Libro de la Degustación*, 2nd ed.; Mundi-Prensa: Madrid, Spain, 2002.
- Parpinello, G. P.; Versari, A.; Chinnici, F.; Galassi, S. Relationship among sensory descriptors, consumer preference and color parameters of Italian Novello red wines. *Food Res. Int.* **2009**, *42*, 1389–1395.
- Coulon-Leroy, C.; Pouzalgues, N.; Cayla, L.; Symoneaux, R.; Masson, G. Is the typicality of “Provence Rosé wines” only a matter of color? *OENO One* **2018**, *52*, 317–331.
- Issa-Issa, H.; Hernández, F.; Lipan, L.; López-Lluch, D.; Carbonell-Barrachina, Á. A. Quality, Nutritional, Volatile and Sensory Profiles and Consumer Acceptance of Fondillón, a Sustainable European Protected Wine. *Agronomy* **2021**, *11*, 1701.
- Bundesrat. *Änderungen und Entschließung zur Vierundzwanzigsten Verordnung zur Änderung der Weinverordnung*; Bundesrat: Berlin, Germany, March 26, 2021.
- Organisation Internationale de la Vin et du Vin (OIV). *Compendium of International Methods of Wine and Must Analysis*, 2021st ed.; OIV: Paris, France, 2021; Vol. 1.
- Commission Internationale de l’Eclairage (CIE). *CIE S 017/E:2020*; CIE: Vienna, Austria, 2020.
- Ribéreau-Gayon, P. *Handbook of Enology: Vol. 2: The Chemistry of Wine Stabilization and Treatments*, 2nd ed.; John Wiley: Chichester, U.K., 2006.
- Glories, Y. La couleur des vins rouges. 2e partie: mesure, origine et interprétation. *OENO One* **2016**, *18*, 253.
- Wenzel, K.; Ditttrich, H. H.; Heimfarth, M. Die Zusammensetzung der Anthocyane in den Beeren verschiedener Rebsorten. *Vitis* **1987**, *26* (2), 65–78.
- Durner, D.; Nickolaus, P.; Weber, F.; Trieu, H.-L.; Fischer, U. Evolution of Anthocyanin-Derived Compounds during Micro-Oxygenation of Red Wines with Different Anthocyanin-Flavanol Ratios. In *Advances in Wine Research*; Ebeler, S. B., Sacks, G., Vidal, S., Winterhalter, P., Eds.; American Chemical Society (ACS): Washington, D.C., 2015; ACS Symposium Series, Vol. 1203, Chapter 16, pp 253–274, DOI: 10.1021/bk-2015-1203.ch016.
- Puértolas, E.; Hernández-Orte, P.; Sladaña, G.; Álvarez, I.; Raso, J. Improvement of winemaking process using pulsed electric fields at pilot-plant scale. Evolution of chromatic parameters and phenolic content of Cabernet Sauvignon red wines. *Food Res. Int.* **2010**, *43*, 761–766.
- Pérez-Magariño, S.; González-San José, M. L. Evolution of flavanols, anthocyanins, and their derivatives during the aging of red wines elaborated from grapes harvested at different stages of ripening. *J. Agric. Food Chem.* **2004**, *52*, 1181–1189.
- Somers, C. *The Wine Spectrum: An Approach Towards Objective Definition of Wine Quality*; Winetitles: Broadview, South Australia, Australia, 1998.
- Hensel, M.; Di Nonno, S.; Mayer, Y.; Scheiermann, M.; Fahrner, J.; Durner, D.; Ulber, R. Specification and Simplification of Analytical Methods to Determine Wine Color. *Processes* **2022**, *10*, 2707.
- Pérez-Magariño, S.; González-San José, M. L. Application of absorbance values used in wineries for estimating CIELAB parameters in red wines. *Food Chem.* **2003**, *81*, 301–306.
- Vaimakis, V.; Roussis, I. G. Must Oxygenation and Polyphenoloxidase Inhibition and the Oxidation of White Wine. *LWT - Food Sci. Technol.* **1993**, *26*, 133–137.
- Ellis, L. P.; Kok, C. Colour Changes in Blanc de Noir Wines during Ageing at Different Temperatures and its Colour Preference Limits. *S. Afr. J. Enol. Vitic.* **2017**, *8*, 16–22.
- Wright, W. D. A re-determination of the mixture curves of the spectrum. *Trans. Opt. Soc.* **1930**, *31*, 201–218.
- Wright, W. D. A re-determination of the trichromatic coefficients of the spectral colours. *Trans. Opt. Soc.* **1929**, *30*, 141–164.
- Guild, J. The colorimetric properties of the spectrum. *Philos. Trans. R. Soc. London, Ser. A* **1931**, *230*, 149–187.
- Stiles, W. S.; Burch, J. M. N.P.L. Colour-matching Investigation: Final Report (1958). *Int. J. Opt.* **1959**, *6*, 1–26.
- Bakker, J.; Bellworthy, S. J.; Reader, H. P.; Watkins, S. J. Effect of Enzymes During Vinification on Color and Sensory Properties of Port Wines. *Am. J. Enol. Vitic.* **1999**, *50*, 271–276.
- Harbertson, J. F.; Parpinello, G. P.; Heymann, H.; Downey, M. O. Impact of exogenous tannin additions on wine chemistry and wine sensory character. *Food Chem.* **2012**, *131*, 999–1008.
- Skinkis, P. A.; Bordelon, B. P.; Butz, E. M. Effects of Sunlight Exposure on Berry and Wine Monoterpenes and Sensory Characteristics of Traminette. *Am. J. Enol. Vitic.* **2010**, *61*, 147–156.
- Lan, H.; Li, S.; Yang, J.; Li, J.; Yuan, C.; Guo, A. Effects of light exposure on chemical and sensory properties of storing Meili Rosé wine in colored bottles. *Food Chem.* **2021**, *345*, 128854.
- McRae, J. M.; Schulkin, A.; Kassara, S.; Holt, H. E.; Smith, P. A. Sensory properties of wine tannin fractions: implications for in-mouth sensory properties. *J. Agric. Food Chem.* **2013**, *61*, 719–727.
- Commission Internationale de l’Eclairage (CIE). *CIE 015:2018 Colorimetry*, 4th ed.; CIE: Vienna, Austria, 2018.
- Deutsches Institut für Normung e.V. (DIN). *Colorimetry—Part 6: CIEDE2000 Colour-Difference Formula*, 12th ed.; Beuth Verlag GmbH: Berlin, Germany, 2016.
- McDonald, R.; Smith, K. J. CIE94—A new colour-difference formula. *J. Soc. Dyers Colour.* **1995**, *111*, 376–379.
- Brun, L.; Sharma, G.; Allebach, J. P.; Sharma, G.; de Queiroz, R. L.; Bala, R.; Klassen, R. V.; Spaulding, K. E.; Parulski, K. A.; Baqai, F. A.; Johnson, G. M.; Fairchild, M. D.; Emmel, P.; Morovic, J.; Balasubramanian, R.; Treméau, A.; Agar, U.; Giorgianni, E.; Madden, T.; Knox, K.; Hains, C.; Wang, S.-g. *Digital Color Imaging Handbook*; CRC Press: Boca Raton, FL, 2003; DOI: 10.1201/9781420041484.
- Roehrig, H.; Hashmi, S. F.; Dallas, W. J.; Krupinski, E. A.; Rehm, K.; Fan, J. Medical color displays and their color calibration: Investigations of various calibration methods, tools, and potential improvement in color difference  $\Delta E$ . *Proceedings of the SPIE Optical Engineering + Applications*; San Diego, CA, Aug 3–5, 2010; DOI: 10.1117/12.864145.
- Martínez, J. A.; Melgosa, M.; Pérez, M. M.; Hita, E.; Negueruela, A. I. Note. Visual and Instrumental Color Evaluation in Red Wines. *Food Sci. Technol. Int.* **2001**, *7*, 439–444.
- García, P. A.; Huertas, R.; Melgosa, M.; Cui, G. Measurement of the relationship between perceived and computed color differences. *J. Opt. Soc. Am. A* **2007**, *24*, 1823–1829.
- Melgosa, M.; Huertas, R.; Berns, R. S. Performance of recent advanced color-difference formulas using the standardized residual sum of squares index. *J. Opt. Soc. Am. A* **2008**, *25*, 1828–1834.
- Melgosa, M. *CIE/ISO New Standard*; CIEDE2000: Leeds, U.K., July 4, 2013.
- Panchal, G.; Mehta, A.; Nair, G.; Patel, D.; Naik, S. A Comparative Study of Color Perception in Young Males and Females. *Southeast Asian J. Case Rep. Rev.* **2012**, *1*, 30.
- Fairchild, M. The colors of wine. *Int. J. Wine Res.* **2018**, *10*, 13–31.
- Deutsches Institut für Normung e.V. (DIN). *DIN 5033-7:2014-10, Farbmessung - Teil 7: Messbedingungen für Körperfarben*; Beuth Verlag GmbH: Berlin, Germany, 2014.
- Ishihara, K.; Ishihara, S.; Nagamachi, M.; Hiramatsu, S.; Osaki, H. Age-related decline in color perception and difficulties with daily activities-measurement, questionnaire, optical and computer-graphics simulation studies. *Int. J. Indust. Ergon.* **2001**, *28*, 153–163.
- Farnsworth, D. The Farnsworth-Munsell 100-Hue and Dichotomous Tests for Color Vision. *J. Opt. Soc. Am.* **1943**, *33*, 568–578.
- Durner, D.; Weber, F.; Neddermeyer, J.; Koopmann, K.; Winterhalter, P.; Fischer, U. Sensory and Color Changes Induced by

Microoxygenation Treatments of Pinot noir before and after Malolactic Fermentation. *Am. J. Enol. Vitic.* **2010**, *61*, 474–485.

(44) Zhang, Y.-S.; Du, G.; Gao, Y.-T.; Wang, L.-W.; Meng, D.; Li, B.-J.; Brennan, C.; Wang, M.-Y.; Zhao, H.; Wang, S.-Y.; Guan, W.-Q. The Effect of Carbonic Maceration during Winemaking on the Color, Aroma and Sensory Properties of 'Muscat Hamburg' Wine. *Molecules* **2019**, *24*, 3120.

(45) Luo, M. R.; Cui, G.; Rigg, B. The development of the CIE 2000 colour-difference formula: CIEDE2000. *Color Res. Appl.* **2001**, *26*, 340–350.



## Supplementary Data

Table S 1: Variety, origin, vintage, and producer of the red wines used in the study

	<b>Variety</b>	<b>Origin</b>	<b>Vintage</b>	<b>Producer</b>
1	<b>Cabernet Sauvignon</b>	France	2018	Foncalieu
2	<b>Cabernet Sauvignon</b>	Italy	2018	Castel Firmian
3	<b>Cabernet Sauvignon</b>	USA	2014	Two Vines
4	<b>Cabernet Sauvignon</b>	Germany	2018	Weingut Metzger
5	<b>Cabernet Sauvignon</b>	Australia	2016	Taylors/Wakefield
6	<b>Cabernet Sauvignon</b>	Spain	2016	Bodega Miguel Torres
7	<b>Cabernet Sauvignon</b>	Italy	2019	Feudo Arancio
8	<b>Merlot</b>	Italy	2018	Cielo e Terra
9	<b>Merlot</b>	France	2018	Monteran
10	<b>Merlot</b>	USA	2018	Two Vines
11	<b>Merlot</b>	France	2016	Jean Pierre Moueix
12	<b>Merlot</b>	New Zealand	2018	Sileni Cellar
13	<b>Merlot</b>	France	2019	SAS Aussières
14	<b>Merlot</b>	USA	2017	Two Vines
15	<b>Pinot Noir</b>	France	2019	Les Producteurs Reunis
16	<b>Pinot Noir</b>	Germany	2019	Weingut Peth-Wetz
17	<b>Pinot Noir</b>	Germany	2013	Weingut Spiess
18	<b>Pinot Noir</b>	Germany	2017	Weingut Stern
19	<b>Pinot Noir</b>	Germany	2018	Kloster Eberbach
20	<b>Pinot Noir</b>	New Zealand	2018	Private Bin
21	<b>Pinot Noir</b>	New Zealand	2019	Rapaura Springs
22	<b>Primitivo</b>	Italy	2020	San Marziano Vino
23	<b>Primitivo</b>	Italy	2017	Hess Collection winery

<b>24</b>	<b>Primitivo</b>	Italy	2020	S.R.L Ortona
<b>25</b>	<b>Primitivo</b>	Italy	2019	Cantine Torrovento
<b>26</b>	<b>Primitivo</b>	Italy	2017	Cantine Torrovento
<b>27</b>	<b>Primitivo</b>	USA	2018	Michael David winery
<b>28</b>	<b>Primitivo</b>	USA	2017	Boutinot
<b>29</b>	<b>Syrah</b>	Australia	2018	Koonunga Hill
<b>30</b>	<b>Syrah</b>	Italy	2019	Feudo Arancio
<b>31</b>	<b>Syrah</b>	USA	2017	Charles Smith winery
<b>32</b>	<b>Syrah</b>	France	2019	Gerard Bertrand
<b>33</b>	<b>Syrah</b>	New Zealand	2019	Sileni Cellar
<b>34</b>	<b>Syrah</b>	France	2020	Château de San Cosme
<b>35</b>	<b>Syrah</b>	France	2019	La Vie Joulie Rouge
<b>36</b>	<b>Tempranillo</b>	Spain	2015	Boedgas Faustino
<b>37</b>	<b>Tempranillo</b>	Spain	2015	Pagos del Rey
<b>38</b>	<b>Tempranillo</b>	Spain	2018	Pagos del Rey
<b>39</b>	<b>Tempranillo</b>	Spain	2019	Hammeken Cellars
<b>40</b>	<b>Tempranillo</b>	Spain	2017	Bodegas Faustino
<b>41</b>	<b>Tempranillo</b>	Spain	2018	Gran Sello
<b>42</b>	<b>Tempranillo</b>	Spain	2012	Marqué de Sandoval
<b>43</b>	<b>Vernatsch</b>	Germany	2018	Collegium Wirttemberg
<b>44i</b>	<b>Vernatsch</b>	Germany	2018	Collegium Wirttemberg
<b>45</b>	<b>Vernatsch</b>	Italy	2019	Kellerei St. Michael Eppan
<b>46</b>	<b>Vernatsch</b>	Italy	2019	Schlosskellerei Turmhof Tiefenbrunner
<b>47</b>	<b>Vernatsch</b>	Italy	2019	Weingut Castelfelder
<b>48</b>	<b>Vernatsch</b>	Italy	2019	Kellerei Caltern
<b>49</b>	<b>Vernatsch</b>	Germany	2018	Weingut Uhlbacher

<b>50</b>	<b>Lemberger</b>	Austria	2018	Weingut Ernst
<b>51</b>	<b>Lemberger</b>	Germany	2018	Graf Neipperg
<b>52</b>	<b>Lemberger</b>	Austria	2017	Weingut Ernst
<b>53</b>	<b>Lemberger</b>	Germany	2019	Weingut Karl Haidle
<b>54</b>	<b>Lemberger</b>	Germany	2016	Collegium Wirttemberg
<b>55</b>	<b>Lemberger</b>	Germany	2019	Weingut Wöhrwag
<b>56</b>	<b>Lemberger</b>	Germany	2019	Weingut Aldinger

*Table S 2: Variety, origin, vintage, and producer of the white wines used in the study*

	<b>Variety</b>	<b>Origin</b>	<b>Vintage</b>	<b>Producer</b>
<b>1</b>	<b>Chardonnay</b>	France	2018	Côtes des Roses
<b>2</b>	<b>Chardonnay</b>	Italy	2020	Feudo Arancio
<b>3</b>	<b>Chardonnay</b>	Germany	2018	Domäne Bergstraße
<b>4</b>	<b>Chardonnay</b>	Spain	2019	Bodega Pago de Cirsus
<b>5</b>	<b>Chardonnay</b>	Germany	2016	Von Buhl
<b>6</b>	<b>Chardonnay</b>	Australia	2019	Koonunga Hill
<b>7</b>	<b>Chardonnay</b>	USA	2018	Canyon Road winery
<b>8</b>	<b>Riesling</b>	Germany	2015	Weingut Schätzel
<b>9</b>	<b>Riesling</b>	France	2017	Cave Vinicole de Cléebourg
<b>10</b>	<b>Riesling</b>	USA	2013	Chateau Ste Michelle
<b>11</b>	<b>Riesling</b>	Germany	2018	Weingut M. Pfaffmann

<b>12</b>	<b>Riesling</b>	Germany	2019	Weingut Stern
<b>13</b>	<b>Riesling</b>	Germany	2020	Weingut Bassermann-Jordan
<b>14</b>	<b>Riesling</b>	Germany	2020	Weingut Rings
<b>15</b>	<b>Sweet Riesling</b>	Germany	2018	Weingut Winkel
<b>16</b>	<b>Sweet Riesling</b>	Germany	2019	Weingut Metzger
<b>17</b>	<b>Sweet Riesling</b>	Germany	2019	Weingut St. Urbans-Hof Saar
<b>18</b>	<b>Sweet Riesling</b>	Germany	2018	Weingut Winkel
<b>19</b>	<b>Sweet Riesling</b>	Germany	2019	Weingut Dr. Loosen
<b>20</b>	<b>Sweet Riesling</b>	Germany	2018	Weingut Hamm
<b>21</b>	<b>Sweet Riesling</b>	Germany	2019	Weingut Rosch Leiwener
<b>22</b>	<b>Pinot gris</b>	Italy	2019	S.R.L Ortona
<b>23</b>	<b>Pinot gris</b>	Italy	2020	Grifòn
<b>24</b>	<b>Pinot gris</b>	France	2017	Les Prince Abbes
<b>25</b>	<b>Pinot gris</b>	New Zealand	2018	Sileni Cellar
<b>26</b>	<b>Pinot gris</b>	Germany	2019	Winzergenossenschaft Königschaffhauser
<b>27</b>	<b>Pinot gris</b>	Germany	2019	Hundertmorgen
<b>28</b>	<b>Pinot gris</b>	Germany	2019	Weingut Kaiser

<b>29</b>	<b>Sauvignon blanc</b>	France	2019	La Fleur Saint-Michel
<b>30</b>	<b>Sauvignon blanc</b>	Germany	2019	Weingut M Pfaffmann
<b>31</b>	<b>Sauvignon blanc</b>	Italy	2019	I Palazzi
<b>32</b>	<b>Sauvignon blanc</b>	USA	2019	Noble Wines
<b>33</b>	<b>Sauvignon blanc</b>	New Zealand	2020	Sileni Cellar
<b>34</b>	<b>Sauvignon blanc</b>	France	2018	Mas Le Chevalière
<b>35</b>	<b>Sauvignon blanc</b>	Germany	2018	Weingut Michael Schneider
<b>36</b>	<b>Pinot blanc</b>	Germany	2019	Weingut Bickensohler
<b>37</b>	<b>Pinot blanc</b>	Germany	2020	Weingut Karl Pfaffmann
<b>38</b>	<b>Pinot blanc</b>	Germany	2018	Weingut Stefan Winter
<b>39</b>	<b>Pinot blanc</b>	Germany	2017	Weingut Schwedhelm
<b>40</b>	<b>Pinot blanc</b>	Germany	2020	Weingut Bassermann-Jordan
<b>41</b>	<b>Pinot blanc</b>	Italy	2019	Schreckbichl Colterenzio
<b>42</b>	<b>Pinot blanc</b>	France	2017	Les Princes Abbes
<b>43</b>	<b>Gewürztraminer</b>	Germany	2019	Weingut Kesselring
<b>44</b>	<b>Gewürztraminer</b>	France	2018	Domaine Paul Blanck
<b>45</b>	<b>Gewürztraminer</b>	Spain	2018	Vinos Divertidos
<b>46</b>	<b>Gewürztraminer</b>	Germany	2018	Weingut Philipp Kuhn
<b>47</b>	<b>Gewürztraminer</b>	Germany	2020	Weingut Kesselring

<b>48</b>	<b>Gewürztraminer</b>	Italy	2019	St. Michel Eppan
<b>49</b>	<b>Gewürztraminer</b>	Italy	2019	Weingut Kornell
<b>50</b>	<b>Sweet Gewürztraminer</b>	France	2019	André Stentz
<b>51</b>	<b>Sweet Gewürztraminer</b>	France	2019	Antoine Heinrich
<b>52</b>	<b>Sweet Gewürztraminer</b>	Spain	2019	Vinas del Vero
<b>53</b>	<b>Sweet Gewürztraminer</b>	Germany	2016	Weingut Tina Pfaffmann
<b>54</b>	<b>Sweet Gewürztraminer</b>	Germany	2019	Weingut Hensel
<b>55</b>	<b>Sweet Gewürztraminer</b>	Germany	2018	P.J. Valkenberg
<b>56</b>	<b>Sweet Gewürztraminer</b>	Germany	2019	Weingut A. Diehl

---

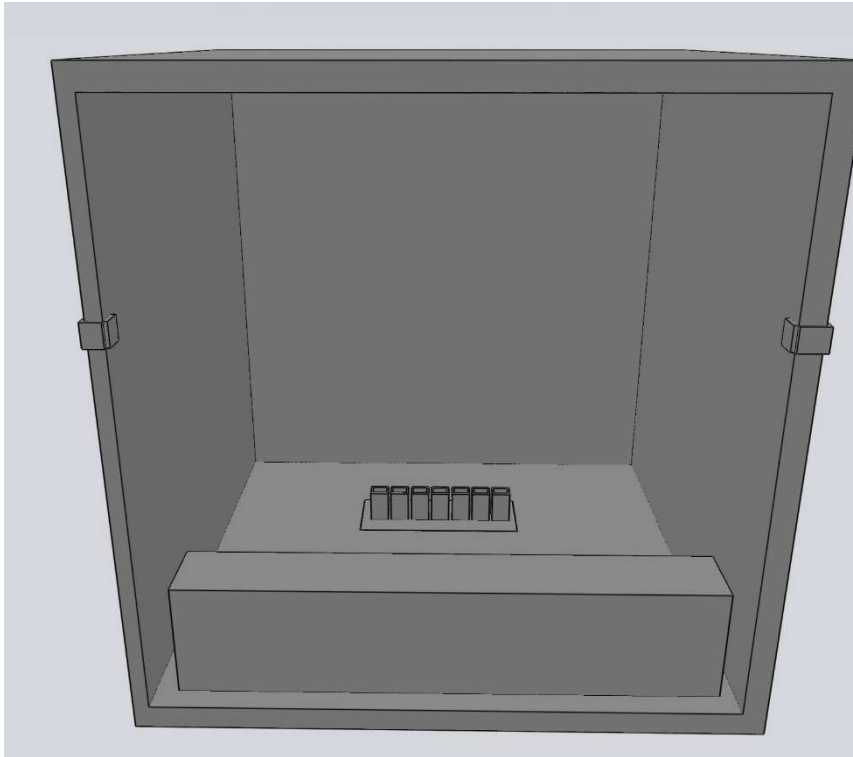


Figure S1: 3D Model of the sensory box.

### CIEDE2000 transformation according to literature

$$L' = L^*$$

$$a' = (1 + G) \cdot a^*$$

$$b' = b^*$$

$$C' = (a'^2 + b'^2)^{1/2}$$

$$\Delta L' = L'_P - L'_R$$

$$\Delta C' = C'_P - C'_R$$

$$G = 0,5 \cdot \left( 1 - \sqrt{\frac{(C_{ab}^*)^7}{(C_{ab}^*)^7 + 25^7}} \right)$$

$$\overline{C_{ab}^*} = \frac{C_P^* + C_R^*}{2}$$

$$\Delta h' = \begin{cases} 0^\circ; C'_R \cdot C'_P = 0 \\ h'_P - h'_R; C'_P \cdot C'_R \neq 0; |h'_R - h'_P| \leq 180^\circ \\ h'_P - h'_R - 360^\circ; C'_R \cdot C'_P \neq 0; (h'_R - h'_P) > 180^\circ \\ h'_P - h'_R + 360^\circ; C'_R \cdot C'_P \neq 0; (h'_R - h'_P) < 180^\circ \end{cases}$$

$$h' = \begin{cases} \arctan\left(\frac{b'}{a'}\right) a' > 0; b' \geq 0 \\ \arctan\left(\frac{b'}{a'}\right) + 360^\circ, a' > 0; b' < 0 \\ \arctan\left(\frac{b'}{a'}\right) + 180^\circ, a' < 0 \\ 90^\circ, a' = 0; b' > 0 \\ 270^\circ, a' = 0; b' < 0 \\ 0, a' = 0; b' = 0 \end{cases}$$

$$\Delta H' = 2 \cdot (C'_R \cdot C'_P)^{\frac{1}{2}} \cdot \sin\left(\frac{\Delta h'}{2}\right)$$



$$S_L = 1 + 0,015 \cdot \frac{(\bar{L}' - 50)^2}{\sqrt{20 + (\bar{L}' - 50)^2}}$$

$$S_C = 1 + 0,045 \cdot \bar{C}'$$

$$S_H = 1 + 0,015 \cdot \bar{C}' \cdot T$$

$$T = 1 - 0,17 \cdot \cos(\bar{h}' - 30^\circ) + 0,24 \cdot \cos(2 \cdot \bar{h}') + 0,32 \cdot \cos(3 \cdot \bar{h}' + 6^\circ) - 0,2 \cdot (4 \cdot \bar{h}' - 63^\circ)$$

$$R_T = -\sin(2\Delta\theta) \cdot R_C$$

$$\Delta\theta = 30^\circ \cdot \exp\left(-\left[\frac{\bar{h}' - 275^\circ}{2}\right]^2\right)$$

$$R_C = 2 \cdot \sqrt{\frac{(\bar{C}')^7}{(\bar{C}')^7 + 25^7}}$$

$$\bar{h}' = \begin{cases} \frac{h'_R + h'_P}{2}; |h'_R - h'_P| \leq 180^\circ; \bar{C}'_R \cdot \bar{C}'_P \neq 0 \\ \frac{h'_R + h'_P + 360^\circ}{2}; |h'_R - h'_P| > 180^\circ; (h'_R + h'_P) < 360^\circ; \bar{C}'_R \cdot \bar{C}'_P \neq 0 \\ \frac{h'_R + h'_P - 360^\circ}{2}; |h'_R - h'_P| > 180^\circ; (h'_R + h'_P) \geq 360^\circ; \bar{C}'_R \cdot \bar{C}'_P \neq 0 \\ h'_R + h'_P; \bar{C}'_R \cdot \bar{C}'_P = 0 \end{cases}$$

### **5.3 Conference Proceeding- Is there a need to redefine the methods to evaluate wine color?**

This publication extends the knowledge provided in section 5.1 and 5.2. To validate the comparison between the CIE L\*a\*b\* color space and Glories' color measurement, the experiment for red wine was recreated using 925 red wines. A total number of 617 wines were considered dark red wine and 306 wines were considered light red wine. The results from section 5.1 could be validated as even with this high number of samples the two methods correlate well with each other in dark red wine and very poorly in light red wine and white wine.

Contrary to the approach seen in section 5.2, this publication considers the wine-to-wine approach. Using different regressions, the perceived color was compared to the analytical color. The results show a linear dependence between the perceived lightness and the analytical lightness ( $L^*$  and CI). Using the CIE L\*a\*b\* color space, the regression is somewhat linear in the low lightness area, but deviated from the perceived color in lighter red wines. This explains the discrepancies seen in with lighter red wines in section 5.2. The color according to Glories shows no correlation on a wine-to-wine basis. The results for white wine are similar. The CIE L\*a\*b\* color space resembles the perceived color better than Glories and the color according to Glories shows a weak correlation.

## **Conference Proceeding- Is there a need to redefine the methods to evaluate wine color?**

Authors:

Dominik Durner<sup>1</sup>, Marcel Hensel<sup>1</sup>

Affiliations:

<sup>1</sup> Weincampus Neustadt, Dienstleistungszentrum ländlicher Raum Rheinpfalz, 67435 Neustadt an der Weinstraße, Germany

### **Disclaimer:**

**The presented work was exclusively performed by the DLR Rheinpfalz (Marcel Hensel and master thesis Marina Scheiermann).**

Citation:

Is there a need to re-define the methods to evaluate wine color?, Dominik Durner, Marcel Hensel, BIO Web Conf. 68 02008 (2023), DOI: 10.1051/bioconf/20236802008.

Copyright:

©2023 by the authors. Licensee EDP Science. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license.

## Is there a need to redefine the methods to evaluate wine color?

Dominik Durner<sup>1</sup>, and Marcel Hensel<sup>1\*</sup>

<sup>1</sup>Weincampus Neustadt, Dienstleistungszentrum Ländlicher Raum Rheinpfalz, 67435 Neustadt an der Weinstraße, Germany

**Abstract.** Both the Glories method and the CIE  $L^*a^*b^*/L^*C^*h^0$  system are used in analytical protocols to express the chromatic characteristics of wine [1–4]. Both methods allow for the interpretation of color intensity, color hue, and other chromatic characteristics. They are utilized in scientific studies to compare experimental treatments and many innovations had been assessed on whether they could retain a higher color intensity or yield more color brilliance. The validity of the two methods and the explanatory power of the respective color parameters were investigated by comparing the two methods with data from a sensory study. Data intervals from 0.5 to 5 nm and different scan speeds were tested as spectrophotometer settings. A trained panel evaluated 112 dark red, light red and white wines from different grape varieties, origins, and vintages. The correlation between Glories' system and the CIE color space was found to be strong for dark red wines ( $r > -0.9$ ). Lighter red wines and white wines ( $L^* > 20$ ) do not correlate with the same quality, with the exception of the correlation between CI and  $L^*$  ( $r = -0.9$ ). In comparison to the human perception, the red and white wines could not be distinguished well with Glories' system. The CIE color space was found to be more suited to depict the perceived color for red and white wines. In recent years, the CIE re-defined the color distance calculation within the CIE  $L^*a^*b^*$  color space due to such non-uniformities. The CIEDE2000 color distance  $\Delta E_{00}$  was proposed [4–6] for a better approximation of the perceived color than the Euclidean color distance  $\Delta E_{ab}$ . A just noticeable difference (JND) test was carried out using triangle testing for the wines segmented into three wine color areas: dark red, light red, and light yellow. For dark red wines, the JND was  $\Delta E_{00} = 1.4$  or  $\Delta E_{ab} = 3.1$ , respectively. For light yellow wines, the JND yielded much lower values at  $\Delta E_{00} = 0.64$  or  $\Delta E_{ab} = 0.60$ , respectively. And for light red wines, the JND was found to be very high at  $\Delta E_{00} = 3.4$  or  $\Delta E_{ab} = 8.1$ , respectively. The CIEDE2000 formula resulted in lower JND values than the Euclidean distance. Although still not uniform for the huge diversity of wine colors, the  $\Delta E_{00}$  provides better comparability and is suggested for future studies.

### 1 Introduction

The color of wine is known to alter sensorial impressions like aroma and mouthfeel. The relationship between aroma and color was established by Morrot *et al.*, where panelists described

\*e-mail: dominik.durner@hs-kl.de

a red colored white wine with typical aroma descriptors for red wine, while the native uncolored white wine was described with typical white wine aroma descriptors [7]. A recent study by Nguyen *et al.* was conducted to measure the impact of the knowledge of the wine color on the aromatic profiles of white and red wines. The intensity of the odors green fruit, citrus fruit, and stone fruit increased in coherence with the knowledge of the color in white wine. The odors red fruit, dark fruit, and oak increased in coherence with the color in red wine [8]. Apart from the correlation between color and aroma profiles, the color of wine is also important for consumer preference, perceived typicality, and regulations [9–11]. The results of these findings and regulations suggest the importance of an objective, generally applicable, and reproducible method. The wine color can be evaluated with three different methods. The color of a wine can be evaluated sensorially. The advantages of this method are the general applicability and the independence from mathematical approximations. A panel can be used for every purpose, research question, and wine type. However, this approach requires a sufficient number of trained panelists to maintain objectivity as prior studies demonstrated the subjectivity of a sensory evaluation [8]. Furthermore, it is very time consuming and requires a vast amount of organization. Other methods like the Glories color measurement and the CIE color spaces use photometry to evaluate the wine color [1, 12]. The advantages of these methods lie in their objectivity, their simplicity, and the required time. The disadvantages of the photometric methods differs depending on the method. Glories color measurement comprising of the absorbance measurements at 420 nm, 520 nm, and 620 nm is used to calculate the Color Intensity (CI), the hue (T), and the Brilliance (dA(%)), has a disadvantage regarding the general applicability. The Glories method was designed with and for dark red wines [12]. Therefore, the use of Glories measurement for any other wine type or style has to be called into question. It is worth to notice that the Glories color measurement is recommended by the OIV [3]. Besides the Glories measurement the CIE  $L^*a^*b^*/L^*C^*h^0$  color space is also widely used and recommended by the OIV for the measurement of wine color [3]. Here, the complete transmission spectrum, the relative sensitivity of the human eye, known as the color matching function, as well as the emitted light of the standard illuminant D65 is used to calculate the Cartesian coordinates  $L^*$ ,  $a^*$ , and  $b^*$ . The Cartesian coordinates can also be transformed to polar coordinates  $C^*$  and  $h^0$ . In both coordinate systems represents the  $L^*$  coordinate the lightness of an object, whereas  $a^*$  the redness if positive, or the greenness if negative of an object represents. The  $b^*$  coordinate describes either the yellowness if positive, or the blueness if negative of an object. The  $C^*$  coordinate, also known as the color vector, describes the saturation of an object. Another CIE  $L^*a^*b^*/L^*C^*h^0$  parameter  $h^0$ , also known as the hue angle, starts at  $0^\circ$  ( $360^\circ$ ) with the redness. At  $90^\circ$  the yellow hue is noted and  $180^\circ$  represents the greenness, whereas  $270^\circ$  stands for blue [2, 4, 13]. In general, the CIE  $L^*a^*b^*/L^*C^*h^0$  color space is considered more accurate, since it uses the relative sensitivity of the human eye to mimic the color response of the human eye, whereas Glories' measurement is easier to understand [12]. For the transmission spectrum measurement, the International Organization for Wine and Vine (OIV) recommends a data interval of 5 nm. However, no recommendations are given for any other photometer parameter [3]. The crucial difference between Glories' color measurement and CIE  $L^*a^*b^*/L^*C^*h^0$  is the ability to calculate the distance between two color points. For wine color the OIV represents the CIE76 Euclidean color distance. However, the CIE recommends the CIEDE2000 color distance [4?–6]. It has yet to be established how the CIEDE 2000 color distance formula impacts the JND in wine.

The objective of the presented studies is to contribute to the better understanding of color measurement of wine. It needs to be established how and if the parameter settings on the photometer and different photometers affects the CIE  $L^*a^*b^*/L^*C^*h^0$  color space [14]. Since both, Glories' color measurement and CIE  $L^*a^*b^*/L^*C^*h^0$ , are recommended for the use in

wine, it needs to be evaluated if the methods can be used interchangeably and which of the methods depict the human color perception better [14, 15]. Previous studies indicating that this is not the case for light red colored wine. With a sample size of 56 red wines a correlation was performed for dark red wine with  $L^* \leq 20$  ( $n=34$ ) and light red wines with  $L^* > 20$  ( $n=11$ ) [14]. These results have to be validated with a bigger sample size. The last part is to re-evaluate the JND with the CIEDE2000 color distance formula and compare the results to the Euclidean color difference [15].

## 2 Materials and Methods

### 2.1 Evaluating the impact of the photometric parameters

#### 2.1.1 Samples

Seven Merlot wines as an example for red wine were used in the study to measure the effect of the photometric parameters. Seven Chardonnay wine were used as an example for white wines. The red wines originated from France, Italy, United States of America, and New Zealand. The vintages ranged from 2016 to 2019. The white wines originated from France, Italy, Germany, Australia, and the USA. Here, the vintages also ranged from 2016 to 2019 [14].

#### 2.1.2 Photometric measurement

Complete transmission spectra from 200 to 900 nm were recorded on a JASCO double-beam photometer. The CIE  $L^*a^*b^*/L^*C^*h^0$  calculations were performed between 360 nm and 830 nm using a flow-through quartz-cuvette with a path length of 1 mm for red wine, whereas for white wines 10 mm polystyrene cuvettes were used. The experiments were done in triplicates and validated on a Varian Cary 100. The 14 wines were measured with data interval ranging from 0.5 nm, 1 nm, 5 nm measured with the a scan speed of 100 nm/min and 1000 nm/min (Table 1)[14]. The JASCO V-730 is equipped with a Silicon photodiode, which generates a current based on the irradiation of the n-p junction of a semiconductor [16]. The Varian Cary 100 uses a photomultiplier tube, which is capable of generating a current when the incident light is low. This is possible due to the amplification of the incident light [17].

**Table 1.** Photometer comparison regarding data intervals and scan speeds used with the two different photometers.

Data interval [nm]	Scan speed [nm/min] (V-730 )	Scan speed [nm/min] (Cary 100)
0.5	100	100
0.5	1000	1000
1	100	100
1	1000	1000
5	100	-
5	1000	n.a

## 2.2 Comparing different color measurements

### 2.2.1 Comparing CIE $L^*a^*b^*/L^*C^*h^0$ and Glories

For the comparison of CIE  $L^*a^*b^*/L^*C^*h^0$  and Glories color measurement 925 red wines were examined. Following the previous published classification, the data set was furthermore divided into dark red wines with an  $L^* \leq 20$  and light red wines with  $L^* > 20$ . The dataset of the dark red wines consisted of 617 wines and the light red wines consisted of 306 wines. For white wine the sample set consisted of 56 white. The sample set is smaller, due to the lower variance of white wine color. The red wine samples were obtained during the Mundus Vini summer tasting 2022 and the color was measured immediately after opening the bottle. For the measurements a StellarRad™ (StellarNet Inc. FA, USA) handheld colorimeter in combination with a dip probe was used to record transmission spectra. The dip probe was equipped with a tip containing a 2 mm pathlength. The transmission spectra were recorded between 200 and 1100 nm. The transmission spectra were referred to a 10 mm pathlength and the CIE  $L^*a^*b^*/L^*C^*h^0$  coordinates were calculated between 360 and 830 nm. Also, the color vector and the hue angle were calculated. The primary Glories parameters were extracted from the transmission spectra. After calculation of the secondary Glories parameters a Spearman correlation was performed between the Glories and CIE  $L^*a^*b^*/L^*C^*h^0$  parameters. The White wines consisted of 7 wines per grape variety. The five grape varieties were used two of whom were evaluated in the two wine styles dry and sweet. The wines originated from the Germany, France, Italy, Spain, United States of America, New-Zealand, and Australia as the red wines except for Austria and the vintages ranged from 2013 to 2020 were measured as described in the previous section.

### 2.2.2 Comparing the human perception with analytical methods

For the comparison of the human perception and the photometric methods a total of 112 wines were used, 56 of whom were red wines and the 56 white wines described in section 2.2.1. Seven red wines each from eight different grape varieties were used and the wines originated from Germany, France, Italy, Spain, United States of America, New Zealand, Australia, and Austria. The vintages ranged from 2012 to 2020. The wines were examined in a specialized sensory box that reduce as much interfering stimuli as possible. The evaluation was performed with the help of two linear line scales. The first linear line scale evaluated the darkness, respectively the lightness of a wine. The second linear line scale evaluated the perceived hue and ranged from violet to red to orange in red wine and from green to yellow to orange in white wine. The CIE  $L^*a^*b^*/L^*C^*h^0$  and Glories parameters were calculated as stated in literature. The sensory panel consisted of 23 trained individuals, 11 of whom were male and 12 of whom were female. The panel was chosen demographically between 18 and 60 years. The panel had to proof their color vision and were trained with different hues of wine color[15].

### 2.2.3 Re-evaluating the just noticeable difference

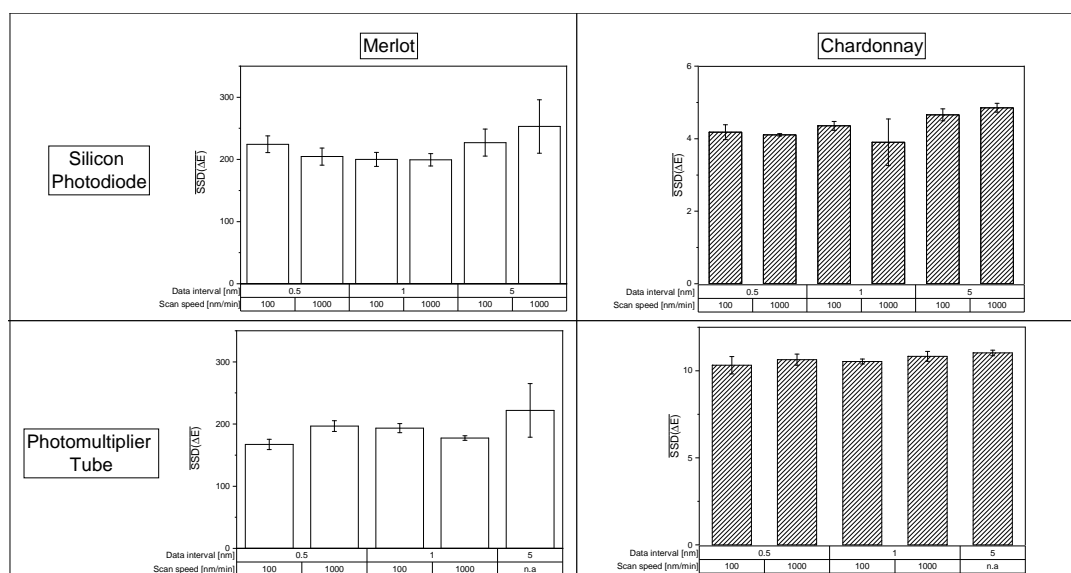
To evaluate the JND the triangle test according to DIN EN ISO 4210:2021 was performed. For this, different blends of a Cabernet Sauvignon wine and one Riesling wines were created to manipulate the hue in the dark red, light red, and light yellow color area. Undiluted Cabernet Sauvignon was used to as base wine for the dark red wine, an 8-fold dilution of Cabernet Sauvignon and Riesling was used to create the base wine of the light red color area. For the light yellow color area the undiluted Riesling was used as base wine. The same panel as described in section 2.2.2 was used [15].

## 2.3 Statistical Analysis

For the evaluation of the impact of photometer settings on the calculation of the CIE  $L^*a^*b^*/L^*C^*h^0$  coordinates normality testing, as well as ANOVA, and Tukey Honest significant difference post hoc test, as well as the linear and non linear regressions were performed in Origin(Pro) (2020b, OriginLab, USA). XLSTAT (2020, ADDINSOFT, France) was used to calculate Bartlett's test to evaluate homoscedasticity. Correlation between Glories and CIE  $L^*a^*b^*/L^*C^*h^0$  were calculated and visualized in Python using the pandas [18], NumPy [19], matplotlib [20], and seaborn [21] libraries.

## 3 Results

In Figure 1 the impact of the data interval and scan speed is depicted for two different photometers. The bars are the averaged sum of squares deviation (SSD) and therefore show the distinguishability between the wines of the photometer settings. The error bars are the standard deviations and show the reproducibility of the measurement for each photometer setting [14]. In general, the ability to distinguish between the wines was not affected significantly. However, the reproducibility was affected by the data interval and scan speed. A higher data interval of 5 nm resulted in the lowest reproducibility. In Figure 1 it was possible to reproduce most of these findings disregardless of the detector system. A lower data interval could minimize the error and yields a higher reproducibility. Furthermore, the results show that the data interval should not be too small. The reproducibility of the photometer with the Silicon photodiode decreased when a data interval of 0.5 nm was used. This is not the case for the photometer using a photomultiplier tube. The photomultiplier tube amplifies the signal while this is not the case for the Silicon photodiode. Therefore, it is possible that the noise of the detector has different impacts on the spectra measurement. A data interval of 1 nm in combination with 1000 nm/min yielded the highest reproducibility in both photometers.

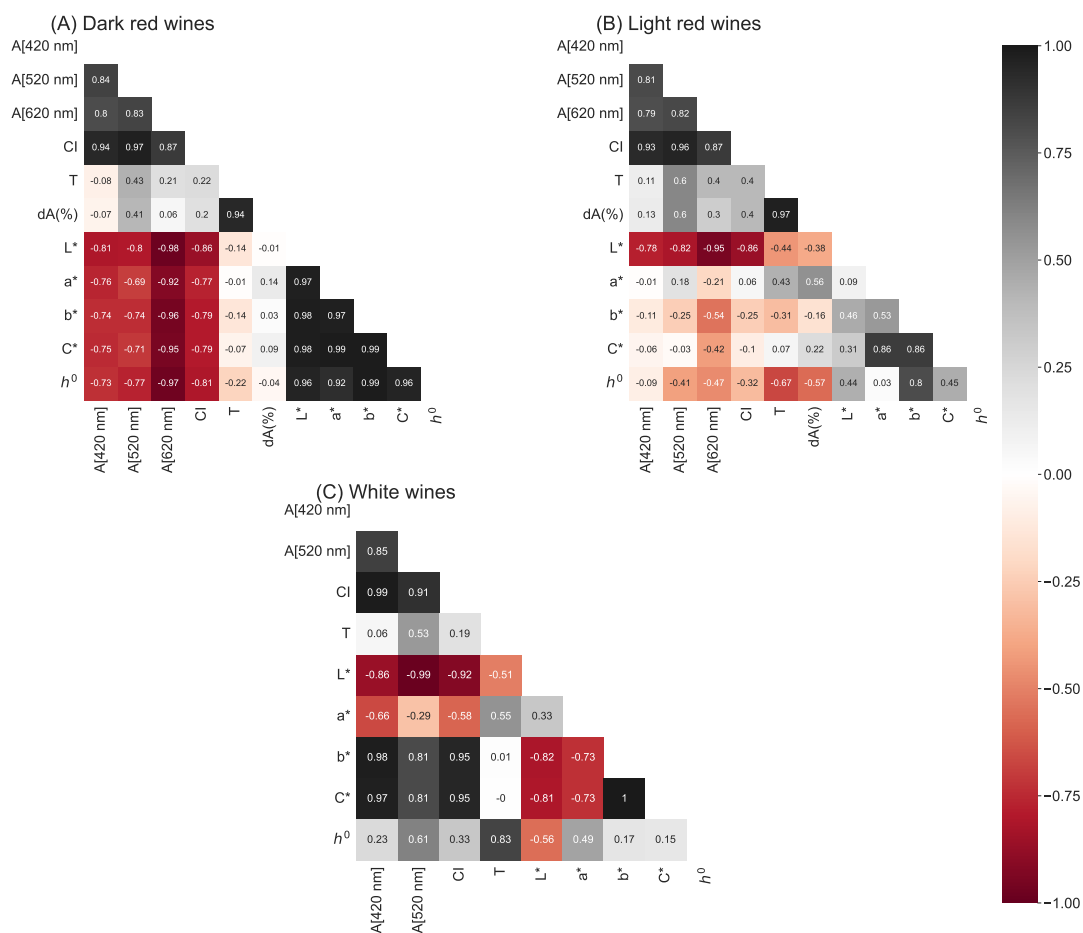


**Figure 1.** Influence of the photometer settings data interval and scan speed on the CIE  $L^*a^*b^*/L^*C^*h^0$  coordinates on using a Silicon photodiode and a photomultiplier tube as detector systems. The mean sum of squared deviations are shown with error bars (SD;  $n = 3$ ;  $\alpha = 0.05$ ).



### 3.1 Correlation between CIE $L^*a^*b^*/L^*C^*h^0$ and Glories

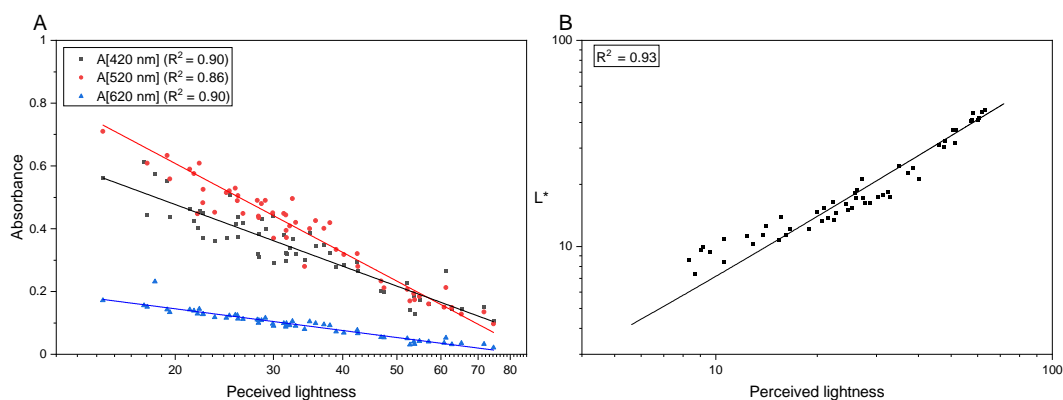
The correlation of 617 dark red wines and 306 light red wines as well as 56 white wines are displayed in Figure 2. In Figure 2A the examination of the dark red wines revealed a strong correlation between the optical density of the Glories parameters and the CIE color space. This is not the case for the secondary Glories parameters T and  $dA(\%)$ . Here, the correlation with the optical density of the primary Glories parameters and the CIE color space is very weak. However, the correlation between the  $L^*$ ,  $a^*$ , and  $b^*$  coordinates as well as the correlation between the absorbance values of absorbance values of the Glories parameters is strong. In Figure 2B the same correlation is depicted for the light red wines. In contrast to the correlation of the dark red wines, the correlation between the Glories color measurement and the CIE color space is in general weak. The only exception of this is the  $L^*$  coordinates that shows a strong correlation with CI of the Glories parameters. These results extend the findings of previous studies that the CIE color space and Glories' color measurement are not interchangeably usable. Prior studies focused hereby on the absorbance values of the Glories method. It is now evident that the other parameters also do not correlate with the CIE  $L^*a^*b^*/L^*C^*h^0$  color space. To evaluate which of the methods should be used, the alignment with the human perception should be evaluated.



**Figure 2.** Spearman correlation for dark red wines (n=617), light red wines (n=308), and white wines (n= 56).

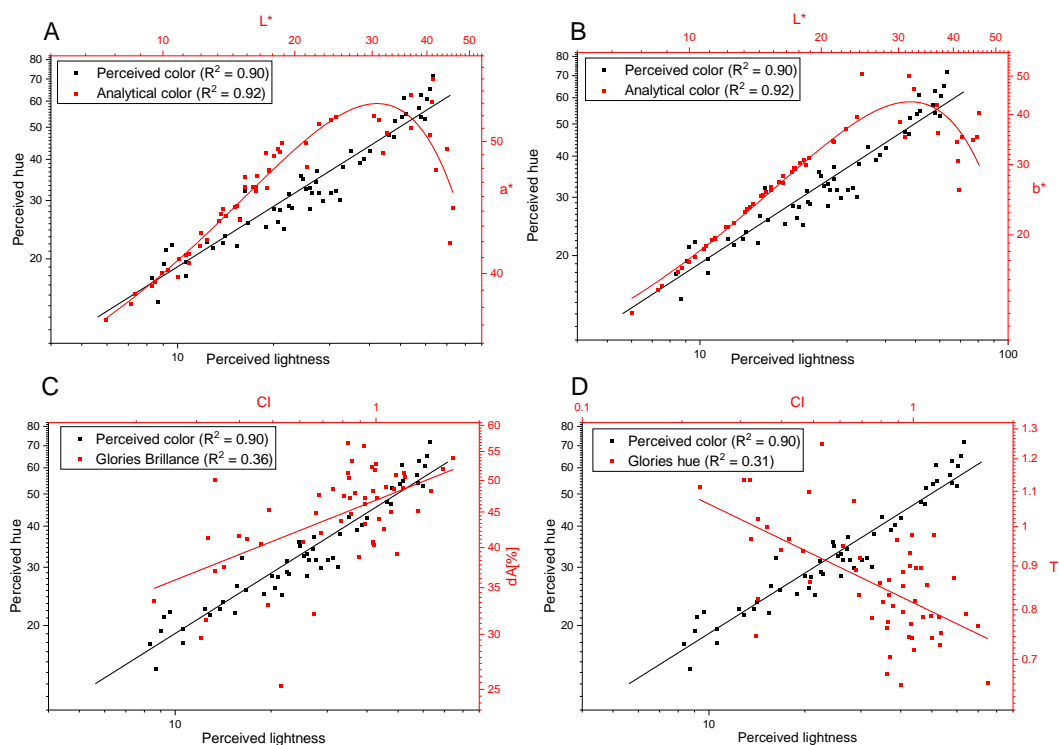
### 3.2 Comparing the human perception with the analytical methods

To evaluate which of the analytical methods depicts the human perception better, the red and white wines were evaluated by a panel and compared to the analytical methods of Glories' color measurement and CIE color spaces. In Figure 3A the correlation between the primary absorbance values of Glories color measurement correlate very good with the perceived lightness. Additionally, in Figure 3B the correlation of the  $L^*$  coordinate with the perceived lightness yielded similar results.



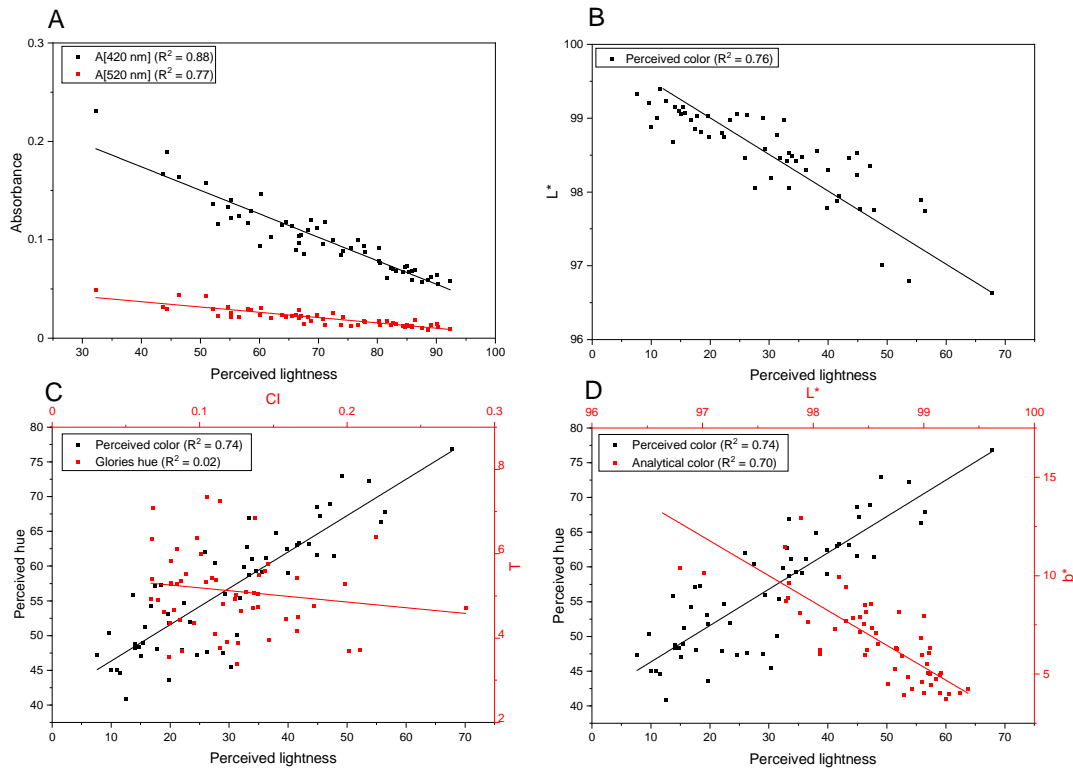
**Figure 3.** Comparison of Glories absorbance values (A) and the  $L^*$  coordinate of the CIE color space (B) regarding the perceived lightness in red wine.

Figure 4A describes the correlation between the perceived color and the perceived hue compared to the correlation of the  $L^*$  coordinate against the  $a^*$  coordinate. Figure 4B shows the perceived color against the correlation between the  $L^*$  coordinate and the  $b^*$  coordinate. The correlation between the perceived color and the color obtained from the CIE  $L^*a^*b^*/L^*C^*h^0$  coordinates deviates from each other as the lightness increases. Regarding the perceived color the relationship between the perceived lightness and the perceived hue is linear, whereas the relationship between the  $L^*$  coordinate and the  $a^*$ , and  $b^*$  coordinate respectively is non-linear. The coefficient of determination in both cases is  $R^2 \geq 0.9$ . However, this is not the case for the comparison between the perceived color and the Glories color measurement. Glories' secondary parameters  $dA(\%)$  and  $T$  are not suited to depict the human perception regarding red wine.



**Figure 4.** Regressions of perceived hue and perceived lightness compared to the regressions of the  $L^*a^*$  projection (A),  $L^*b^*$  projection (B), Glories' CI and T (C), as well as Glories CI and  $dA(\%)$ (D).

In Figure 5A the relationship between the perceived lightness and the absorbance values of the Glories measurement system is shown. The same relationship between the perceived lightness is shown in Figure 5B. Due to the low absorbance values at 620 nm only the absorbance values at 420 nm and 520 nm were used. Regarding white wines the findings were similar to the findings in red wine. Both, the Glories color measurement and  $L^*$  coordinate have a similar capacity to depict the human perception. In Figure 5C the perceived color is compared with Glories' hue and according to the coefficient of determination Glories hue parameter is also not suited to depict the perceived color of white wine in coherence with the findings in red wine. In Figure 5D the comparison between the perceived color and the CIE  $L^*a^*b^*/L^*C^*h^0$  color space is shown revealing a better suitability to depict the human perception. However, the regressions of white wine are not as good as the regressions for red wine.



**Figure 5.** Regressions of Glories absorbance values and perceived lightness (A), as well as the Regression of the  $L^*$  coordinate and the perceived lightness (B). Also, the regressions of the perceived hue and perceived lightness compared to the regression of CI and T (C), and the  $L^*b^*$  projection (D).

### 3.3 Re-evaluating the just noticeable difference

For the JND the color difference was calculated with the Euclidean color difference ( $\Delta E_{ab}$ ) and the color distance according to CIEDE2000 ( $\Delta E_{00}$ ). The color differences are shown in table 2. It is visible that the prior used  $\Delta E_{ab} = 3$  established by Martinez *et al.* could be replicated [22]. However, the visible color threshold changes throughout the color areas, yielding a higher visible color threshold in the light red color area and a lower visible color threshold in the light yellow color area. Melgosa *et al.* established the Standardized Residual Sum of Squares (STRESS) to measure the uniformity of color spaces [5, 6, 23]. According to these findings and the findings of Luo *et al.*, where the results showed tolerance ellipses in the CIE color space, the Euclidean color distance yields lower equidistance than the modernized CIEDE2000 color distance formula. However, this color distance formula is also not completely uniform, which explains why the visible color threshold is not the same disregarding of the color area. However, the range between the lowest visible color threshold and the highest visible color threshold decreases 2.7-fold, resulting in a higher degree of uniformity. Therefore, the CIEDE2000 color distance should be used instead of the Euclidean color difference with the evaluated color thresholds as a reference point.

**Table 2.** The JND expressed as the Euclidean color difference and the CIEDE2000 color distance [15].

Color area	$\Delta E_{ab}$	$\Delta E_{00}$
Dark red	3.1	1.4
Light red	8.1	3.4
Light yellow	0.60	0.64
$\Delta E_{Max} - \Delta E_{Min}$	<b>7.5</b>	<b>2.8</b>

## 4 Conclusion

In the presented work new insights on how color measurement is affected by photometers, how different photometric color measurements are correlated, and how this is in coherence with the results of the sensory evaluation were gained. For the CIE color space the currently recommended data interval of 5 nm yielded the lowest reproducibility. A higher reproducibility was achieved by using a lower data interval of 1 nm. Furthermore, the presented work validates our understanding about the correlation between Glories' color measurement and the CIE color space. And since this section indicates that Glories and CIE  $L^*a^*b^*/L^*C^*h^0$  cannot be used interchangeably the comparison showed which method better aligns with the perception of the human eye the CIE color space is recommended before Glories color system. Additionally, due to better understandings about the uniformity or lack thereof it was possible to redefine new JND's that differs in dependence of the color area. The JND expressed as the updated CIEDE2000 color distance formula is advised to use.

## References

- [1] Y. Glories, *OENO One* **18**, 195 (2016)
- [2] CIE, *Cie s 017/e:2020*
- [3] *Compendium of international methods of wine and must analysis: Volume 1*, edition 2021 edn. (OIV, Paris, 2021), ISBN 978-2-85038-034-1
- [4] DIN e.V., *Colorimetry - part 6: Ciede2000 colour-difference formula* (2016)
- [5] M. Melgosa, R. Huertas, R.S. Berns, *Journal of the Optical Society of America. A, Optics, image science, and vision* **25**, 1828 (2008)
- [6] M.R. Luo, G. Cui, B. Rigg, *Color Research & Application* **26**, 340 (2001)
- [7] G. Morrot, F. Brochet, D. Dubourdieu, *Brain and language* **79**, 309 (2001)
- [8] Thi H. Nguyen, Dominik Durner, *Food Quality and Preference* **107**, 104844 (2023)
- [9] G.P. Parpinello, A. Versari, F. Chinnici, S. Galassi, *Food Research International* **42**, 1389 (2009)
- [10] C. Coulon-Leroy, N. Pouzalgues, L. Cayla, R. Symoneaux, G. Masson, *OENO One* **52** (2018)
- [11] Bundesrat, *Änderungen und entschließung zur vierundzwanzigsten verordnung zur änderung der weinverordnung* (26.03.2021)
- [12] P. Ribéreau-Gayon, *Handbook of enology: Volume 2: The chemistry of wine stabilization and treatments*, 2nd edn. (John Wiley, Chichester, West Sussex, England, 2006), ISBN 978-0-470-01038-9, <https://ebookcentral.proquest.com/lib/kxp/detail.action?docID=257673>
- [13] CIE, *Cie 015:2018 colorimetry, 4th edition*
- [14] M. Hensel, S. Di Nonno, Y. Mayer, M. Scheiermann, J. Fahrer, D. Durner, R. Ulber, *Processes* **10** (2022)

- [15] M. Hensel, Marcel Scheiermann, J. Fahrer, D. Durner (2023)
- [16] *Silicon photodiode*, [https://www.hamamatsu.com/content/dam/hamamatsu-photonics/sites/documents/99\\_SALES\\_LIBRARY/ssd/si\\_pd\\_kspd9001e.pdf](https://www.hamamatsu.com/content/dam/hamamatsu-photonics/sites/documents/99_SALES_LIBRARY/ssd/si_pd_kspd9001e.pdf)
- [17] *Photomultiplier tube* (09.02.2020), [https://www.hamamatsu.com/content/dam/hamamatsu-photonics/sites/documents/99\\_SALES\\_LIBRARY/etd/PMT\\_handbook\\_v4E.pdf](https://www.hamamatsu.com/content/dam/hamamatsu-photonics/sites/documents/99_SALES_LIBRARY/etd/PMT_handbook_v4E.pdf)
- [18] Jeff Reback, Wes McKinney, jbrockmendel, Joris Van den Bossche, Tom Augspurger, Phillip Cloud, gfyong, Sinhrks, Adam Klein, Matthew Roeschke et al., *pandas-dev/pandas: Pandas 1.0.3* (2020)
- [19] C.R. Harris, K.J. Millman, S.J. van der Walt, R. Gommers, P. Virtanen, D. Cournapeau, E. Wieser, J. Taylor, S. Berg, N.J. Smith et al., *Nature* **585**, 357 (2020)
- [20] Thomas A Caswell, Michael Droettboom, Antony Lee, Elliott Sales de Andrade, Tim Hoffmann, John Hunter, Jody Klymak, Eric Firing, David Stansby, Nelle Varoquaux et al., *matplotlib/matplotlib: Rel: v3.5.1* (2021)
- [21] M. Waskom, *Journal of Open Source Software* **6**, 3021 (2021)
- [22] J.A. Martínez, M. Melgosa, M.M. Pérez, E. Hita, A.I. Negueruela, *Food Science and Technology International* **7**, 439 (2001)
- [23] P.A. García, R. Huertas, M. Melgosa, G. Cui, *Journal of the Optical Society of America. A, Optics, image science, and vision* **24**, 1823 (2007)

#### **5.4 Evaluation of different machine learning algorithms to build an application classifying Blanc de noir wine with spectrophotometric data**

The publication presented in section 5.1 and 5.2 lay the foundation for the following publication. The previous publications showed that the CIE  $L^*a^*b^*$  color space is better suited for color measurement in wine than Glories' color measurement, since it more closely resembles the human perception of wine color. The main objective of this work is to find new applications regarding wine color apart from the purely descriptive nature of the parameter. In Germany, the color of Blanc de noir is regulated by state law. In order to obtain a permit to sell a potential Blanc de noir, the wine has to be assessed by wine controlling entities in a subjective sensory evaluation. In order to help the winemaker and objectify this process, the ability to classify Blanc de noir wines from other wine styles like white wine and rosé wine, the CIE  $L^*a^*b^*$  color space was combined with different ML algorithms of varying complexity. Furthermore, to measure the applicability of CIE  $L^*a^*b^*$  coordinates with ML algorithms, the raw transmission spectra were also used to train the algorithm. CIE  $L^*a^*b^*$  coordinates were used to train Logistic regression and an SVM, while the transmission spectra were used to train an SVM and XGBoost. The models were optimized via hyperparameter optimization and validated by iterating the optimization process over 100 different training and test datasplits.

The results show that SVM on CIE  $L^*a^*b^*$  coordinates provided the optimal classification model for this task. Logistic regression on CIE  $L^*a^*b^*$  coordinates show a high amount of overfitting, as the optimized models performance differs between training and test data. SVM on transmission spectra also showed some overfitting. While XGBoost on transmission spectra performs good, the validation process shows, that it should not be considered optimal, as it produces unstable results, when different training datasets are used. To make SVM on CIE  $L^*a^*b^*$  accessible a browser-based web application was developed that enables winemakers and wine laboratories to evaluate their potential Blanc de noir wine with the algorithm.

This study proves that CIE  $L^*a^*b^*$  coordinates can be used classification purposes. Future studies should extend this work to other wine styles like orange wine, red wine, or local specialties.

## **Evaluation of different machine learning algorithms to build an application classifying Blanc de noir wine with spectrophotometric data**

Authors:

Marcel Hensel<sup>1</sup>, Jochen Vestner<sup>1</sup>, Jörg Fahrer<sup>3</sup>, Dominik Durner<sup>1</sup>, Affiliations:

<sup>1</sup> Weincampus Neustadt, Dienstleistungszentrum ländlicher Raum Rheinpfalz, 67435 Neustadt an der Weinstraße, Germany

<sup>2</sup> Division of Food Chemistry and Toxicology, Department of Chemistry, University of Kaiserslautern-Landau, 67663 Kaiserslautern, Germany

### **Disclaimer:**

**The presented work was exclusively performed by the DLR Rheinpfalz (Marcel Hensel). Measurements of the transmission spectra for the white wines were conducted by Marina Scheiermann.**

Copyright

Reprinted with permission: Hensel, Marcel; Vestner, Jochen; Fahrer, Jörg; Durner, Dominik (2024): Evaluation of different machine learning algorithms to build an application classifying Blanc de noir wine with spectrophotometric data American Journal of Enology and Viticulture. Submitted manuscript. Copyright ©2024 by the American Society of Enology and Viticulture. All rights reserved.



1 **Article Type**

2 Research Report

3

4 **Article Title**

5 Evaluation of Machine Learning Algorithms to classify Blanc De Noir wines with Spectrophotometric

6 Data

7

8 **Author names**

9 Marcel Hensel<sup>1</sup>, Jochen Vestner<sup>1</sup>, Jörg Fahrer<sup>2</sup>, Dominik Durner<sup>1\*</sup>

10

11 **Author affiliations**

12 <sup>1</sup>Dienstleistungszentrum Ländlicher Raum Rheinpfalz, Weincampus Neustadt, Breitenweg 71, 67435

13 Neustadt an der Weinstraße, Germany

14 <sup>2</sup>Division of Food Chemistry, Department of Chemistry, Rheinland-Pfälzische Universität Kaiserslautern-

15 Landau, Erwin-Schrödinger-Straße 52, 67663 Kaiserslautern, Germany

16

17 **\*Corresponding author** (Dominik Durner: dominik.durner@hs-kl.de)

18

19 **ORCID numbers**

20 Marcel Hensel: 0000-0003-4360-5328

21 Jörg Fahrer: 0000-0001-5225-2718

22 Dominik Durner: 0009-0009-0525-4548

23

24 **Acknowledgments**

25 The authors thank the working group of Maren Scharfenberger-Schmeer as well as Roland Ulber and  
26 Sarah Di Nonno from RPTU Kaiserslautern-Landau for their input and support as project partners.  
27 Furthermore, the authors thank Thi Nguyen for their assistance with the English language and his ideas  
28 during the finalization of this manuscript. Lastly, the authors thank Marc Weber for his support with  
29 programming and the initial discussions leading to this manuscript. This work was financially supported  
30 by the Arbeitsgemeinschaft Industrieller Forschungsvereinigungen (AiF) “Otto von Guericke” e.V.  
31 (project number: AIF 20964 N).

32

33 **Short version of title:** Machine Learning in Enology

34

35 **Abstract:**

36 **Background and goals**

37 Machine learning (ML) and statistical modeling have emerged as important innovations in science. In  
38 wine research, ML is often used to predict abstract parameters such as wine quality based on complex  
39 instrumental chemical analysis.

40 **Methods and key findings**

41 The present study uses spectrophotometric data and CIE L\*a\*b\* coordinates from 176 commercial wines  
42 to distinguish Blanc de noir from rosé wine, and white wine. The transmission spectra were used to train  
43 extreme gradient-boosted trees and a support vector machine (SVM). CIE L\*a\*b\* coordinates were used  
44 to train SVM and logistic regression. After parameter hyperparameter tuning, the combination of SVM on  
45 CIE L\*a\*b\* data provided the optimal classification with a cross-validated accuracy of 0.88 and an F1  
46 score of 0.93.

47

48 **Conclusions and significance**

49 This study compares the applicability of different machine learning algorithms, namely logistic regression,  
50 support vector machine (SVM), and eXtreme gradient boosting (XGBoost), to distinguish Blanc de noir  
51 from white and rosé wine based on spectrophotometric data. Dimensional reduction from transmission  
52 data to CIE L\*a\*b\* coordinates was utilized as a method to train machine learning models. This  
53 transformation, simplifying the use of machine learning algorithms for potential users, was compared to  
54 raw/preprocessed transmission data as input to train machine learning models. The final classification  
55 model is deployed in a browser-based user-friendly dashboard for winemakers and other users, such as  
56 wine laboratories.

57 **Keywords:** Machine Learning, Blanc de Noir, Transmission Data, CIE L\*a\*b\*, Wine color, Support  
58 Vector Machine

59

## 60 **Introduction**

61 Machine learning (ML) and statistical modeling are among the most important recent innovations in  
62 scientific research. The goal of statistical modeling is to draw conclusions about a population based on  
63 the observed samples, though it uses the traditional approach of making assumptions about the underlying  
64 dataset using mathematical formulation. On the other hand, ML focuses on developing models that can  
65 learn patterns from a given dataset without being explicitly programmed to do so, seeking generalization  
66 rather than inference. In ML, algorithms including statistical models are designed to make the most  
67 accurate qualitative or quantitative predictions. Qualitative predictions include classification, where ML  
68 models use a given dataset to predict the presence or absence of a feature. Quantitative predictions rely on  
69 regression, where the cardinality of the target feature is predicted based on the given dataset. These  
70 algorithms learn from structured or unstructured input data and apply the results from this learning process  
71 to new data. ML is already an established tool in many different domains of natural sciences. For example,  
72 Hsieh et al. (2011) introduced an ML model based on random survival forests, a tree-based method  
73 derived from the random forest approach (Breiman 2001; Ishwaran et al. 2008), which identified risk  
74 factors for the survival of patients with systolic heart failure (Hsieh 2011). Another study uses ML as a  
75 tool to improve breast cancer diagnostics (Andrew et al. 2011). Medicinal ML models also use  
76 spectroscopic data such as Fourier-Transformation Infrared Spectroscopy, which predicts sarcomas and  
77 glial abnormal growth (Ramirez 2021). In wine research, the impact of ML algorithms is growing. The  
78 main focus of current literature is to predict abstract parameters such as wine quality based on complex  
79 instrumental analytical procedures. (Kumar et al. 2020; Bhardwaj et al. 2022; Shaw et al. 2020). Kumar  
80 et al. (2020) used a red wine dataset of the University College London (UCL) ML repository, which  
81 contains analytical parameters like acidity, sugar, free and total sulfur dioxide, and many more, to predict  
82 the perceived wine quality using different algorithms, like Naive Bayes, Random Forest, and Support

83 Vector Machine (Kumar et al. 2020). Bhardwaj et al. (2022) used a total of 47 volatile components (mainly  
84 alcohols and esters) and other parameters like pH, alcohol, anthocyanin level etc. to predict the quality of  
85 New Zealand Pinot noir using a mixture of real and synthetic data (Bhardwaj et al. 2022). Another study  
86 conducted by Shaw et al. also used the UCL ML repository red wine dataset to predict the wine quality  
87 with support vector machine, random forest, and multilayer perceptrons (Shaw et al. 2020). The cited  
88 publications predict quality parameters that are important when the wine is already produced. However,  
89 no publications could be found that used ML during wine production and with feasible parameters like  
90 spectroscopic data as input as it is established in other fields (Ramirez 2021).

91 Wine color is an important analytical quality parameter. Color is known to influence the olfactory and  
92 gustatory perception of wine (Nguyen und Durner 2023; Morrot et al. 2001). Wine color is also an  
93 important purchase driver as well as a typicality parameter for consumers (Issa-Issa et al. 2021; Coulon-  
94 Leroy et al. 2018; Parpinello et al. 2009). Additionally, the color of some wines such as Blanc de noir is  
95 regulated by law in some countries (Bundesrat 2021). Blanc de noir is a very light, unsaturated colored  
96 wine produced from red grapes processed in the same way as white grapes (Bundesrat 2021). Since Blanc  
97 de noir wines are made from red grapes, the very light color of Blanc de noir wines directly influences  
98 consumer acceptance (Ellis und Kok 2017). It is sometimes very difficult to distinguish between Blanc de  
99 noir and rosé as well as between Blanc de noir and white wine by the human eye. Commercial Blanc de  
100 noirs range from rosé-like to white wine-like color, making it challenging to establish a universal color  
101 threshold to distinguish the color of Blanc de noir from rosé and white wines.

102 For color determination, high-dimensional data from transmission spectra can be reduced to only three  
103 dimensions via the  $L^*a^*b^*$  color space of the Commission Internationale de l'Eclairage (CIE). The CIE  
104  $L^*a^*b^*$  color space is a Cartesian coordinate system, where the  $L^*$  value represents the lightness of an  
105 object, the  $a^*$  value represents either the red component ( $+a^*$ ) or the green component ( $-a^*$ ) of a color,

106 whereas the  $b^*$  value represents either the yellow component ( $+b^*$ ) or the blue component ( $-b^*$ ) of an  
107 object. The CIE  $L^*a^*b^*$  color space is recommended by the International Organization for Vine and Wine  
108 (OIV) for the determination of wine color (OIV 2021). The CIE  $L^*a^*b^*$  color space is also an effective  
109 tool to depict the visually perceived color of wine (Hensel et al. 2023). Besides the transmission spectrum,  
110 a color-matching function of the standard observer (DIN EN ISO/CIE 11664-1) and the power distribution  
111 of a standard illuminant (DIN EN ISO/CIE 11664-2) are needed for the CIE  $L^*a^*b^*$  coordinates  
112 calculation (DIN EN ISO/CIE 11664-4). The color-matching functions of the standard observer were  
113 developed experimentally. The spectral sensitivity of the eye was evaluated with the three primary  
114 valences red, green, and blue (Wright 1929, 1930; J. Guild 1931; Stiles und Burch 1959). The currently  
115 used standard is the D65 illuminant representing natural daylight on a sunny day and the  $10^\circ$  standard  
116 observer (DIN EN ISO/CIE 11664-2). Dimensionality reduction by transforming transmission spectra into  
117 CIE  $L^*a^*b^*$  coordinates has the advantage of enhanced comprehensibility for users. This approach allows  
118 a more intuitive representation of the color data aligned with human perception. However, most ML  
119 algorithms do not produce reliable results from datasets with only three dimensions. In contrast, a  
120 transmission spectra provides an abundant number of dimensions, allowing the use of most ML algorithms  
121 (Chopra et al. 2021; Hossain et al. 2020; Irene Marivel Nolasco Perez et al. 2018).

122 One of the most basic models for binary classification is logistic regression, which is a special form of  
123 regression analysis. In comparison to linear regression, nominal-scaled dependent variables are used  
124 instead of interval-scaled dependent variables. Logistic regression calculates the probability of an event  
125 from the natural logarithmic ratio between its occurrence and non-occurrence (log-odds). The function  
126 that converts log-odds to probability is the logistic function. More complex ML algorithms for binary  
127 classification are Support Vector Machines (SVM). SVM is well suited for datasets with a limited number  
128 of variables, such as the well-known Iris-dataset (Fisher RA 1936; Iris - UCI Machine Learning Repository

129 2024). Binary classification is achieved by locating support vectors in a higher dimensional feature space  
130 and constructing a suitable hyperplane based on the location of the support vectors. The simplest form of  
131 the SVM is the maximal margin classifier or hard margin SVM (Boser et al. 1992). The margin in SVM  
132 signifies the smallest distance between the hyperplane separating the two classes and the nearest sample  
133 point to the hyperplane. A hard margin SVM is only applicable to data that is linearly separable and  
134 therefore rarely used. The better option for “real-world” data often with outliers, overlapping classes etc.,  
135 is the soft margin SVM, which allows a certain degree of misclassification. The occurrence of overfitting  
136 is thus reduced and a better generalization is possible (Bennett und Campbell 2000).

137 If transmission spectra in the visual wavelength area are used, more complex algorithms like eXtreme  
138 Gradient Boosting (XGBoost) can be applied. XGBoost is an ensemble learning method using a unique  
139 way to build decision trees sequentially, with every new tree designed to correct the errors made by the  
140 previous tree. For this purpose, XGBoost uses L1 (Lasso regression) and L2 (Ridge regression)  
141 regularization to prevent overfitting. Furthermore, a common problem with “real-world” data is missing  
142 data, for which XGBoost has an intrinsic mechanism to cope (Chen und Guestrin 2016).

143 The goal of this study was to develop an ML application to classify commercial Blanc de noir wines based  
144 on transmission spectra and CIE  $L^*a^*b^*$  coordinates, thus classifying the color appearance of Blanc de  
145 noirs compared to rosé and white wines. For this purpose, the dimensional reduction of transmission  
146 spectra to CIE  $L^*a^*b^*$  coordinates in combination with logistic regression and SVM was compared to the  
147 more complex combination of transmission spectra on SVM and XGBoost modeling. To make the  
148 classification model accessible to wine chemists and winemakers, a browser application was developed.  
149 This includes a dashboard that compares the spectrophotometric data of unknown wine samples to a given  
150 dataset of commercial wines. Spectrophotometric data of unknown wines can easily be obtained from any  
151 photometer or from low-cost photometric devices (Di Nonno und Ulber 2021, 2022; Hensel et al. 2022).

152 **Materials and Methods**

153 Samples

154 The sample set contained 176 commercial wines from Germany, France, Italy, Australia, Spain, Belgium,  
155 and Turkey. The vintages of the wines were 2020 to 2022. The data set consisted of 82 Blanc de noir  
156 wines, 56 white wines, and 38 rosé wines. The grape varieties in the “Blanc de noir” class were Pinot noir,  
157 Cabernet Sauvignon, Pinot Meunier, Dolcetto, Portugieser, Pinot Madeleine, and St. Laurent. White wines  
158 were from Riesling, Gewürztraminer, Sauvignon blanc, Pinot gris, Pinot blanc, and Chardonnay. Rosé  
159 wines were from Pinot noir, Cabernet Sauvignon, Cabernet Franc, Merlot, Sangiovese, Blauer Zweigelt,  
160 Lemberger, Grenache, Cabernet Cortis, Bobal, Muskattrollinger, Montepulciano, and Syrah. The samples  
161 were stored at 8 °C in the 0.75 L bottles as provided from the producers. One bottle per wine was used for  
162 the study. The bottle type, closure material, bottle age, experiment date, as well as wine compositional  
163 data is given in the supporting information.

164 Basic wine analysis

165 Ethanol, residual sugar, titratable total acidity, and pH were determined by Fourier transform infrared  
166 (FTIR) spectroscopy (Foss WineScan FT120 Basic, Hillerød, Denmark). An in-house calibration “white  
167 wine” was used for the white wines and Blanc de noir wines. For rosé wines, an in-house calibration “rosé  
168 wine” was used. Free and total SO<sub>2</sub> were determined with iodometric titration by Ripper method using the  
169 Titrator T50 (Mettler Toledo, Gießen, Germany).

170 Photometric measurement and CIE L\*a\*b\* computation

171 To obtain the color data, transmission spectra of the wines were recorded in triplicate between 360 and  
172 780 nm with a scan speed of 1000 nm/min and a data interval of 1 nm. These parameter settings provided  
173 the most reproducible results (Hensel et al. 2022). The Blanc de noir and white wines were measured in  
174 10 mm cuvettes (Brand, Germany), whereas the rosé wines were measured in a 1 mm flow cuvette



175 (Hellma, Germany). Distilled water was used as a blank measurement. In compliance with the OIV  
176 recommendation, the transmission spectra of the rosé wines were referred to a 10 mm path length (OIV  
177 2021). CIE L\*a\*b\* coordinates were computed from averaged transmission spectra using the 10° standard  
178 observer and the D65 standard illuminant according to literature (DIN EN ISO/CIE 11664-1; DIN EN  
179 ISO/CIE 11664-2).

#### 180 Data pipeline

181 Figure 1 depicts the different computation steps starting from the transmission spectra of all wines. The  
182 transmission spectra of all wines were transformed to CIE L\*a\*b\* coordinates and preprocessed via the  
183 RobustScaler, where the median and the interquartile range is used to scale the data. Logistic regression  
184 was applied only to preprocessed CIE L\*a\*b\* coordinates. SVM was applied to preprocessed CIE L\*a\*b\*  
185 coordinates and the preprocessed transmission spectra. XGBoost was applied only to full transmission  
186 spectra without preprocessing.

187 To find the optimal parameters for each model, a full factorial hyper parameter tuning (grid search) on  
188 70 % of the dataset (training data) with a 5-fold cross validation was used. Application of the optimized  
189 models on the test data facilitated the validation of the models. The accuracy of the five subsets was  
190 averaged and is only referred to as accuracy from this point onwards. Validation was repeated 100 times  
191 including shuffled splitting of training and test data in each repetition (different random seed for each  
192 repetition). This repeated validation allowed for a conservative estimation of the classification variance.  
193 Based on the test dataset, confusion matrices were calculated to compare actual and predicted  
194 classifications. The accuracy parameter considers only the true positive (TP) and true negative (TN)  
195 counts, thus accuracy may be inflated by a high number of TN predictions. Therefore, to validate the  
196 accuracy as a parameter the F1-score that considers also the false positive (FP) and false negative (FN)  
197 counts is calculated with the following equation.

198 
$$F1 = \frac{TP}{TP + \frac{1}{2}(FP + FN)}$$

199 With:

200 TP: True positive count

201 FP: False positive count

202 FN: False negative count

203 Logistic Regression

204 Logistic regression was used as a reference model due to its simplicity. The logistic regression was trained  
205 and tested with the preprocessed CIE L\*a\*b\* data. Hypertuning was performed with a regularization  
206 parameter (C) between -4 and 4, randomly distributed on a logarithmic scale and the L2 penalty function.  
207 To start with the logistic regression, the data is mapped between 0 and 1 with the logistic function. Logistic  
208 regression creates a model that calculates the probability of the dependent variable Y based on independent  
209 variables according to the following equation.

210 
$$P(Y = 1) = \frac{1}{1 + e^{(\beta_0 + \beta_1 x)}}$$

211 with:

212 P(Y = 1): The probability that the dependent variable Y is 1

213  $\beta_0$ : Intercept of y-Axis

214  $\beta_1$ : Coefficient for the corresponding dependent variable x

215 Support Vector Machine

216 To increase the flexibility of the algorithm, a soft margin SVM was used. The optimization problem is  
217 described as the maximization of the margin with the fewest number of misclassifications. The  
218 optimization problem can be solved with the kernel functions listed in Table 1.

219 The following hyperparameters were optimized: kernel (linear, polynomial, radial basis function (RBF),  
220 and sigmoidal), the penalty for misclassification C (between 1 and 50), and polynomial degree (2,3,4 for  
221 the polynomial kernel). The parameter controlling the individual impact of each data point for the  
222 construction of the hyperplane  $\gamma$  was set to different values depending on the data input. For CIE L\*a\*b\*  
223 data,  $\gamma$  was set between 0.01 and 0.9 for the polynomial, RBF, and sigmoidal kernel. Due to the  
224 significantly increased number of features,  $\gamma$  was set more restrictive between 0.001 and 0.009 for full  
225 transmission spectra.

### 226 Extreme Gradient Boosting

227 XGBoost was applied only to the unprocessed transmission spectra. To minimize the possibility of  
228 overfitting due to the flexibility of the algorithm and the large number of variables, a range of conservative  
229 hyperparameters were used for model optimization. The optimization parameters were chosen as follows:  
230 maximum number of estimators (1, 10, 25), maximum depth of estimators (1, 2, 3), learning rate (0.01,  
231 0.05, 0.1), gain (0, 0.25, 1), L2 regularization (0, 1, 10), ratio between positive and negative classes (1, 3,  
232 5). The fraction of the training samples (subsample) was set to 90 %. The fraction of features  
233 (colsample\_bytree) was set to 50 %. The binary logistic objective function was used.

### 234 Software and Data Analysis

235 The computation of the CIE L\*a\*b\* coordinates from the transmission spectra was performed using the  
236 JASCO spectra manager v. 2.15.01 (JASCO, Japan). Data preprocessing, model construction, and model  
237 performance evaluation were then performed using Python 3 with the NumPy (Harris et al. 2020), pandas  
238 (The pandas development team 2024), scikit-learn (Pedregosa F et al. 2011), seaborn (Waskom 2021),  
239 matplotlib (Hunter 2007), and XGB (Chen und Guestrin 2016) libraries. Plotly and Dash libraries were  
240 used to program the browser application “Blanc de Noir Check” (Plotly Technologies Inc. 2015).

241

242 **Results**

243 The accuracy of all optimized models (logistic regression on CIE L\*a\*b\*, SVM on CIE L\*a\*b\*, SVM on  
244 transmission spectra, and XGBoost on transmission spectra) of the training dataset is displayed as boxplots  
245 in Figure 2. During hyperparameter tuning, every possible parameter combination was used to calculate  
246 the accuracy of the prediction. Overall, the best classification was achieved with SVM (RBF kernel) on  
247 full transmission spectra with an accuracy of 0.93, closely followed by the SVM (RBF kernel) on CIE  
248 L\*a\*b\* coordinates with an accuracy of 0.88 (with  $C = 21$  and  $\gamma = 0.8$ ). However, during validation of the  
249 SVM on transmission spectra with the test dataset the accuracy yielded 0.85 and was almost 10 % lower  
250 than the accuracy of the training dataset (0.93) indicating overfitting. Therefore, the hyperparameter  
251 setting with the smallest distance between the accuracy of training and test data was selected as the optimal  
252 parameter setting with an accuracy in the training dataset of 0.88 using SVM on transmission spectra (with  
253  $C = 1$  and  $\gamma = 0.005$ ). The data points in Figure 2 show the accuracy of the optimal parameter settings,  
254 where the accuracy of the training and test data are as close together as possible. With the simplest model,  
255 logistic regression on CIE L\*a\*b\* coordinates, a maximum accuracy of 0.77 (with  $C = 11.51$  and L2  
256 regularization) could be achieved. However, the accuracy of the training data of logistic regression on CIE  
257 L\*a\*b\* coordinates was 8 % higher than the accuracy of the test data also suggesting overfitting (Figure  
258 2).

259 Using SVM on CIE L\*a\*b\* and transmission spectra, the RBF kernel was the best possible kernel for  
260 both data types. When using transmission spectra instead of CIE L\*a\*b\* coordinates to train the SVM,  
261 the accuracy of the linear and polynomial kernel was increased, while the accuracy of the sigmoidal kernel  
262 decreased. The hyperparameter-optimized XGBoost model showed an accuracy of 0.85, which is very  
263 close to the accuracy of 0.88 of the best model (SVM RBF on CIE L\*a\*b\*) (with  $n\_estimators = 25$ ,  
264  $max\_depth = 3$ ,  $learning\ rate = 0.05$ ,  $gain = 0.25$ ).

265 The results displayed in Figure 2 allow an in-depth comparison between the different kernel functions of  
266 SVM, logistic regression, and XGBoost of one training/test data split.

267 To obtain a more conservative estimation of the variance of classification, the complete model  
268 optimization was repeated with 100 random splits of the training and test data. Figure 3 shows the highest  
269 accuracy of the optimized models after hyperparameter tuning for each training and test dataset. The  
270 difference in the accuracy values for the different training/test data splits was similar for all models.  
271 Relative standard deviations (RSD) of the training dataset were as follows: SVM on CIE L\*a\*b\* (1.85 %);  
272 SVM on transmission spectra (1.75 %); XGBoost on transmission spectra (2.83 %); logistic regression on  
273 CIE L\*a\*b\* (3.55 %). To make a generalized assumption, the 100 test datasets were also evaluated. The  
274 spectra-trained SVM showed an RSD of 3.11 % in the test datasets followed by CIE L\*a\*b\* trained SVM  
275 with 3.55 %. This repetitive approach of model optimization revealed that XGBoost had a higher  
276 classification variation than the SVM models. The test datasets showed an RSD of 5.1 %, being closer to  
277 the logistic regression of 6.5 %.

278 To retain further insights into the model performances, confusion matrices were built using the test dataset  
279 for the optimized models. Overfitting of the logistic regression could be substantiated by the confusion  
280 matrix displayed in Figure 4. The F1-score of the logistic regression in the test dataset was 0.72; a higher  
281 value of 0.77 was obtained for the training dataset indicating overfitting. The poor performance could be  
282 explained by the simplicity of the algorithm. SVM is slightly more complex than logistic regression.

283 The confusion matrix of the SVM on transmission spectra yielded an F1-score of 0.94 in the test dataset  
284 and 0.91 in the training dataset. Very similar F1-scores were obtained from SVM on CIE L\*a\*b\*  
285 coordinates with 0.93 for the test dataset and 0.92 for the training dataset. The optimized XGBoost model  
286 resulted in an F1-score of 0.92 in the test dataset and 0.93 in the training dataset.

287 Figure 5 depicts the decision boundaries for the three possible projection planes of the CIE L\*a\*b\* color  
288 space of the optimized SVM model shown in Figure 2. In the L\*a\* color plane (Figure 5A) the decision  
289 boundary and the margin separated the Blanc de noir wines from the rosé and the white wines. This is not  
290 the case in the L\*b\* color plane (Figure 5B), where Blanc de noir wines could not be separated from the  
291 rosé and white wines. The a\*b\* color plane, like the L\*a\* color plane, showed a distinct decision boundary  
292 and margin (Figure 5C).

293

294

295

296 **Discussion**

297 When transmission spectra were used as input data for SVM, the risk of overfitting was relatively high  
298 compared to SVM on CIE L\*a\*b\* coordinates, due to the high number of features and the relatively low  
299 sample size (Figure 2). The risk of overfitting can be caused by co-linearity between the features in the  
300 transmission spectra. A transformation of the high-dimensional transmission spectra (420 dimensions) to  
301 low-dimensional CIE L\*a\*b\* coordinates (3 dimensions) mitigates the co-linearity and decreases the risk  
302 of overfitting. Another publication using a high number of features with a relatively small sample size  
303 also observed overfitting due to co-linearity. Therefore, reducing the features with this dataset seems to  
304 be the optimal approach (Xiong et al. 2019).

305 Despite being not as affected by co-linearity in comparison to SVM and logistic regression, XGBoost was  
306 not the optimal algorithm here, given its flexibility. In another study, the performance of XGBoost  
307 plummets when the dataset is too small (Zou et al. 2022). Although the accuracy of XGBoost was  
308 relatively high in this dataset, the limited sample size could lead to an unstable model that produces  
309 unstable results when new data is presented. This could also explain the relatively high RSD in the test  
310 dataset. The RSDs of the test datasets were generally higher than the RSDs of the training data, likely  
311 because the training datasets are known to the algorithm and the test datasets are unknown to the algorithm  
312 (Figure 3).

313 Dimensionality reduction of full transmission Vis-spectra to CIE L\*a\*b\* coordinates reduced the risk of  
314 overfitting. CIE L\*a\*b\* coordinates are also more intuitive for potential users, as it is easier to visualize  
315 the decision boundary of the SVM for the three dimensions of the CIE L\*a\*b\* color space, compared to  
316 the 420 dimensions of the transmission data. Therefore, SVM with the RBF kernel on CIE L\*a\*b\* was  
317 considered the optimal mode for further development.

318 The confusion matrices, depicted in Figure 4, confirm that the SVM was best suited for the differentiation  
319 Blanc de noir from rosé and white wine. In contrast, logistic regression is not very well suited to  
320 differentiate between the wine styles. Despite the difficulty of comparing the performance of different  
321 algorithms from different studies, the performance of the SVM used in this study is comparable to the  
322 performance of the SVM using the Fisher Iris-dataset and can therefore be considered highly efficient (Iris  
323 - UCI Machine Learning Repository 2024; Fisher RA 1936). A direct comparison with other literature on  
324 wine ML research is not possible. However, an indirect comparison with Bhardwaj et al. 2022 as well as  
325 Kumar et al. 2020 shows that the F1 scores and accuracies of the SVM on CIE L\*a\*b\* is higher, although  
326 the goal of these publications is much more difficult to achieve. The only conclusion to be drawn from  
327 this observation is that SVM on CIE L\*a\*b\* is more suitable for the objective of this study than the  
328 objectives of the referenced literature (Bhardwaj et al. 2022; Kumar et al. 2020).

329 The reason for the low resolution in the L\*b\* color plane in Figure 5B is explained by the absence of red-  
330 colored anthocyanins in white wine. Rosé wines and Blanc de noir wines, however, are produced from  
331 red grapes and therefore contain a certain amount of anthocyanins. Consequently, the a\* value is of higher  
332 importance than the b\* value. The high number of support vectors between Blanc de noir wines and white  
333 wines in Figure 5A and Figure 5C indicate a more difficult differentiation between these two classes. This  
334 aligns with findings from acceptance threshold studies performed by Kok and Ellis, where the  
335 differentiation of Blanc de noir and white wines was difficult for the human eye (Ellis und Kok 2017).

336



337 **Conclusion**

338 This study assessed the applicability of logistic regression, SVM, and XGBoost to classify different wines  
339 based on spectrophotometric data. SVM outperformed logistic regression and XGBoost, either because  
340 the model was too simple (logistic regression) or too flexible (XGBoost). In both cases, the models  
341 produced unstable results. The RBF kernel showed the best performance for the classification of Blanc de  
342 noir wines with SVM on CIE L\*a\*b\* coordinates and with SVM on transmission spectra. The dimensional  
343 reduction from transmission spectra to CIE L\*a\*b\* coordinates using SVM showed no significant  
344 difference to the SVM trained with transmission spectra. The comprehensibility and usability increased  
345 significantly as only three parameters are needed for the classification. Therefore, SVM on CIE L\*a\*b\*  
346 was the best-suited model for the classification. The use of transmission spectra as training input for SVM  
347 also led to a higher risk of overfitting due to co-linearity between the transmission values of different  
348 wavelengths. This could also be mitigated by dimensional reduction to CIE L\*a\*b\* coordinates.  
349 Additionally, visualization of the decision boundaries in the three color planes showed that the a\*  
350 coordinate is essential for the classification of Blanc de noir wines. To make this machine learning  
351 algorithm available to lab technicians and winemakers we developed an interactive, user friendly graphic  
352 user interface (GUI) in form of a dashboard. For this purpose, the Plotly's Dash library was used in Python.  
353 The user only needs to input the CIE L\*a\*b\* coordinates of their wine to the application, and the wine is  
354 classified according to the parameter-optimized SVM model. The browser application can be reached  
355 online using the following link: <https://blancdenoircheck-217f19ee6419.herokuapp.com>.

356 **References**

- 357 Andrew, H. Beck; Ankur R. Sangoi; Samuel Leung; Robert J. Marinelli; Torsten O. Nielsen; Marc J. van  
358 de Vijver et al. (2011): Systematic Analysis of Breast Cancer Morphology Uncovers Stromal Features  
359 Associated with Survival. In: *Science Translational Medicine* 3 (108), 108ra113-108ra113. DOI:  
360 10.1126/scitranslmed.3002564.
- 361 Bennett, Kristin P.; Campbell, Colin (2000): Support vector machines. In: *SIGKDD Explor. Newsl.* 2 (2),  
362 S. 1–13. DOI: 10.1145/380995.380999.
- 363 Bhardwaj, Piyush; Tiwari, Parul; Olejar, Kenneth; Parr, Wendy; Kulasiri, Don (2022): A machine learning  
364 application in wine quality prediction. In: *Machine Learning with Applications* 8, S. 100261. DOI:  
365 10.1016/j.mlwa.2022.100261.
- 366 Boser, Bernhard E.; Guyon, Isabelle M.; Vapnik, Vladimir N. (1992): A Training Algorithm for Optimal  
367 Margin Classifiers. In: Proceedings of the Fifth Annual Workshop on Computational Learning Theory.  
368 New York, NY, USA: Association for Computing Machinery (COLT '92), S. 144–152.
- 369 Breiman, Leo (2001): Random Forest. In: *Machine Learning* 45 (1), S. 5–32. DOI:  
370 10.1023/A:1010933404324.
- 371 Bundesrat 2021: Änderungen und Entschließung zur Vierundzwanzigsten Verordnung zur Änderung der  
372 Weinverordnung (Drucksache 175/21).
- 373 Chen, Tianqi; Guestrin, Carlos (2016): XGBoost. In: Balaji Krishnapuram (Hg.): Proceedings of the 22nd  
374 ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. KDD '16: The 22nd  
375 ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. San Francisco  
376 California USA, 2016. Association for Computing Machinery-Digital Library; ACM Special Interest

- 377 Group on Management of Data; ACM Special Interest Group on Knowledge Discovery in Data. New  
378 York, NY: ACM (ACM Digital Library), S. 785–794.
- 379 Chopra, Hetarth; Singh, Harsh; Bamrah, Manpreet Singh; Mahbubani, Falesh; Verma, Ashish; Hooda,  
380 Nishtha et al. (2021): Efficient Fruit Grading System Using Spectrophotometry and Machine Learning  
381 Approaches. In: *IEEE Sensors J.* 21 (14), S. 16162–16169. DOI: 10.1109/JSEN.2021.3075465.
- 382 Coulon-Leroy, Cécile; Pouzalgues, Nathalie; Cayla, Laure; Symoneaux, Ronan; Masson, Gilles (2018):  
383 Is the typicality of “Provence Rosé wines” only a matter of color? In: *OENO One* 52 (4). DOI:  
384 10.20870/oenone.2018.52.4.2125.
- 385 Cristianini, Nello; Ricci, Elisa (2008): Support Vector Machines. In: Ming-Yang Kao (Hg.): Encyclopedia  
386 of algorithms. With 38 tables. New York, NY: Springer (Springer reference), S. 928–932.
- 387 Di Nonno, Sarah; Ulber, Roland (2021): Smartphone-based optical analysis systems. In: *Analyst* 146 (9),  
388 S. 2749–2768. DOI: 10.1039/D1AN00025J.
- 389 Di Nonno, Sarah; Ulber, Roland (2022): Portuino-A Novel Portable Low-Cost Arduino-Based Photo- and  
390 Fluorimeter. In: *Sensors (Basel, Switzerland)* 22 (20). DOI: 10.3390/s22207916.
- 391 Ellis, L. P.; Kok, C. (2017): Colour Changes in Blanc de Noir Wines during Ageing at Different  
392 Temperatures and its Colour Preference Limits. In: *SAJEV* 8 (1). DOI: 10.21548/8-1-2324.
- 393 DIN EN ISO/CIE 11664-1: Farbmatrik - Teil 1: CIE farbmtrische Normalbeobachter (ISO/CIE 11664-  
394 1:2019).
- 395 DIN EN ISO/CIE 11664-2: Farbmatrik - Teil 2: CIE Normlichtarten (ISO/CIE DIS 11664-2.2:2021).
- 396 DIN EN ISO/CIE 11664-4: Farbmatrik - Teil 4: CIE 1976 L\*a\*b\* Farbraum (ISO/CIE 11664-4:2019);
- 397 Fisher RA (1936): Iris Dataset.

- 398 Harris, Charles R.; Millman, K. Jarrod; van der Walt, Stéfan J.; Gommers, Ralf; Virtanen, Pauli;  
399 Cournapeau, David et al. (2020): Array programming with NumPy. In: *Nature* 585 (7825), S. 357–362.  
400 DOI: 10.1038/s41586-020-2649-2.
- 401 Hensel, Marcel; Di Nonno, Sarah; Mayer, Yannick; Scheiermann, Marina; Fahrner, Jörg; Durner, Dominik;  
402 Ulber, Roland (2022): Specification and Simplification of Analytical Methods to Determine Wine Color.  
403 In: *Processes* 10 (12), S. 2707. DOI: 10.3390/pr10122707.
- 404 Hensel, Marcel; Scheiermann, Marina; Fahrner, Jörg; Durner, Dominik (2023): New Insights into Wine  
405 Color Analysis: A Comparison of Analytical Methods to Sensory Perception for Red and White Varietal  
406 Wines. In: *Journal of agricultural and food chemistry*. DOI: 10.1021/acs.jafc.3c01284.
- 407 Hossain, Sharif; Chow, Christopher W. K.; Hewa, Guna A.; Cook, David; Harris, Martin (2020):  
408 Spectrophotometric Online Detection of Drinking Water Disinfectant: A Machine Learning Approach. In:  
409 *Sensors (Basel, Switzerland)* 20 (22). DOI: 10.3390/s20226671.
- 410 Hsieh, E. (2011): Identifying Important Risk Factors for Survival in Patient With Systolic Heart Failure  
411 Using Random Survival Forests. In: *Circulation: Cardiovascular Quality and Outcomes* 4 (1), S. 39–45.  
412 DOI: 10.1161/CIRCOUTCOMES.110.939371.
- 413 Hunter, J. D. (2007): Matplotlib: A 2D graphics environment. In: *Computing in Science & Engineering* 9  
414 (3), S. 90–95. DOI: 10.1109/MCSE.2007.55.
- 415 Irene Marivel Nolasco Perez; Amanda Teixeira Badaró; Sylvio Barbon; Ana Paula AC Barbon; Marise  
416 Aparecida Rodrigues Pollonio; Douglas Fernandes Barbin (2018): Classification of Chicken Parts Using  
417 a Portable Near-Infrared (NIR) Spectrophotometer and Machine Learning. In: *Appl. Spectrosc.* 72 (12),  
418 S. 1774–1780.

- 419 Iris - UCI Machine Learning Repository (2024). Online verfügbar unter  
420 <https://archive.ics.uci.edu/dataset/53/iris>, zuletzt aktualisiert am 08.05.2024, zuletzt geprüft am  
421 08.05.2024.
- 422 Ishwaran, Hemant; Kogalur, Udaya B.; Blackstone, Eugene H.; Lauer, Michael S. (2008): Random  
423 survival forests. In: *Ann. Appl. Stat.* 2 (3). DOI: 10.1214/08-AOAS169.
- 424 Issa-Issa, Hanán; Hernández, Francisca; Lipan, Leontina; López-Lluch, David; Carbonell-Barrachina,  
425 Ángel A. (2021): Quality, Nutritional, Volatile and Sensory Profiles and Consumer Acceptance of  
426 Fondillón, a Sustainable European Protected Wine. In: *Agronomy* 11 (9), S. 1701. DOI:  
427 10.3390/agronomy11091701.
- 428 J. Guild (1931): The colorimetric properties of the spectrum. In: *Phil. Trans. R. Soc. Lond. A* 230 (681-  
429 693), S. 149–187. DOI: 10.1098/rsta.1932.0005.
- 430 Kumar, Sunny; Agrawal, Kanika; Mandan, Nelshan (2020): Red Wine Quality Prediction Using Machine  
431 Learning Techniques. In: 2020 International Conference on Computer Communication and Informatics.  
432 January 22-24, 2020, Coimbatore, India. 2020 International Conference on Computer Communication and  
433 Informatics (ICCCI). Coimbatore, India, 2020. Sri Shakthi Institute of Engineering and Technology;  
434 Institute of Electrical and Electronics Engineers. Piscataway, NJ: IEEE, S. 1–6.
- 435 Morrot, G.; Brochet, F.; Dubourdieu, D. (2001): The color of odors. In: *Brain and language* 79 (2), S.  
436 309–320. DOI: 10.1006/brln.2001.2493.
- 437 Nguyen, Thi H.; Durner, Dominik (2023): Sensory evaluation of wine aroma: Should color-driven  
438 descriptors be used? In: *Food Quality and Preference* 107, S. 104844. DOI:  
439 10.1016/j.foodqual.2023.104844.

- 440 OIV (Hg.) (2021): Compendium of international methods of wine and must analysis. International  
441 Organization of wine and Vine. Edition 2021. Paris: OIV.
- 442 Parpinello, Giuseppina Paola; Versari, Andrea; Chinnici, Fabio; Galassi, Sergio (2009): Relationship  
443 among sensory descriptors, consumer preference and color parameters of Italian Novello red wines. In:  
444 *Food Research International* 42 (10), S. 1389–1395. DOI: 10.1016/j.foodres.2009.07.005.
- 445 Pedregosa F; Varoquaux G; Gramfort A; Michel V; Thirion B; Grisel O et al. (2011): Scikit-learn:  
446 Machine Learning in Python. In: *Journal of Machine Learning Research* 12 (85), S. 2825–2830. Online  
447 verfügbar unter <http://jmlr.org/papers/v12/pedregosa11a.html>.
- 448 Plotly Technologies Inc. (2015): Plotly. Collaborative data science Publisher: Plotly Technologies Inc.  
449 Montréal, QC: Plotly Technologies Inc. Online verfügbar unter <https://plot.ly>.
- 450 Ramirez, A. M.C. (2021): Applications of machine learning in spectroscopy. In: *Applied Spectroscopy*  
451 *Reviews* 56 (8-10), S. 733–763. DOI: 10.1080/05704928.2020.1859525.
- 452 Shaw, Bipul; Suman, Ankur Kumar; Chakraborty, Biswarup (2020): Wine Quality Analysis Using  
453 Machine Learning. In: *Emerging Technology in Modelling and Graphics*: Springer, Singapore, S. 239–  
454 247. Online verfügbar unter [https://link.springer.com/chapter/10.1007/978-981-13-7403-6\\_23](https://link.springer.com/chapter/10.1007/978-981-13-7403-6_23).
- 455 Stiles, W. S.; Burch, J. M. (1959): N.P.L. Colour-matching Investigation: Final Report (1958). In: *Optica*  
456 *Acta: International Journal of Optics* 6 (1), S. 1–26. DOI: 10.1080/713826267.
- 457 The pandas development team (2024): pandas-dev/pandas: Pandas: Zenodo.
- 458 Waskom, Michael (2021): seaborn: statistical data visualization. In: *JOSS* 6 (60), S. 3021. DOI:  
459 10.21105/joss.03021.

460 Wright, W. D. (1929): A re-determination of the trichromatic coefficients of the spectral colours. In:  
 461 *Trans. Opt. Soc.* 30 (4), S. 141–164. DOI: 10.1088/1475-4878/30/4/301.

462 Wright, W. D. (1930): A re-determination of the mixture curves of the spectrum. In: *Trans. Opt. Soc.* 31  
 463 (4), S. 201–218. DOI: 10.1088/1475-4878/31/4/303.

464 Xiong, Bo; Newton, Sidney; Li, Vera; Skitmore, Martin; Xia, Bo (2019): Hybrid approach to reducing  
 465 estimating overfitting and collinearity. In: *ECAM* 26 (10), S. 2170–2185. DOI: 10.1108/ECAM-08-2018-  
 466 0353.

467 Zou, Miao; Jiang, Wu-Gui; Qin, Qing-Hua; Liu, Yu-Cheng; Li, Mao-Lin (2022): Optimized XGBoost  
 468 Model with Small Dataset for Predicting Relative Density of Ti-6Al-4V Parts Manufactured by Selective  
 469 Laser Melting. In: *Materials (Basel, Switzerland)* 15 (15). DOI: 10.3390/ma15155298.

470

## 471 Tables

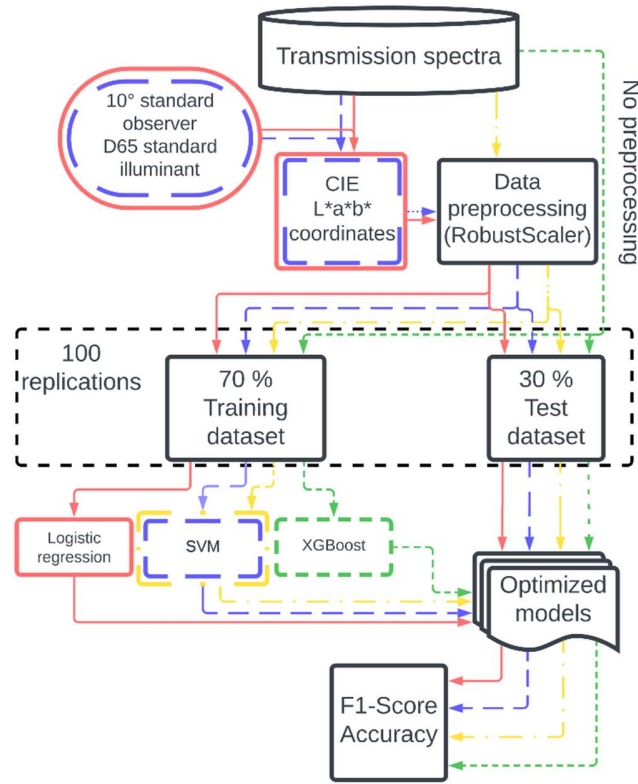
472 *Table 1: Kernels used and the regularization parameters (Cristianini und Ricci 2008; Bennett und Campbell 2000).*

Kernel Name	Kernel Function	Regularization parameter
Linear	$K(x, y) = \langle x, y \rangle$	C
Polynomial	$K(x, y) = (\gamma \langle x, y \rangle + r)^d$	C, $\gamma$ , d
RBF	$K(x, y) = e^{-\gamma \ x_i - y_i\ ^2}$	C, $\gamma$
Sigmoidal	$K(x, y) = \tanh(\gamma \langle x, y \rangle) + r$	C, $\gamma$

473

474

475 **Figures**

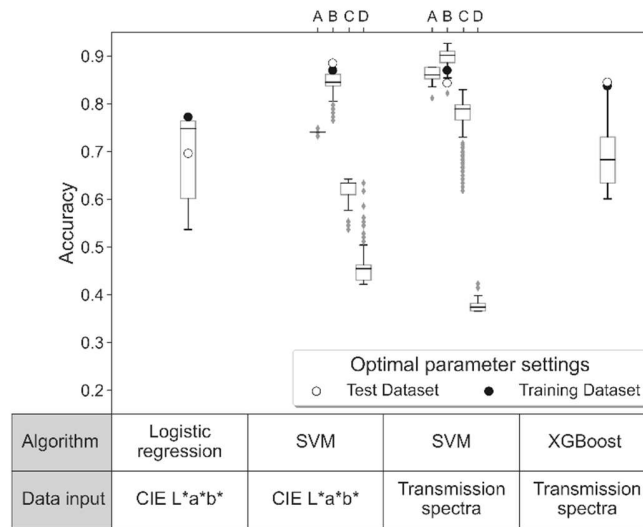


476

477 *Figure 1: Data pipeline starting from the transmission spectra as raw data and following through the different computation steps. Colors*  
478 *indicate different models. Logistic regression on CIE L\*a\*b\* (red), SVM on CIE L\*a\*b\* (blue), SVM on transmission spectra (yellow), and*  
479 *XGBoost on transmission spectra (green)*

480





481

482

483

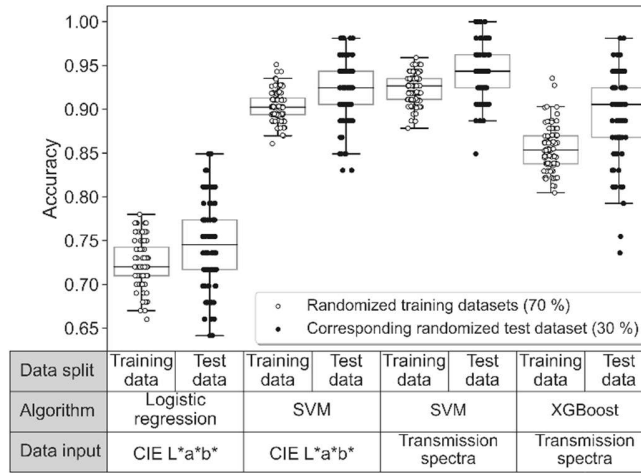
484

485

486

487

Figure 2: Accuracy of the parameter-optimized models after hyperparameter tuning displayed as boxplots comparing logistic regression, SVM, and XGBoost, as well as the data input CIE L\*a\*b\* and transmission spectra. The boxplots contain quantile 1, the median line, and quantile 3. The whiskers display the 1.5-fold inter-quantile range the rhombi are outliers. The data points indicate the most accuracy of the most suited hyperparameter setting. The SVM algorithm is conducted with the (A) linear kernel, (B) RBF kernel, (C) polynomial kernel, and (D) sigmoidal kernel.

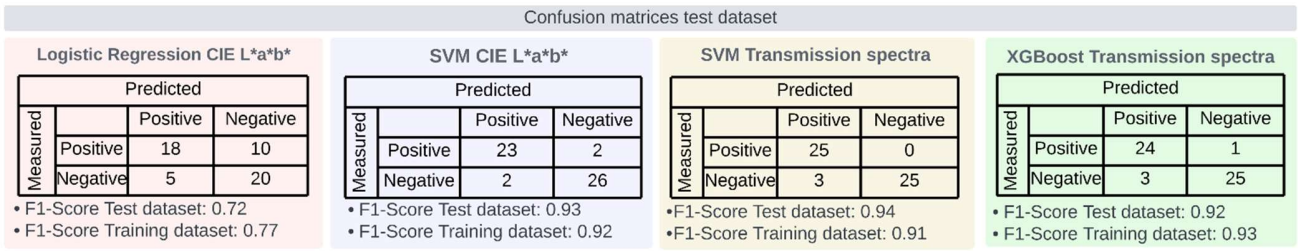


488

489 *Figure 3: Highest accuracy of the optimized models after hyperparameter tuning for each training and test dataset with 100 randomized*  
 490 *training datasets and the corresponding test datasets. The boxplots show quantile 1, median line, quantile 3. The whiskers depict the 1.5-fold*  
 491 *inter-quantile range.*

492

493

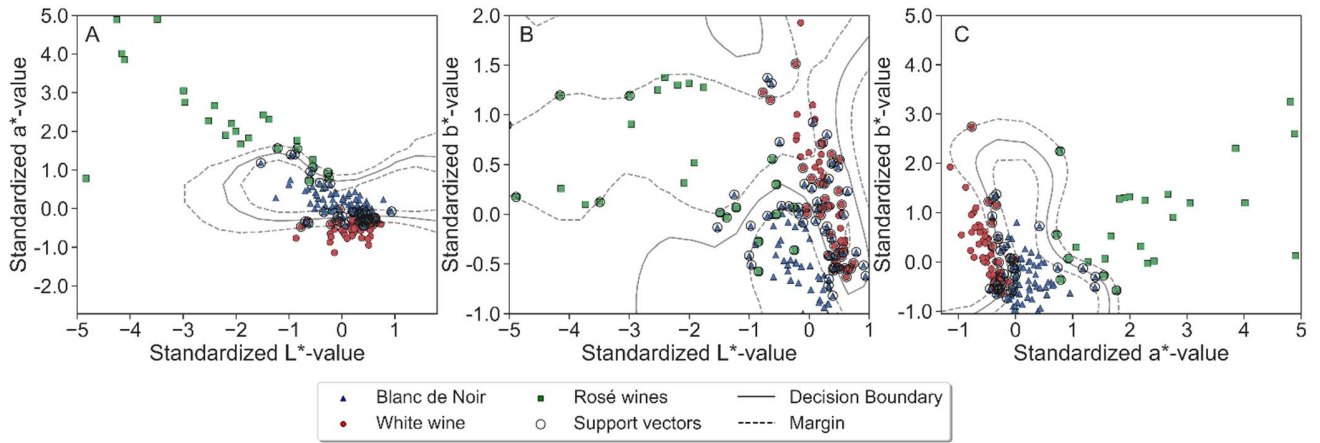


494

495 *Figure 4: Confusion matrices for the test dataset ( $n = 53$ ) of the used algorithms logistic regression on CIE  $L^*a^*b^*$ , SVM on transmission*  
 496 *spectra, SVM on CIE  $L^*a^*b^*$ , and XGBoost on transmission spectra, as well as the F1-scores for test and training dataset ( $n = 128$ ).*

497

498



Sample ID	Wine type	Origin	Vintage	Grape variety	Bottle volume	Bottle type	Closure Type	Alcohol	Residual sugar	pH	Total acidity	Free SO <sub>2</sub>	Total SO <sub>2</sub>	Experiment date	Bottle age at experiment
WW1	white wine	Spain	2019	Chardonnay	0.75 L	Burgundy	Natural cork	11.1	7.6	3.1	6.5	37.5	82.8	04/29/2021	2
WW2	white wine	USA	2018	Chardonnay	0.75 L	Bordeaux	Natural cork	12.6	5.2	3.4	5.2	42.4	77.1	04/29/2021	3
WW3	white wine	Germany	2018	Chardonnay	0.75 L	Bordeaux	BVS	13.5	1.1	3.2	6.2	44.2	86.0	04/29/2021	3
WW4	white wine	Italy	2019	Chardonnay	0.75 L	Burgundy	BVS	12.7	5.2	3.2	6.1	25.7	71.4	04/29/2021	2
WW5	white wine	Australia	2019	Chardonnay	0.75 L	Bordeaux	Synthetic cork	13.2	5.4	3.3	6.3	32.8	94.2	04/29/2021	2
WW6	white wine	Spain	2019	Chardonnay	0.75 L	Burgundy	BVS	11.7	1.7	3.4	5.1	43.9	78.5	04/29/2021	2
WW7	white wine	Germany	2016	Chardonnay	0.75 L	Bordeaux	BVS	11.5	1.8	3.4	5.5	18.5	82.5	04/29/2021	5
WW8	white wine	Germany	2015	Riesling	0.75 L	Alsace	BVS	11.7	3.7	3.2	6.5	9.7	81.4	05/28/2021	6
WW9	white wine	France	2017	Riesling	0.75 L	Bordeaux	BVS	12.8	1.2	3.1	6.9	16.5	60.1	05/28/2021	4
WW10	white wine	USA	2013	Riesling	0.75 L	Burgundy	BVS	11.2	6.1	3.3	6.8	8.9	63.8	05/28/2021	8
WW11	white wine	Germany	2018	Riesling	0.75 L	Alsace	BVS	11.0	3.7	3.1	7.3	18.9	92.2	05/28/2021	3
WW12	white wine	Germany	2019	Riesling	0.75 L	Alsace	Natural cork	11.8	5.3	3.1	7.2	20.7	86.9	05/28/2021	2
WW13	white wine	Germany	2020	Riesling	0.75 L	Alsace	Natural cork	12.7	8.8	3.2	7.0	30.6	70.1	05/28/2021	1
WW14	white wine	Germany	2020	Riesling	0.75 L	Alsace	Natural cork	12.0	8.8	3.3	7.2	35.1	86.4	05/28/2021	1
WW15	white wine	Germany	2019	Riesling	0.75 L	Burgundy	Agglomerated cork	11.8	63.7	3.1	7.2	29.6	118.1	05/28/2021	2
WW16	white wine	Germany	2018	Riesling	0.75 L	Bordeaux	BVS	11.8	56.1	3.3	7.4	43.4	116.9	05/28/2021	3
WW17	white wine	Germany	2019	Riesling	0.75 L	Bordeaux	BVS	11.5	73.3	3.2	7.1	32.3	105.0	05/28/2021	2
WW18	white wine	Germany	2018	Riesling	0.75 L	Bordeaux	BVS	12.8	64.0	3.2	8.0	42.6	119.6	05/28/2021	3
WW19	white wine	Germany	2019	Riesling	0.75 L	Alsace	Natural cork	11.5	47.6	3.2	7.6	37.9	107.7	05/28/2021	2
WW20	white wine	Germany	2019	Riesling	0.75 L	Alsace	Natural cork	12.1	72.0	3.4	7.8	36.8	137.4	05/28/2021	2
WW21	white wine	Germany	2018	Riesling	0.75 L	Alsace	Natural cork	12.4	47.6	3.2	7.7	38.7	101.9	05/28/2021	3
WW22	white wine	Italy	2019	Pinot gris	0.75 L	Bordeaux	BVS	12.4	2.6	3.4	6.4	28.0	74.2	05/28/2021	2
WW23	white wine	Italy	2020	Pinot gris	0.75 L	Alsace	BVS	13.8	8.5	3.2	5.9	43.7	75.5	05/28/2021	1
WW24	white wine	France	2017	Pinot gris	0.75 L	Bordeaux	BVS	12.1	4.4	3.4	6.2	21.1	71.5	05/28/2021	4
WW25	white wine	New Zealand	2018	Pinot gris	0.75 L	Burgundy	BVS	11.8	4.7	3.5	5.9	37.3	79.6	05/28/2021	3
WW26	white wine	Germany	2019	Pinot gris	0.75 L	Bordeaux	BVS	11.8	2.7	3.4	6.6	33.8	68.8	05/28/2021	2
WW27	white wine	Germany	2019	Pinot gris	0.75 L	Alsace	Agglomerated cork	13.6	8.5	3.3	6.5	33.7	74.6	05/28/2021	2
WW28	white wine	Germany	2019	Pinot gris	0.75 L	Alsace	BVS	13.9	1.5	3.2	5.8	33.5	71.0	05/28/2021	2
WW29	white wine	France	2019	Sauvignon Blanc	0.75 L	Bordeaux	BVS	13.0	2.3	3.4	5.9	37.6	60.4	05/31/2021	2
WW30	white wine	Germany	2019	Sauvignon Blanc	0.75 L	Burgundy	BVS	11.4	7.6	3.5	6.7	38.1	72.6	05/31/2021	2
WW31	white wine	Italy	2019	Sauvignon Blanc	0.75 L	Burgundy	BVS	13.6	2.1	3.4	5.9	42.4	83.2	05/31/2021	2
WW32	white wine	USA	2019	Sauvignon Blanc	0.75 L	Bordeaux	Natural cork	12.2	3.1	3.5	6.4	44.0	78.4	05/31/2021	2

WW33	white wine	New Zealand	2020	Sauvignon Blanc	0.75 L	Burgundy	BVS	11.7	2.9	3.4	6.0	36.0	80.8	05/31/2021	1
WW34	white wine	France	2018	Sauvignon Blanc	0.75 L	Bordeaux	Glass cork	13.0	6.9	3.1	6.9	44.9	83.8	05/31/2021	3
WW35	white wine	Germany	2018	Sauvignon Blanc	0.75 L	Burgundy	BVS	13.0	2.6	3.1	6.2	37.5	95.4	05/31/2021	3
WW36	white wine	Germany	2019	Pinot Blanc	0.75 L	Alsace	BVS	11.3	7.1	3.3	6.9	24.5	72.1	05/31/2021	2
WW37	white wine	Germany	2020	Pinot Blanc	0.75 L	Bordeaux	BVS	12.6	2.0	3.3	6.9	21.8	62.5	05/31/2021	1
WW38	white wine	Germany	2018	Pinot Blanc	0.75 L	Burgundy	BVS	13.4	4.1	3.2	6.4	38.3	77.7	05/31/2021	3
WW39	white wine	Germany	2017	Pinot Blanc	0.75 L	Burgundy	BVS	13.8	7.8	3.3	5.2	9.5	55.1	05/31/2021	4
WW40	white wine	Germany	2020	Pinot Blanc	0.75 L	Bordeaux	Natural cork	13.3	6.4	3.4	6.2	43.6	84.4	05/31/2021	1
WW41	white wine	Italy	2019	Pinot Blanc	0.75 L	Bordeaux	BVS	11.6	1.6	3.4	6.7	28.6	94.8	05/31/2021	2
WW42	white wine	France	2017	Pinot Blanc	0.75 L	Alsace	Natural cork	11.3	2.2	3.2	6.5	15.2	84.5	05/31/2021	4
WW43	white wine	Germany	2016	Gewürztraminer	0.75 L	Alsace	Agglomerated cork	11.9	76.4	3.1	5.5	25.8	125.5	05/31/2021	5
WW44	white wine	Germany	2018	Gewürztraminer	0.75 L	Burgundy	BVS	12.5	54.1	3.2	5.5	35.1	106.8	05/31/2021	3
WW45	white wine	France	2019	Gewürztraminer	0.75 L	Alsace	BVS	13.0	77.8	3.4	5.5	29.6	136.3	05/31/2021	2
WW46	white wine	Germany	2019	Gewürztraminer	0.75 L	Bordeaux	BVS	13.1	64.1	3.1	6.0	36.7	128.8	05/31/2021	2
WW47	white wine	France	2019	Gewürztraminer	0.75 L	Bordeaux	BVS	13.4	77.9	3.5	5.2	32.8	109.8	05/31/2021	2
WW48	white wine	Spain	2019	Gewürztraminer	0.75 L	Burgundy	BVS	12.3	77.2	3.4	6.1	28.8	111.3	05/31/2021	2
WW49	white wine	Germany	2019	Gewürztraminer	0.75 L	Alsace	BVS	11.6	53.5	3.5	6.1	27.5	126.7	05/31/2021	2
WW50	white wine	Italy	2019	Gewürztraminer	0.75 L	Alsace	BVS	13.7	3.7	3.3	5.6	25.6	72.7	05/31/2021	2
WW51	white wine	Italy	2019	Gewürztraminer	0.75 L	Burgundy	BVS	12.6	4.3	3.2	6.7	44.4	62.5	05/31/2021	2
WW52	white wine	Spain	2018	Gewürztraminer	0.75 L	Bordeaux	Natural cork	11.5	1.4	3.4	6.5	27.5	96.2	05/31/2021	3
WW53	white wine	Germany	2019	Gewürztraminer	0.75 L	Alsace	BVS	11.6	4.7	3.1	5.9	23.1	93.0	05/31/2021	2
WW54	white wine	France	2018	Gewürztraminer	0.75 L	Bordeaux	BVS	11.0	2.5	3.2	5.8	33.5	87.5	05/31/2021	3
WW55	white wine	Germany	2019	Gewürztraminer	0.75 L	Burgundy	BVS	11.3	6.6	3.1	5.4	38.9	71.0	05/31/2021	2
WW56	white wine	Germany	2020	Gewürztraminer	0.75 L	Alsace	BVS	13.1	7.2	3.1	6.3	29.7	77.1	05/31/2021	1
RW1	Rose	Germany	2022	Pinot Noir	0.75 L	Alsace	BVS	12.4	4.1	3.3	6.6	29.4	95.7	05/10/2023	1
RW2	Rose	Germany	2018	Muskatrollinger	0.75 L	Burgundy	BVS	12.2	5.3	3.1	6.0	16.5	58.3	05/10/2023	5
RW3	Rose	France	2020	Cabernet Franc	0.75 L	Bordeaux	BVS	12.0	7.3	3.5	5.2	22.1	61.2	05/10/2023	3
RW4	Rose	Germany	2018	Domfelder	0.75 L	Bordeaux	Natural cork	13.3	6.9	3.2	5.9	21.7	66.9	05/10/2023	5
RW5	Rose	France	2019	Cabernet Franc	0.75 L	Alsace	Natural cork	13.1	1.7	3.2	5.1	29.9	75.2	05/10/2023	4
RW6	Rose	Germany	2019	Domfelder	0.75 L	Burgundy	Natural cork	12.9	7.3	3.5	5.1	18.9	79.6	05/10/2023	4
RW7	Rose	Germany	2022	Tempranillo	0.75 L	Burgundy	Agglomerated cork	12.9	6.0	3.5	6.3	31.1	91.6	05/10/2023	1
RW8	Rose	Germany	2021	Pinot Noir	0.75 L	Alsace	Synthetic cork	13.3	1.7	3.4	6.3	24.9	86.9	05/26/2023	2
RW9	Rose	Germany	2020	Pinot Noir	0.75 L	Alsace	Synthetic cork	12.0	5.7	3.2	6.5	36.2	93.3	05/26/2023	3

RW10	Rose	Germany	2022	Pinot Noir	0.75 L	Burgundy	Agglomerated cork	13.1	6.0	3.2	5.8	41.4	79.9	05/26/2023	1
RW11	Rose	Germany	2021	Blauer Zweigelt	0.75 L	Bordeaux	Natural cork	12.7	3.1	3.1	6.4	41.1	92.6	05/26/2023	2
RW12	Rose	Germany	2022	Cabernet Sauvignon	0.75 L	Bordeaux	BVS	11.8	6.0	3.3	5.7	24.5	95.0	05/26/2023	1
RW13	Rose	Germany	2022	Lemberger	0.75 L	Bordeaux	Agglomerated cork	12.5	6.9	3.1	6.5	20.3	97.7	05/26/2023	1
RW14	Rose	Spain	2020	Syrah	0.75 L	Burgundy	Natural cork	13.8	3.3	3.5	5.4	39.6	85.6	05/26/2023	3
RW15	Rose	France	2021	Cabernet Franc	0.75 L	Bordeaux	Synthetic cork	11.2	1.0	3.2	5.2	24.6	87.0	05/26/2023	2
RW16	Rose	France	2022	Merlot	0.75 L	Burgundy	Agglomerated cork	11.4	7.0	3.1	5.7	27.8	97.1	05/26/2023	1
RW17	Rose	Turkey	2021	Sangiovese	0.75 L	Burgundy	BVS	13.9	5.0	3.5	5.0	39.9	64.9	05/26/2023	2
RW18	Rose	Belgium	2022	Cabernet Cortis	0.75 L	Burgundy	BVS	12.5	3.5	3.3	5.7	40.2	96.2	05/30/2023	1
RW19	Rose	Germany	2022	Pinot Noir	0.75 L	Burgundy	BVS	12.0	4.0	3.4	6.1	32.9	86.4	05/30/2023	1
RW20	Rose	Germany	2021	Pinot Noir	0.75 L	Bordeaux	Agglomerated cork	12.9	5.8	3.3	6.2	43.6	88.5	05/30/2023	2
RW21	Rose	Germany	2020	Cabernet Cortis	0.75 L	Alsace	BVS	13.1	6.5	3.1	5.2	33.4	63.1	05/30/2023	3
RW22	Rose	Germany	2021	Lemberger	0.75 L	Alsace	Natural cork	13.6	6.6	3.5	6.7	45.0	69.5	05/30/2023	2
RW23	Rose	Spain	2020	Bobal	0.75 L	Burgundy	Agglomerated cork	13.3	6.0	3.1	5.3	31.1	78.7	05/30/2023	3
RW24	Rose	Spain	2021	Grenache	0.75 L	Bordeaux	Natural cork	13.3	7.6	3.2	5.6	34.3	93.8	05/30/2023	2
RW25	Rose	France	2022	Cabernet Franc	0.75 L	Bordeaux	Agglomerated cork	13.1	6.4	3.3	6.4	25.4	71.7	05/30/2023	1
RW26	Rose	France	2022	Cabernet Sauvignon	0.75 L	Bordeaux	Agglomerated cork	13.4	2.0	3.3	5.6	21.9	61.8	05/30/2023	1
RW27	Rose	Germany	2022	Muskatrollinger	0.75 L	Bordeaux	BVS	12.3	1.1	3.4	6.4	26.8	86.9	05/30/2023	1
RW28	Rose	Chile	2021	Pinot Noir	0.75 L	Burgundy	Natural cork	12.9	7.9	3.3	5.2	37.0	72.5	08/16/2023	2
RW29	Rose	Spain	2021	Grenache	0.75 L	Burgundy	Synthetic cork	12.9	1.9	3.3	5.7	22.5	70.4	08/16/2023	2
RW30	Rose	France	2021	Grenache	0.75 L	Bordeaux	Natural cork	12.4	4.8	3.4	6.0	39.4	62.6	08/16/2023	2
RW31	Rose	Italy	2021	Montepulciano	0.75 L	Bordeaux	BVS	12.6	5.5	3.4	6.1	27.8	77.0	08/16/2023	2
RW32	Rose	Italy	2021	Montepulciano	0.75 L	Burgundy	BVS	11.9	5.7	3.3	5.7	33.7	60.3	08/18/2023	2
RW33	Rose	Germany	2020	Pinot Noir	0.75 L	Bordeaux	Agglomerated cork	11.6	3.6	3.2	6.6	28.0	89.7	08/18/2023	3
RW34	Rose	France	2021	Pinot Noir	0.75 L	Bordeaux	Natural cork	13.7	4.8	3.1	6.5	36.0	92.6	08/18/2023	2
RW35	Rose	France	2021	Cabernet Franc	0.75 L	Burgundy	Synthetic cork	11.9	1.3	3.3	5.9	29.7	76.2	08/18/2023	2
RW36	Rose	Italy	2021	Pinot Noir	0.75 L	Burgundy	BVS	11.5	2.1	3.2	6.3	21.6	72.1	08/18/2023	2
RW37	Rose	France	2021	Grenache	0.75 L	Bordeaux	Agglomerated cork	11.3	1.6	3.5	6.3	36.3	67.8	08/18/2023	2
Bdn1	Blanc de Noir	Germany	2021	Pinot Noir	0.75 L	Burgundy	Agglomerated cork	12.5	8.6	3.1	5.8	20.5	92.8	04/08/2022	1
Bdn2	Blanc de Noir	Germany	2021	Pinot Noir	0.75 L	Bordeaux	BVS	12.5	7.9	3.2	5.1	29.2	67.1	04/08/2022	1
Bdn3	Blanc de Noir	Germany	2021	Pinot Noir	0.75 L	Burgundy	BVS	12.0	7.7	3.4	6.3	22.9	87.3	04/08/2022	1

Bdn4	Blanc de Noir	Germany	2021	Pinot Noir	0.75 L	Burgundy	BVS	12.5	7.5	3.2	5.9	22.2	88.2	04/08/2022	1
Bdn5	Blanc de Noir	Germany	2021	Pinot Noir	0.75 L	Bordeaux	BVS	12.5	7.3	3.3	6.1	42.8	72.8	04/08/2022	1
Bdn6	Blanc de Noir	Germany	2021	Pinot Noir	0.75 L	Bordeaux	BVS	12.5	6.8	3.2	6.7	29.2	75.1	04/08/2022	1
Bdn7	Blanc de Noir	Germany	2021	Pinot Noir	0.75 L	Bordeaux	BVS	12.5	6.4	3.3	6.4	32.9	62.6	04/08/2022	1
Bdn8	Blanc de Noir	Germany	2021	Pinot Noir	0.75 L	Burgundy	BVS	11.5	5.6	3.3	7.1	44.7	71.8	04/08/2022	1
Bdn9	Blanc de Noir	Germany	2021	Pinot Noir	0.75 L	Alsace	BVS	12.0	4.9	3.4	6.6	29.9	91.9	04/08/2022	1
RW38	Rose	Germany	2021	Pinot Noir	0.75 L	Alsace	Agglomerated cork	11.5	0.0	3.3	7.5	22.6	80.6	04/08/2022	1
Bdn11	Blanc de Noir	Germany	2021	Pinot Noir	0.75 L	Alsace	BVS	10.0	44.0	3.3	7.4	40.7	130.1	04/08/2022	1
Bdn12	Blanc de Noir	Germany	2021	Pinot Noir	0.75 L	Burgundy	BVS	12.0	8.1	3.5	6.9	22.4	63.1	04/08/2022	1
Bdn13	Blanc de Noir	Germany	2021	Pinot Noir	0.75 L	Bordeaux	BVS	13.0	7.9	3.3	6.4	26.6	93.4	04/08/2022	1
Bdn14	Blanc de Noir	Germany	2021	Pinot Noir	0.75 L	Burgundy	BVS	12.0	6.3	3.4	6.1	38.7	87.5	04/08/2022	1
Bdn15	Blanc de Noir	Germany	2021	Pinot Noir	0.75 L	Alsace	BVS	12.5	6.0	3.3	6.1	25.5	78.5	04/08/2022	1
Bdn16	Blanc de Noir	Germany	2021	Pinot Noir	0.75 L	Bordeaux	BVS	11.0	6.0	3.2	6.3	29.8	96.3	04/08/2022	1
Bdn17	Blanc de Noir	Germany	2021	Pinot Noir	0.75 L	Burgundy	BVS	13.0	6.0	3.4	7.3	34.6	66.0	04/08/2022	1
Bdn18	Blanc de Noir	Germany	2021	Pinot Noir	0.75 L	Burgundy	BVS	12.5	5.9	3.4	7.4	40.4	67.9	04/08/2022	1
Bdn19	Blanc de Noir	Germany	2021	Pinot Noir	0.75 L	Burgundy	BVS	13.0	5.8	3.3	5.3	23.9	66.0	04/08/2022	1
Bdn20	Blanc de Noir	Germany	2021	Pinot Noir	0.75 L	Burgundy	BVS	11.5	5.3	3.3	7.5	43.5	60.1	04/08/2022	1
Bdn21	Blanc de Noir	Germany	2021	Pinot Noir	0.75 L	Bordeaux	Natural cork	12.0	1.8	3.5	6.6	25.3	70.1	04/08/2022	1
Bdn22	Blanc de Noir	Germany	2020	Pinot Noir	0.75 L	Bordeaux	BVS	12.0	8.3	3.5	6.5	39.4	71.9	04/08/2022	2
Bdn23	Blanc de Noir	Germany	2020	Pinot Noir	0.75 L	Burgundy	BVS	12.5	0.6	3.4	5.3	21.1	65.0	05/03/2022	2
Bdn24	Blanc de Noir	Germany	2019	Pinot Noir	0.75 L	Alsace	BVS	12.5	2.3	3.3	5.9	25.1	78.1	05/03/2022	3
Bdn25	Blanc de Noir	Germany	2021	Pinot Meunier	0.75 L	Alsace	BVS	12.0	7.5	3.2	7.4	42.6	98.9	05/03/2022	1
Bdn26	Blanc de Noir	Germany	2021	Pinot Meunier	0.75 L	Bordeaux	BVS	12.5	6.6	3.3	7.1	33.4	93.0	05/03/2022	1
Bdn27	Blanc de Noir	Germany	2021	Pinot Meunier	0.75 L	Burgundy	BVS	12.0	5.6	3.1	5.8	31.0	80.8	05/03/2022	1
Bdn28	Blanc de Noir	Germany	2020	Pinot Meunier	0.75 L	Bordeaux	BVS	12.0	6.1	3.2	7.1	32.7	94.6	05/03/2022	2
Bdn29	Blanc de Noir	Germany	2021	Pinot Noir	0.75 L	Bordeaux	BVS	12.2	4.8	3.3	5.2	34.7	87.7	05/03/2022	1
Bdn30	Blanc de Noir	Germany	2021	Pinot Noir	0.75 L	Alsace	BVS	12.4	6.7	3.3	5.2	34.3	69.1	05/03/2022	1



Bdn31	Blanc de Noir	Germany	2021	Pinot Noir	0.75 L	Alsace	BVS	12.0	5.3	3.2	6.5	38.4	89.4	05/03/2022	1
Bdn32	Blanc de Noir	Germany	2021	Pinot Noir	0.75 L	Bordeaux	Agglomerated cork	13.8	4.3	3.1	6.1	34.3	70.9	05/03/2022	1
Bdn33	Blanc de Noir	Germany	2021	Pinot Noir	0.75 L	Burgundy	BVS	11.5	4.5	3.4	5.2	32.1	69.4	05/03/2022	1
Bdn34	Blanc de Noir	Germany	2021	Pinot Noir	0.75 L	Bordeaux	BVS	11.2	5.3	3.4	5.2	28.2	70.4	05/03/2022	1
Bdn35	Blanc de Noir	Germany	2021	Portugieser	0.75 L	Burgundy	BVS	12.2	2.2	3.5	6.4	38.8	94.2	05/03/2022	1
Bdn36	Blanc de Noir	Germany	2021	Pinot Noir	0.75 L	Alsace	BVS	13.4	1.4	3.2	5.3	33.3	77.5	05/03/2022	1
Bdn37	Blanc de Noir	Germany	2021	Pinot Noir	0.75 L	Bordeaux	BVS	13.5	3.6	3.2	6.9	36.9	68.8	05/03/2022	1
Bdn38	Blanc de Noir	Germany	2021	Pinot Noir	0.75 L	Burgundy	BVS	12.6	5.9	3.2	6.8	31.7	99.3	05/03/2022	1
Bdn39	Blanc de Noir	Germany	2021	Pinot Noir	0.75 L	Burgundy	BVS	11.3	6.9	3.4	6.5	23.3	62.1	05/03/2022	1
Bdn40	Blanc de Noir	Germany	2021	Pinot Noir	0.75 L	Burgundy	BVS	12.7	3.8	3.3	6.8	40.1	85.8	05/03/2022	1
Bdn41	Blanc de Noir	Germany	2021	Pinot Noir	0.75 L	Burgundy	Synthetic cork	12.6	4.8	3.2	6.8	35.3	73.7	05/03/2022	1
Bdn42	Blanc de Noir	Germany	2021	Pinot Noir	0.75 L	Bordeaux	BVS	13.1	7.3	3.3	5.1	43.6	95.1	05/03/2022	1
Bdn43	Blanc de Noir	Germany	2021	Pinot Noir	0.75 L	Bordeaux	BVS	12.3	1.8	3.3	5.3	23.4	88.7	05/03/2022	1
Bdn44	Blanc de Noir	Germany	2021	Pinot Noir	0.75 L	Burgundy	BVS	13.3	8.0	3.2	6.5	42.7	73.5	05/03/2022	1
Bdn45	Blanc de Noir	Germany	2021	Dolcelto	0.75 L	Bordeaux	Natural cork	12.8	3.3	3.1	6.9	29.6	97.4	05/03/2022	1
Bdn46	Blanc de Noir	Germany	2021	Pinot Noir	0.75 L	Alsace	BVS	12.6	5.3	3.3	5.6	23.2	91.4	05/03/2022	1
Bdn47	Blanc de Noir	Germany	2021	Pinot Noir	0.75 L	Bordeaux	Agglomerated cork	12.8	4.9	3.1	6.8	39.4	70.8	05/03/2022	1
Bdn48	Blanc de Noir	Germany	2021	Pinot Noir	0.75 L	Burgundy	Natural cork	13.5	6.4	3.3	5.2	36.3	88.2	05/03/2022	1
Bdn49	Blanc de Noir	Germany	2021	Pinot Noir	0.75 L	Bordeaux	BVS	11.3	3.8	3.2	6.5	42.5	65.3	05/03/2022	1
Bdn50	Blanc de Noir	Germany	2021	Pinot Noir	0.75 L	Bordeaux	BVS	13.4	5.0	3.2	6.1	25.0	72.4	05/03/2022	1
Bdn51	Blanc de Noir	Germany	2021	Pinot Noir	0.75 L	Alsace	BVS	11.4	6.4	3.1	5.3	35.1	83.5	05/03/2022	1
Bdn52	Blanc de Noir	Germany	2021	Pinot Noir	0.75 L	Burgundy	BVS	12.6	3.1	3.2	6.5	41.0	67.1	05/03/2022	1
Bdn53	Blanc de Noir	Germany	2021	Pinot Meunier	0.75 L	Bordeaux	Synthetic cork	13.4	3.1	3.1	6.5	41.8	63.3	05/03/2022	1
Bdn54	Blanc de Noir	Germany	2021	Pinot Meunier	0.75 L	Bordeaux		13.0	1.3	3.3	5.8	42.0	95.1	05/03/2022	1
Bdn55	Blanc de Noir	Germany	2021	Pinot Meunier	0.75 L	Bordeaux	BVS	13.2	6.1	3.3	5.0	38.1	81.2	05/03/2022	1
Bdn56	Blanc de Noir	Germany	2021	Pinot Meunier	0.75 L	Burgundy	Natural cork	12.0	5.7	3.5	5.6	43.3	93.3	05/03/2022	1
Bdn57	Blanc de Noir	Germany	2021	Pinot Meunier	0.75 L	Alsace	Synthetic cork	11.4	7.4	3.5	6.5	34.8	90.9	05/03/2022	1

Bdn58	Blanc de Noir	Germany	2021	Pinot Meunier	0.75 L	Burgundy	BVS	11.3	6.7	3.3	5.2	34.6	94.8	05/03/2022	1
Bdn59	Blanc de Noir	Germany	2021	Pinot Meunier	0.75 L	Bordeaux	BVS	12.0	1.9	3.3	6.0	23.7	74.8	11/10/2022	1
Bdn60	Blanc de Noir	Germany	2021	Pinot Meunier	0.75 L	Burgundy	BVS	12.5	2.8	3.3	5.8	27.6	84.6	11/10/2022	1
Bdn61	Blanc de Noir	Germany	2021	Cabernet Sauvignon	0.75 L	Bordeaux	BVS	11.0	7.0	3.3	6.3	42.8	64.1	11/10/2022	1
Bdn62	Blanc de Noir	Germany	2021	Pinot Noir	0.75 L	Bordeaux	Synthetic cork	11.2	1.2	3.5	6.3	39.8	82.1	11/10/2022	1
Bdn63	Blanc de Noir	Germany	2021	Pinot Noir	0.75 L	Burgundy	BVS	13.5	5.0	3.5	6.2	43.4	64.4	11/10/2022	1
Bdn64	Blanc de Noir	Germany	2021	Pinot Noir	0.75 L	Bordeaux	BVS	13.5	4.3	3.1	5.1	21.3	60.5	11/10/2022	1
Bdn65	Blanc de Noir	Germany	2021	Pinot Noir	0.75 L	Bordeaux	BVS	12.7	1.2	3.1	6.2	39.7	97.1	11/10/2022	1
Bdn66	Blanc de Noir	Germany	2021	Pinot Noir	0.75 L	Burgundy	Agglomerated cork	12.4	4.1	3.1	5.6	30.5	84.1	11/10/2022	1
Bdn67	Blanc de Noir	Germany	2021	Pinot Noir	0.75 L	Alsace	BVS	13.6	2.3	3.2	5.5	24.4	75.0	11/10/2022	1
Bdn68	Blanc de Noir	Germany	2021	St. Laurent	0.75 L	Alsace	BVS	11.0	7.4	3.2	5.5	30.7	70.3	11/10/2022	1
Bdn69	Blanc de Noir	Germany	2021	Pinot Noir	0.75 L	Alsace	BVS	11.0	4.0	3.2	6.2	34.3	75.3	11/10/2022	1
Bdn70	Blanc de Noir	Germany	2021	Pinot Noir	0.75 L	Burgundy	Synthetic cork	12.3	2.9	3.1	6.5	26.5	71.1	11/10/2022	1
Bdn71	Blanc de Noir	Germany	2021	Pinot Noir	0.75 L	Burgundy	BVS	12.5	3.7	3.4	5.3	36.3	60.7	11/10/2022	1
Bdn72	Blanc de Noir	Germany	2021	Pinot Noir	0.75 L	Bordeaux	BVS	11.8	3.5	3.3	5.0	21.5	95.9	11/10/2022	1
Bdn73	Blanc de Noir	Germany	2021	Pinot Noir	0.75 L	Bordeaux	BVS	12.2	4.1	3.4	5.9	40.3	70.2	11/10/2022	1
Bdn74	Blanc de Noir	Germany	2021	Pinot Madeleine	0.75 L	Burgundy	BVS	13.9	3.2	3.2	6.6	20.6	80.4	11/10/2022	1
Bdn75	Blanc de Noir	Germany	2021	Pinot Noir	0.75 L	Bordeaux	BVS	11.2	4.9	3.3	6.4	36.3	61.5	11/10/2022	1
Bdn76	Blanc de Noir	Germany	2021	Pinot Noir	0.75 L	Burgundy	BVS	13.0	1.9	3.5	6.3	42.4	90.5	11/10/2022	1
Bdn77	Blanc de Noir	Germany	2021	Pinot Noir	0.75 L	Bordeaux	Natural cork	12.9	7.0	3.1	7.0	29.1	61.5	11/10/2022	1
Bdn78	Blanc de Noir	Germany	2021	Pinot Noir	0.75 L	Burgundy	Synthetic cork	11.1	3.0	3.3	5.2	36.1	72.6	11/10/2022	1
Bdn79	Blanc de Noir	Germany	2021	Pinot Noir	0.75 L	Burgundy	BVS	11.9	3.6	3.5	6.7	42.8	70.1	11/10/2022	1
Bdn80	Blanc de Noir	Germany	2021	Pinot Noir	0.75 L	Burgundy	BVS	12.5	4.0	3.3	6.9	28.4	70.5	11/10/2022	1
Bdn81	Blanc de Noir	Germany	2021	Pinot Noir	0.75 L	Bordeaux	BVS	13.4	4.8	3.4	6.9	21.2	94.4	11/10/2022	1
Bdn82	Blanc de Noir	Germany	2021	Pinot Noir	0.75 L	Bordeaux	BVS	13.5	2.8	3.1	5.8	22.3	78.9	11/10/2022	1
Bdn83	Blanc de Noir	Germany	2021	Pinot Noir	0.75 L	Bordeaux	BVS	12.9	5.9	3.2	5.5	34.4	96.0	11/10/2022	1

## 6 Additional relevant unpublished results

### 6.1 Blanc de Noir Check Web Application

In section 5.4 a web application was introduced to actually make the results from this study useful to a broader audience. This section focuses on the built and the functions of the app. By accessing the aforementioned website, the application starts by loading the dependencies of the app from the internet. An open source CSS stylesheet (Parmer 2024) controls the shape and appearance of the buttons and the needed Python packages are loaded from a virtual environment, running on an external server. The training dataset and a plotting reference dataset is stored in a GitHub repository (Hensel 2024). While the training dataset only contains the target id (0: Not Blanc de Noir, 1: Blanc de Noir), the plotting reference dataset subdivides the target to rosé, white wine, and Blanc de Noir. This is necessary as the reference data points of rosé wines and white wines are depicted in the two-dimensional and three-dimensional projections. The app proceeds to plot boxplots from the training dataset to show the general distribution of the two classes. From there, both datasets are scaled by the aforementioned RobustScaler and split into training and test data. The in section 5.4 described hyperparameters were used to fit the SVM to achieve the same output as described. After fitting, the two-dimensional projections ( $L^*a^*$ -,  $L^*b^*$ -, and  $a^*b^*$ -projection) are plotted. The App has loaded completely and is now ready for user interaction.

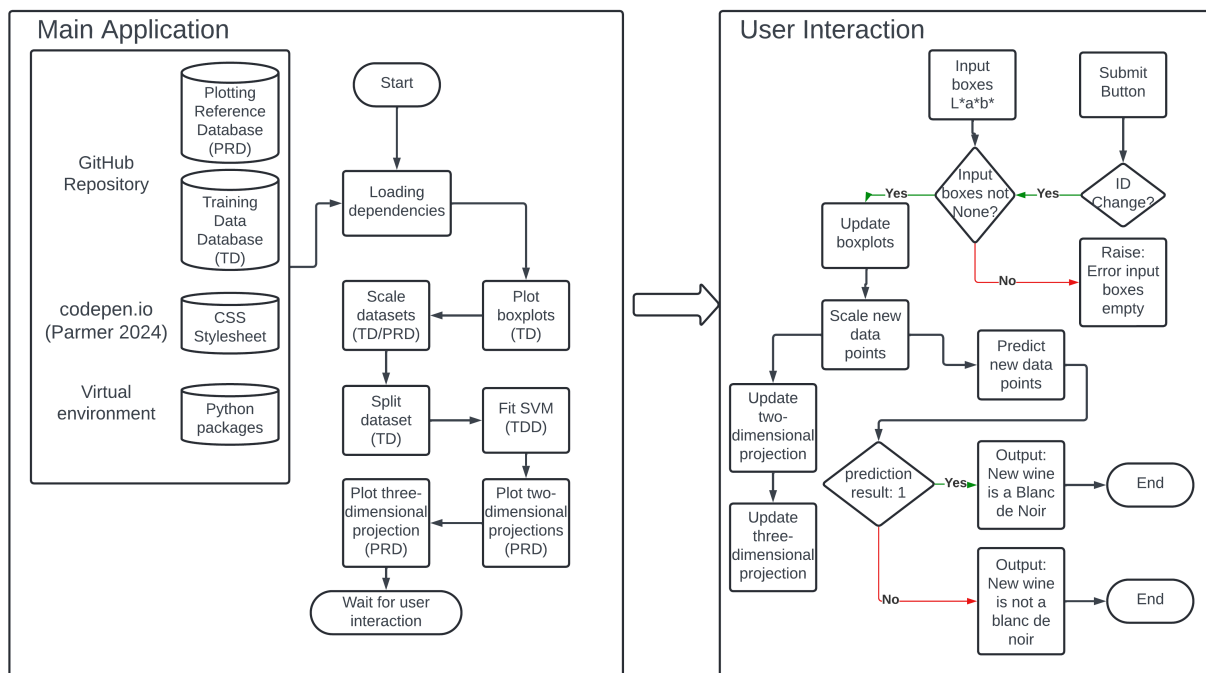


Figure 37: Flowchart that describes the workflow of the web application.

The user has three input boxes at the top of the app. By inputting the measured CIE  $L^*a^*b^*$  coordinates and clicking the "Submit" button, the app checks if the input boxes are filled with valid data. If not, the app raises an error, stating that the boxes are empty or not readable (Figure 37, Figure 38a). If the condition is met, the new data points are added to the boxplots. The data is now scaled under the same conditions as the dataset prior and also added as red dots to the two dimensional and three dimensional projections. Afterwards the trained model is used to predict the wine style of the new wine. Depending

on the result, the algorithm raises an output message stating whether the new wine is considered a Blanc de Noir (Figure 37, Figure 38b, Figure 38c). A complete framework is shown in Appendix I. Due to the Dash API (Plotly Technologies Inc. 2015), it is possible to interact with the graphs. It is possible to zoom into certain areas of interest, change the labels of the legend and titles, and scale the axis to a desired outcome. By hovering the cursor of the mouse over a data point, the app provides additional info about the data point. The data points can be isolated, filtered, and the graphs can be exported as picture files for quality control documentation.

## Welcome to Blanc de Noir check

Please insert the CIE L\*a\*b\* parameters in the boxes below

Error 404: CIE L\*a\*b\* coordinates not found or readable. Please insert your CIE L\*a\*b\* coordinates

(a) Error message if no input boxes are empty or not properly filled.

Please insert the CIE L\*a\*b\* parameters in the boxes below

this wine is most likely a Blanc de Noir.

(b) Output message if the prediction of the new wine is positive.

Please insert the CIE L\*a\*b\* parameters in the boxes below

this wine is most likely NOT a Blanc de Noir.

(c) Output message if the prediction of the new wine is negative.

Figure 38: Output messages of the web application depending on the user interaction and the prediction result of the SVM

## 6.2 The Classification of Red wines

Differentiating Blanc de Noir wines from other visually similar wine styles is a rather simple target, since the difference is visible to some extent. The developed application is also very niche and specifically developed for the German wine industry. To evaluate the limitations of spectrophotometric data combined with machine learning, a study was carried out investigating if it is possible to classify barrel aged red wines. Due to barrel aging, the oxygen uptake of the wine is increased in comparison to steel tank fermentation and therefore oxidation occurs more often and more intense, leading to changes of the anthocyanin profile, the spectrophotometric properties and more precise a difference in the b\*-coordinate (Jackson 2008; Ribéreau-Gayon 2006; Pfahl et al. 2021; Durner et al. 2010). Thus, it was hypothesized that the alteration of the spectrophotometric properties are sufficient to classify the wines accordingly. Possible application for users could be quality control to investigate if a new barrel aged wine is coherent with the typical appearance of barrel-aged wine. Another application would be for food control entities as an early detection system that provides a cost-effective evaluation of a wine. To investigate this, the algorithm presented in section 5.4 is used to classify red wines. Since the spectra differ so much between Blanc de Noir, rosé wine, white wine and red wine, transmission spectra, absorption spectra, and CIE L\*a\*b\* coordinates were used to as input in order to evaluate, which of the data types were most suited

for red wine classification. In order to improve classification, absorption spectra were used to train a neural network.

### 6.2.1 Samples and photometric measurement

For this study a total of 366 wines from the grape varieties Cabernet Sauvignon (112), Tempranillo (86), Merlot (62), Primitivo (53), and Sangiovese (53). The wines were measured on-site during the Mundus Vini Summer Tasting 2022 in Neustadt a.d.W. Transmission spectra were recorded with a StellarRad<sup>TM</sup> (StellarNet Inc, FA, USA) handheld colorimeter with a dip probe. The dip probe was equipped with a 2 mm pathlength tip. The transmission spectra were recorded between 200 and 1100 nm. Afterwards the spectra were referred to a 10 mm pathlength and the CIE L\*a\*b\* coordinates were calculated between 360 and 830 nm. Furthermore, the recorded transmission spectra were transformed into absorption spectra. The wines were separated into two classes: barrel aged (Wood), and Not barrel aged (INOX). The barrel aged class consisted of wines from all different kinds of barrels like barrique, tonneau, Halbstück etc. The other class consisted of wines aged in steel tanks (INOX). A complete breakdown of the used wines is shown in Appendix II.

### 6.2.2 Support Vector Machine Modeling

Figure 39 depicts the computation of the initial transmission spectra. The transmission spectra were transformed to absorption spectra and CIE L\*a\*b\* coordinates, using the D65 standard illuminant and the 10° standard observer. All datasets were preprocessed with the RobustScaler, where the median and the interquartile range was used to scale the data. The dataset was split into 70% training and 30% test data. The training data was used to optimize the SVM by randomized search and the test data was used to validate the optimized model. The following hyper parameters were optimized in 150 iterations: Penalty of misclassification (C) between 1 and 50 and the parameter  $\gamma$  between  $10^{-5}$  and 1 on a log-scale. The used kernel is the radial basis function (RBF) since this was the best suited kernel in section 5.4. The performance of the parameter optimized model was evaluated with the accuracy, recall, precision as well as the ROC curves.

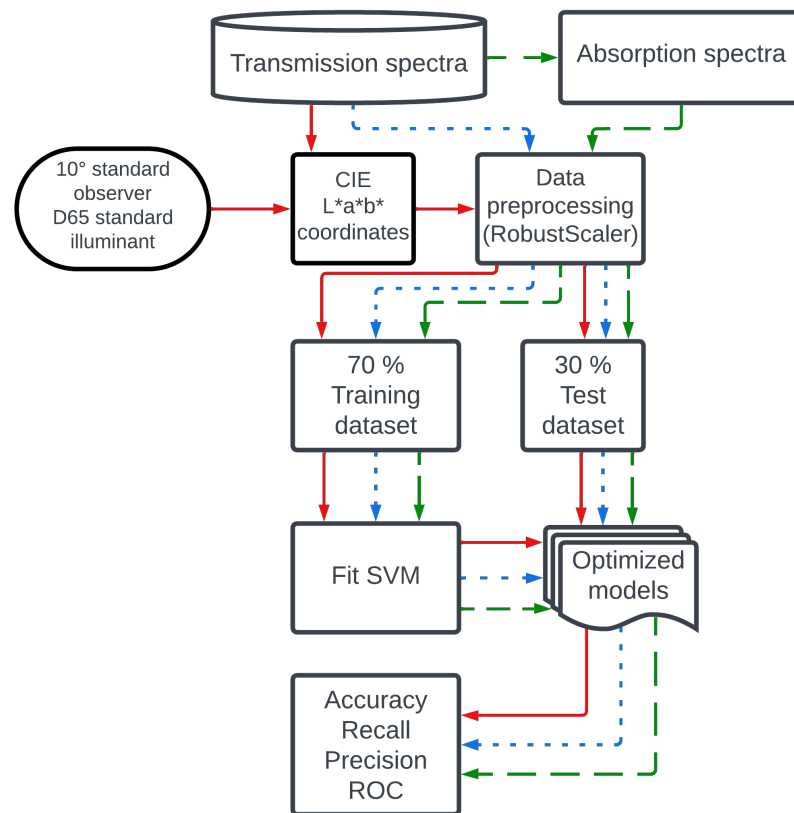


Figure 39: Data pipeline starting from the transmission spectra as raw data through the different computation steps. Color indicates different spectrophotometric input data. All input datasets were used to fit and optimize an SVM. The optimized models were then evaluated with performance parameters.

### 6.2.3 Neural Network architecture

Figure 40 shows the transformation of the dataset and the following deep learning process. CIE  $L^*a^*b^*$  coordinates are not suited for this type of model, because of their limited amount of features. The initial spectrophotometric data is preprocessed using the RobustScaler (Pedregosa F et al. 2011). To counter co-linearity and the risk of overfitting, dimensionality reduction with principle component analysis (PCA) was used. The initial scaled dataset with 471 dimensions is reduced to 100 principle components, which then were used to feed the neural network or multilayer perceptron model (MLP) using Keras (Chollet et al. 2015) SciKeras (Badaracco 2024), as well as TensorFlow (TensorFlow Developers 2024). An initial input layer with corresponding dimensions was equipped with an L2 regularization term. Both input and hidden layer contain dropout nodes to furthermore prevent overfitting. For the input and hidden layer the ReLU activation function was used, whereas the output layer was activated with a sigmoidal function. The model was compiled with the binary crossentropy loss function and the Adam optimizer Kingma and Ba 2014. Accuracy was the used metric for monitoring the performance of the model during training. The model was fitted with 200 epochs, where 10% of the data was used for validation and an early stopping callback that monitors the validation loss. To measure the performance of the model the ROC curves were evaluated as well as the typical performance parameters recall, precision, and F1-score.

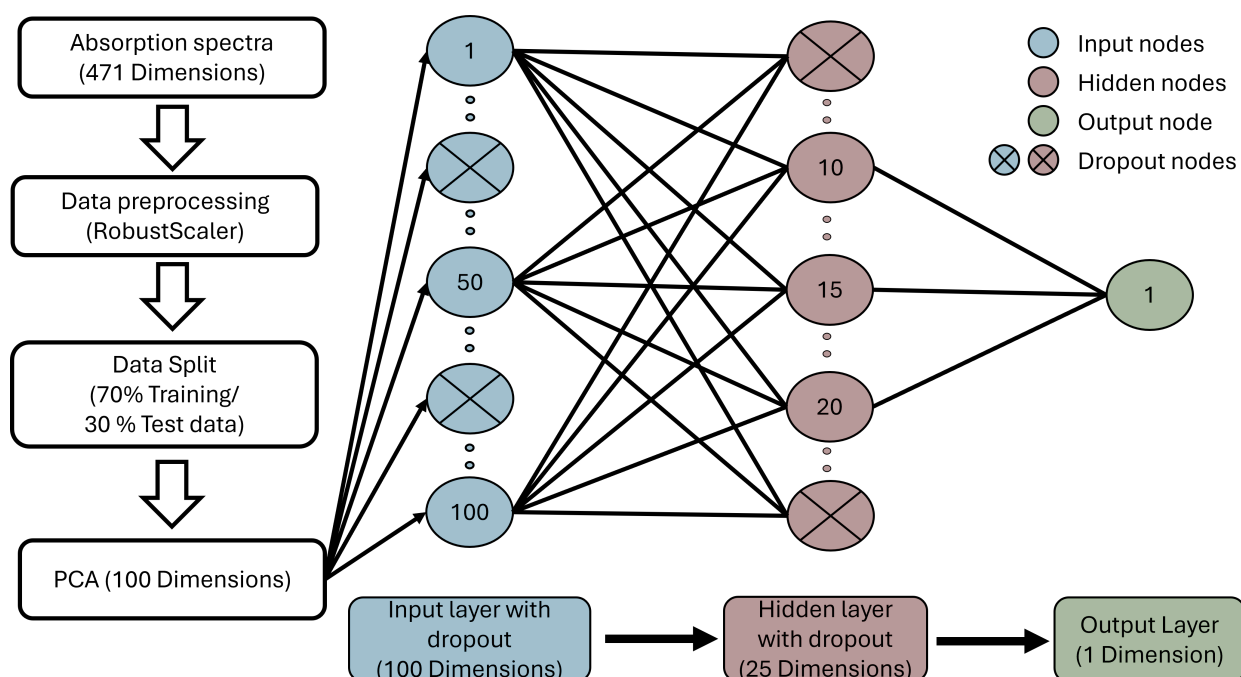


Figure 40: Depiction of the data processing starting by scaling the data and using PCA for dimensionality reduction. The principle components of the PCA are then used as input for the MLP model with one hidden layer.

#### 6.2.4 Results and Discussion

To evaluate the the performance of the models and the input datatype, the ROC curves were used (Figure 41). This tool plots the false positive rate (FPR), or false alarm rate, against the true positive rate (TPR), or sensitivity. The diagonal line represents the random classifier (AUC=0.5). Figure41A shows the test dataset. SVM on transmission spectra yields the lowest performance of the models with AUC = 0.60. SVM on absorption spectra (AUC = 0.71) and SVM on CIE L\*a\*b\* coordinates (AUC = 0.71) show similar performance. The best performance in the test dataset is achieved by MLP on absorption spectra.

Evaluating the training data, the performance is generally higher. Here, the lowest performance is achieved by MLP on absorption spectra (AUC = 0.85) . SVM on CIE L\*a\*b\* coordinates (AUC = 0.89) and SVM on absorption spectra (AUC = 0.90) yield similar performance. SVM on transmission spectra yields a perfect classification model in the training dataset. However, comparing the results of the training datasets with the results of the test dataset, severe overfitting is visible in this model. All SVM models indicate signs of overfitting, but not with the severity as SVM on transmission spectra. The best suited model to classify barrel aged red wine is MLP on absorption spectra, where the difference between training dataset and test dataset is under 10%.

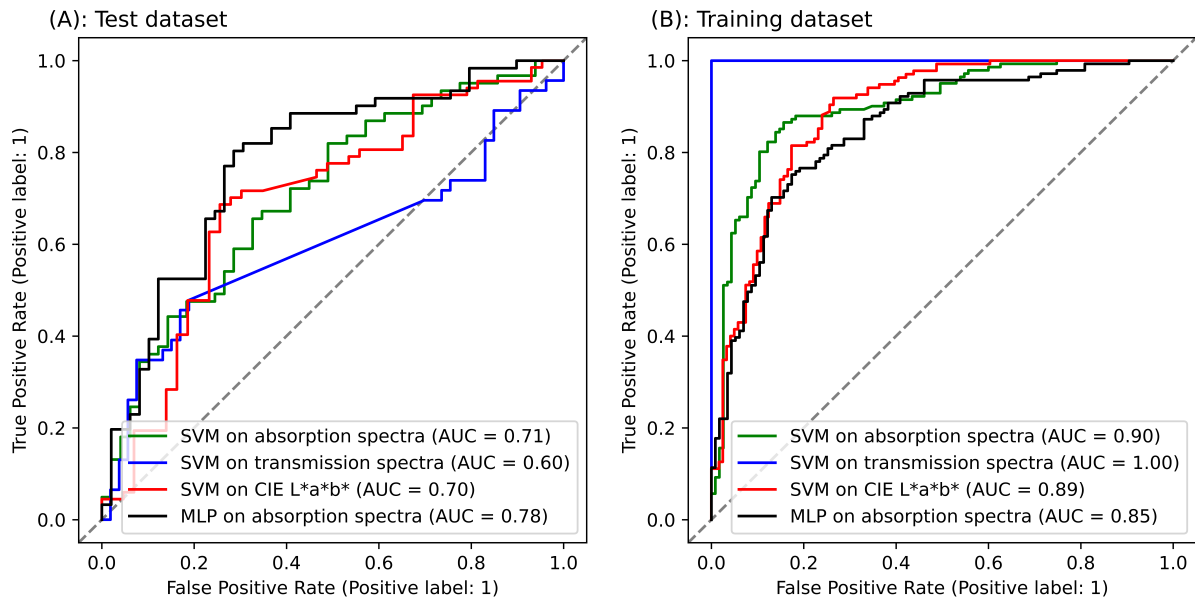


Figure 41: Depiction of the ROC and for the test dataset (A) and training dataset (B).

Table 6 shows the performance parameter precision, recall, and F1-score for the training and test dataset for the SVM and MLP models. Precision represents a parameter that measures the accuracy of positive prediction. A high precision model is making few false positive predictions. Recall measures the performance of a model to find all positive instances. A model with high recall is good at finding all relevant positive cases. The harmonic mean between precision and recall is the F1-score. A high F1-score corresponds to a model with high precision and recall. Table 6 shows that in general the performance parameters are on the same scale of different magnitude. In the training dataset SVM on transmission spectra shows the highest performance, which plummets during model validation, substantiating the conclusion that severe overfitting occurs. SVM on absorption spectra also shows overfitting, because precision, recall, and F1-score in the training dataset are over 20% higher than in the test dataset. The difference between training and test dataset is not as severe with SVM on CIE L\*a\*b\*, where the precision parameters are just 10% higher in the training dataset. MLP on absorption spectra show the lowest amount of overfitting, with under 10% difference between training and test dataset.

Table 6: Performance parameters precision, recall, and F1-score of SVM on absorption spectra, transmission spectra, and CIE L\*a\*b\* coordinates as well as MLP on absorption spectra for training and test dataset.

Model	Data input	Training data			Test data		
		Precision	Recall	F1-score	Precision	Recall	F1-score
SVM	Absorption spectra	0.84	0.82	0.82	0.69	0.66	0.67
	Transmission spectra	1	1	1	0.48	0.49	0.38
	CIE L*a*b*	0.81	0.81	0.81	0.70	0.71	0.70
MLP	Absorption spectra	0.77	0.77	0.77	0.75	0.74	0.74

The reason for the inferior performance of SVM on transmission spectra is of chemical nature. In Figure 42 the scaled mean absorption spectra (A) and the scaled mean transmission spectra (B) are shown with their respective standard deviation. These are the spectra on which the models are trained.



The area of greatest variance is between 400 and 550 nm in Figure 42A and between 650 to 800 nm in Figure 42B. The absorbance maxima of the coloring agents in wine are between 400 and 550 nm. Therefore the area of highest variance in transmission spectra is in an area where limited molecular information is available in wine. In general, areas of high variance have a much higher impact than areas of low variance. Since transmission values are on a relative scale, the differences in certain wavelength areas can be skewed. By transforming the transmission spectra to absorption spectra the relative scale is transformed to absolute values and therefore to a more precise depiction of the variance. By transforming the transmission spectra to absorption spectra the areas of high variance shift to different wavelength areas leading the models to weight the input data differently. High variance in transmission spectra is caused by lack of molecular information and therefore SVM on transmission spectra is rather trained on noise than molecular information. SVM on absorption spectra on the other hand is trained on the molecular information stored in the spectra, so the logic the SVM follows also aligns with the chemical logic behind absorption spectra.

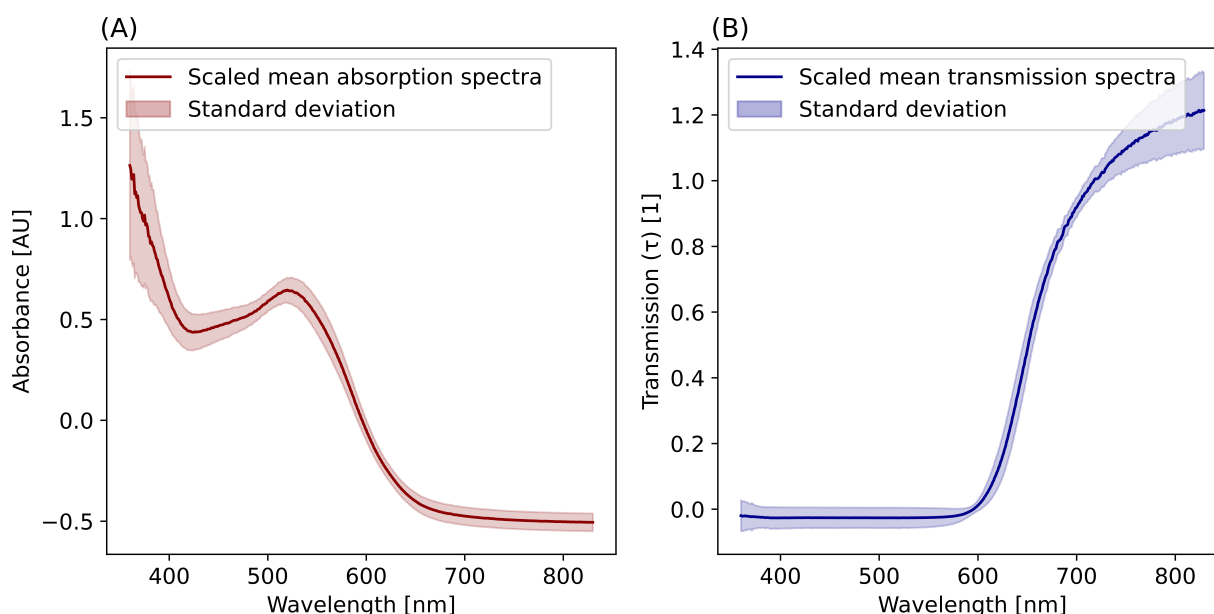


Figure 42: Depiction of the mean absorption spectra (A) and the mean transmission spectra (B) in the dataset ( $n=366$ ). The unsaturated areas are the standard deviation.

To summarize this study, the algorithm presented in section 5.4 can be used for red wine classification as well with moderate success. CIE  $L^*a^*b^*$  coordinates as input dataset is better suited than absorption and transmission spectra. However, using absorption spectra to train an MLP outperforms SVM on CIE  $L^*a^*b^*$ . Still, the performance of the used models are not good enough yet to build an application that can be used by wineries, food control entities etc. More work has to be done, expanding the dataset or combine different parameters like FTIR and UV/Vis spectra.

## 7 Concluding remarks

Wine is one of the most consummated alcoholic beverages and it is not only seen as a consumer good, but also has a cultural component. Quality wine is celebrated all over the world and the sensory impressions when drinking wine are manifold. Not only the taste and the smell are decisive, but also the appearance, color, texture, and viscosity play a decisive role, contributing to the wine experience. The color of a wine in particular is known to create expectations about the aroma and the taste of a wine. It is furthermore known to influence purchase decisions, typicality and quality ratings. In wineries, the wine color is used as an internal quality assessment parameter. In wine research it is also a control parameter to enhance the quality of sensory color evaluation ratings.

The first part was to elaborate the two different methods to measure the wine color. For this, it was demonstrated that the photometer settings data interval and scan speed have a significant impact on the reproducibility of the CIE  $L^*a^*b^*$  color space. The photometer settings of 1 nm data interval and 1000 nm/min yielded the most reliable results. In a next step the color measurement according to Glories and the CIE  $L^*a^*b^*$  color space were compared. The results showed a strong correlation between the two methods in dark red wine with  $L^* \leq 20$ . In light red wines however, the correlation of the CIE hue parameters ( $a^*$  and  $b^*$ ) and Glories' color tint parameters *hue* and *Brilliance* was very weak. The same applies to the correlation in white wine. Furthermore, the results of the color measurement according to Glories in white wine indicate that a violation of the first Grassmann law as one of the three parameters (OD620) could not be obtained from white wine. Grassmann's law states here, that in order to measure the color, three independent variables are needed. Furthermore, Glories absorbance values correlated with themselves in all three wine categories, which indicates a dependency between the parameters in violation of Grassmann's first law. The CIE  $L^*a^*b^*$  color space is potentially more reliable for color calculation in wine. However, the parameters also correlate with each other in dark red wine, indicating a dependence and a violation of Grassman's law. Initially, a dataset of 56 red wines and 56 white wines was used. However, to verify the results of the experiments a more thorough evaluation with 617 dark red wines and 308 light red wines was performed. The results of this validation experiment were coherent with the previous performed work.

Since both methods cannot be interchangeably used for wine color, it was imperative to evaluate, which of the methods depicts the "correct" color. For this, the results of the photometric color measurement was compared to the results of a descriptive sensory evaluation. The study extends knowledge from previous publications using a wide range of commercial wines from dark red to light red color as well as from yellow to orange hue. The premise of the CIE color spaces is that it approximates the human color vision. This study showed, that not every CIE coordinate is able to depict the perceived color and not every Glories parameter is unable to do it. Glories' CI and the  $L^*$  coordinate show a similar performance regarding the comparison with the perceived color. However, the hue parameters according to Glories do not resemble the perceived color of a wine. Contrarily to the definition, the  $a^*$  coordinate (red component) is also not very well suited to depict the perceived color, leading to the conclusion that humans distinguish between red wines of different hues regardless of the  $a^*$  value. Accordingly, it is searched for deviations apart from the red color. The best alignment with the perceived color had the  $b^*$  coordinate, followed by  $C^*$  and  $h^0$ . The CIE  $L^*a^*b^*$  color space in total had a higher alignment with the perceived color for red wines. However, there are some discrepancies in the high saturated light red color area. There, the CIE  $L^*a^*b^*$  color space has problems to represent the perceived color. There are

similarities between red and white wines, because Glories' CI and  $L^*$  could both represent the perceived lightness. Also,  $b^*$  was the CIE coordinate that approximated the perceived hue best. However, the  $h^0$  coordinate had no resolution for white wine as it had for red wine. In this evaluation, the performance of these parameters was evaluated by their ability to differentiate grape varieties. In another approach a wine to wine comparison was evaluated using a regression analysis in the different CIE color planes. Here, the experiment yielded similar results. While the regression of CI,  $L^*$ , and the perceived lightness were similar, this was not the case for the hue parameters. The discrepancies visible in the original approach can be substantiated here. On a wine to wine comparison level, the hue parameters of the CIE  $L^*a^*b^*$  color space follow the general linear appearance of the regression of the perceived color. However, in the light red area, the analytical color and perceived color disperse, leading to an irregularity in the high saturated light red color area. Despite this, the fit of the CIE  $L^*a^*b^*$  coordinates show a better coherence with the perceived color than the fit of the Glories parameters, which were unable to match the perceived color.

In order to make future assumptions about the visual difference between two wines of different colors, the JND's for wine color were re-evaluated with the currently recommended CIEDE2000 standard. With the CIEDE2000 color distance formula the uniformity is increased, but the color space is still not completely equidistant and further studies should investigate the smallest visible color threshold throughout the complete color space. However, until further work has been carried out, it is advised to use the CIEDE2000 color distance formula and the presented visible color thresholds as guidelines for future studies.

The results of the two studies have led to the conclusion that the most suited color measurement method is the CIE  $L^*a^*b^*$  color space. The next step was to find new applications for color measurement. For this, the applicability of logistic regression, SVM, and XGBoost were assessed based on spectrophotometric data. The results show that the CIE  $L^*a^*b^*$  coordinates can be used for ML classification via SVM or logistic regression. The performance of the algorithm trained with CIE  $L^*a^*b^*$  was similar to the algorithm trained with the conventional spectrophotometric data. The classification of Blanc de Noir wines proved not only that ML assisted classification is possible with wine color data, the transfer to a browser based application shows that wine color data can be possibly used to build powerful applications that are still easy to understand and operate for untrained personnel.

The classification of barrel-aged red wine is possible to a certain extend. The performance of the ML algorithms is not as good as in the classification of Blanc de noir. Future studies should use additional parameters to obtain better results.

To summarize this work, the protocols regarding the photometric color measurement of the CIE color spaces could be optimized. Furthermore, the method according to Glories and the CIE  $L^*a^*b^*$  color space were compared to each other and a few discrepancies were discovered, mainly in the light red and white wine color area. Both methods were compared to the human color perception and the JND regarding wine color was re-evaluated with the CIEDE2000 color difference formula. The CIE  $L^*a^*b^*$  color space was most suited for further investigation and was used to establish a new application for wine color: Wine classification. An algorithm was utilized that incorporated wine color data to differentiate Blanc de Noir from rosé and white wine and in order to make this algorithm accessible a browser application was developed. This enables untrained personnel like lab technicians, winemakers etc. to use the algorithm. In future studies these results can be utilized for other new targets for example red wine classification. Furthermore, future studies should also evaluate the applicability of wine color data in combination

with ML-assisted regression analysis like Support Vector Regression and XGBoost with the squared loss function approach.

There are still some open questions about the wine color measurement. For example, "data interval", and "scan speed" are not the only parameters that can be set by the user in photometers. There are others like "response time" or "bandwidth". It is still unknown if these parameters influence the wine color measurement. Regarding the correlation between the CIE L\*a\*b\* color space and Glories' color measurement, in dark red wine both methods correlate with the other parameters of the same system, like L\* with a\*, and b\* as well as CI and T. This can be explained by the regression approach with the wine-to-wine comparison. Here, the analytical color indicates a dependence between lightness and hue in low lightness areas as the regression between L\* and b\*, and L\* and a\* is roughly linear in those areas, leading to this visible dependence in the correlation. This leads to another problem, CIE L\*a\*b\* is designed to comply to Grassman's law, where the color must be described by three independent parameters, which is not the case in this study. The reason for this dependence when there should be independence cannot yet be explained. Future studies could evaluate if this dependence derives from the wine making process or the color measurement.

One recurring pattern was that light red wine stood out throughout this work. Glories color measurement and CIE L\*a\*b\* correlate well in dark red wine, but the worst correlation could be observed in light red wine. In light red wine, the perceived color and the measured color deviate from each other, and the JND is the highest in light red wine. It is unknown why light red wine appears to be the odd one, which can be addressed in future studies.

Regarding the use of ML with spectrophotometric data. Another approach that could be part of future studies is the reasoning of these algorithms. What leads them to their results and how does this differ from the human evaluation process. Do human and algorithm depend the result on the same reasons or are they completely different?

## 8 References

- Alcalde-Eon, Cristina, María Teresa Escribano-Bailón, et al. (2006). "Changes in the detailed pigment composition of red wine during maturity and ageing". In: *Analytica Chimica Acta* 563.1-2, pp. 238–254. ISSN: 00032670. DOI: 10.1016/j.aca.2005.11.028.
- Alcalde-Eon, Cristina, Rebeca Ferreras-Charro, et al. (2019). "Monitoring the effects and side-effects on wine colour and flavonoid composition of the combined post-fermentative additions of seeds and mannoproteins". In: *Food Research International* 126, p. 108650. ISSN: 0963-9969. DOI: <https://doi.org/10.1016/j.foodres.2019.108650>. URL: <https://www.sciencedirect.com/science/article/pii/S0963996919305368>.
- Asenstorfer, Robert E., Yoji Hayasaka, and Graham P. Jones (2001). "Isolation and Structures of Oligomeric Wine Pigments by Bisulfite-Mediated Ion-Exchange Chromatography". In: *Journal of Agricultural and Food Chemistry* 49.12, pp. 5957–5963. ISSN: 0021-8561. DOI: 10.1021/jf011261a.
- Asenstorfer, Robert E., David F. Lee, and Graham P. Jones (2006). "Influence of structure on the ionisation constants of anthocyanin and anthocyanin-like wine pigments". In: *Analytica Chimica Acta* 563.1-2, pp. 10–14. ISSN: 00032670. DOI: 10.1016/j.aca.2005.09.040.
- Atkins, Peter W., Julio de Paula, and James J. Keeler (2022). *Physikalische Chemie*. 6. Auflage. Weinheim: WILEY-VCH. ISBN: 9783527828333. URL: <https://resolver.vitalsource.com/9783527828326>.

- Badaracco, Adrian Garcia (2024). *Welcome to SciKeras's documentation! — SciKeras 0.13.0 documentation*. URL: <https://adriangb.com/scikeras/stable/index.html>.
- Ballester, Jordi et al. (2009). "The Odor of Colors: Can Wine Experts and Novices Distinguish the Odors of White, Red, and Rosé Wines?" In: *Chemosensory Perception* 2.4, pp. 203–213. ISSN: 1936-5802. DOI: 10.1007/s12078-009-9058-0.
- Bate-Smith, E. C. (1954). "Leuco-anthocyanins. 1. Detection and identification of anthocyanidins formed leuco-anthocyanins in plant tissues". In: *The Biochemical journal* 58.1, pp. 122–125. ISSN: 0264-6021. DOI: 10.1042/bj0580122.
- Belitz, Hans-Dieter, Werner Grosch, and Peter Schieberle (2001). *Lehrbuch der Lebensmittelchemie*. Fünfte, vollständig überarbeitete Auflage. Springer eBook Collection Life Science and Basic Disciplines. Berlin, Heidelberg and s.l.: Springer Berlin Heidelberg. ISBN: 978-3-662-08303-1. DOI: 10.1007/978-3-662-08302-4.
- Bennett, Kristin P. and Colin Campbell (2000). "Support vector machines". In: *ACM SIGKDD Explorations Newsletter* 2.2, pp. 1–13. ISSN: 1931-0145. DOI: 10.1145/380995.380999.
- Bhardwaj, Piyush et al. (2022). "A machine learning application in wine quality prediction". In: *Machine Learning with Applications* 8, p. 100261. ISSN: 26668270. DOI: 10.1016/j.mlwa.2022.100261. URL: <https://www.sciencedirect.com/science/article/pii/S266682702200007X>.
- Bhowmik, Pritom (2019). "Research Study on basic Understanding of Artificial Neural Networks". In: *Global journal of computer science and technology*. URL: <https://api.semanticscholar.org/CorpusID:214277491>.
- Boser, Bernhard E., Isabelle M. Guyon, and Vladimir N. Vapnik (1992). "A Training Algorithm for Optimal Margin Classifiers". In: *Proceedings of the Fifth Annual Workshop on Computational Learning Theory*. Ed. by ACM. COLT '92. New York, NY, USA: Association for Computing Machinery, pp. 144–152. ISBN: 089791497X. DOI: 10.1145/130385.130401.
- Boulton, Roger (2001). "The Copigmentation of Anthocyanins and Its Role in the Color of Red Wine: A Critical Review". In: *American Journal of Enology and Viticulture* 52.2, pp. 67–87. ISSN: 0002-9254. DOI: 10.5344/ajev.2001.52.2.67.
- Brochet, F. and D. Dubourdieu (2001). "Wine descriptive language supports cognitive specificity of chemical senses". In: *Brain and Language* 77.2, pp. 187–196. ISSN: 0093-934X. DOI: 10.1006/brln.2000.2428. URL: <https://www.sciencedirect.com/science/article/pii/S0093934X00924283>.
- Brossaud, F., Veronique Cheynier, and A. C.N.N. Noble (2001). "Bitterness and astringency of grape and wine polyphenols". In: *Australian Journal of Grape and Wine Research* 7.1, pp. 33–39. ISSN: 1322-7130. DOI: 10.1111/j.1755-0238.2001.tb00191.x.
- Bühler, Peter, Patrick Schlaich, and Dominik Sinner (2018). *Digitale Farbe: Farbgestaltung – Color-management – Farbverarbeitung*. 1. Aufl. 2018. Bibliothek der Mediengestaltung. Berlin, Heidelberg: Springer Berlin Heidelberg. ISBN: 9783662546079. URL: <http://nbn-resolving.org/urn:nbn:de:bsz:31-epflicht-1520973>.
- Bührle, Franziska, Anita Gohl, and Fabian Weber (2017). "Impact of Xanthylum Derivatives on the Color of White Wine". In: *Molecules* 22.8, p. 1376. ISSN: 1420-3049. DOI: 10.3390/molecules22081376. URL: <https://www.mdpi.com/1420-3049/22/8/1376>.
- Bundesrat Drucksache 175/21 (2021). *Änderungen und Entschließung zur Vierundzwanzigsten Verordnung zur Änderung der Weinverordnung*.

- Chen, Tianqi and Carlos Guestrin (2016). "XGBoost". In: *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. Ed. by Balaji Krishnapuram. ACM Digital Library. New York, NY: ACM, pp. 785–794. ISBN: 9781450342322. DOI: 10.1145/2939672.2939785.
- Chollet, François et al. (2015). *Keras*. <https://keras.io>.
- CIE (2019a). *Farbmetrik - Teil 1: CIE farbmetrische Normalbeobachter (ISO/CIE 11664-1:2019)*. Berlin. DOI: 10.31030/3091836.
- (2019b). *Farbmetrik - Teil 3: CIE tristimulus values (ISO/CIE 11664-3:2019)*. DOI: 10.25039/DS11664-3.2019.
- (2019c). *Farbmetrik - Teil 4: CIE 1976 L\*a\*b\* Farbraum (ISO/CIE 11664-4:2019)*; Berlin. DOI: 10.31030/3092072.
- (2021). *Farbmetrik - Teil 2: CIE Normlichtarten (ISO/CIE DIS 11664-2.2:2021)*. Berlin.
- Clarke, F. J. J., R. McDonald, and B. Rigg (1984). "Modification to the JPC79 Colour-difference Formula". In: *Journal of the Society of Dyers and Colourists* 100.4, pp. 128–132. ISSN: 0037-9859. DOI: 10.1111/j.1478-4408.1984.tb00969.x.
- Coulon-Leroy, Cécile et al. (2018). "Is the typicality of "Provence Rosé wines" only a matter of color?" In: *OENO One* 52.4. DOI: 10.20870/oenone.2018.52.4.2125.
- Czerny, M. and A. F. Turner (1930). "ber den Astigmatismus bei Spiegelspektrometern". In: *Zeitschrift für Physik* 61.11-12, pp. 792–797. DOI: 10.1007/BF01340206. URL: <https://link.springer.com/article/10.1007/BF01340206#article-info>.
- Czibulya, Zsuzsanna et al. (2012). "Unexpected effect of potassium ions on the copigmentation in red wines". In: *Food Research International* 45.1, pp. 272–276. ISSN: 0963-9969. DOI: 10.1016/j.foodres.2011.10.040. URL: <https://www.sciencedirect.com/science/article/pii/S0963996911006211>.
- Danilewicz, John C. (2003). "Review of Reaction Mechanisms of Oxygen and Proposed Intermediate Reduction Products in Wine: Central Role of Iron and Copper". In: *American Journal of Enology and Viticulture* 54.2, pp. 73–85. ISSN: 0002-9254. DOI: 10.5344/ajev.2003.54.2.73.
- Darias-Martin, Jacinto et al. (2001). "Enhancement of red wine colour by pre-fermentation addition of copigments". In: *Food Chemistry* 73.2, pp. 217–220. ISSN: 0308-8146. DOI: 10.1016/S0308-8146(00)00286-7. URL: <https://www.sciencedirect.com/science/article/pii/S0308814600002867>.
- Durner, Dominik et al. (2010). "Sensory and Color Changes Induced by Microoxygenation Treatments of Pinot noir before and after Malolactic Fermentation". In: *American Journal of Enology and Viticulture* 61.4, pp. 474–485. ISSN: 0002-9254. DOI: 10.5344/ajev.2010.09122.
- Eiro, Maarit J. and Marina Heinonen (2002). "Anthocyanin color behavior and stability during storage: effect of intermolecular copigmentation". In: *Journal of agricultural and food chemistry* 50.25, pp. 7461–7466. ISSN: 0021-8561. DOI: 10.1021/jf0258306.
- Fairchild, Mark (2018). "The colors of wine". In: *International Journal of Wine Research* Volume 10, pp. 13–31. DOI: 10.2147/IJWR.S161891.
- Fenton, H. J. H. (1894). "LXXIII.—Oxidation of tartaric acid in presence of iron". In: *J. Chem. Soc., Trans.* 65.0, pp. 899–910. ISSN: 0368-1645. DOI: 10.1039/CT8946500899.
- Friedman, Jerome, Trevor Hastie, and Robert Tibshirani (2000). "Additive logistic regression: a statistical view of boosting (With discussion and a rejoinder by the authors)". In: *The Annals of Statistics* 28.2. ISSN: 0090-5364. DOI: 10.1214/aos/1016218223.

- fulcrand, Hélène et al. (1996). "Study of the acetaldehyde induced polymerisation of flavan-3-ols by liquid chromatography-ion spray mass spectrometry". In: *Journal of Chromatography A* 752.1-2, pp. 85–91. ISSN: 00219673. DOI: 10.1016/S0021-9673(96)00485-2.
- Gao, L. et al. (1997). "Changes in Anthocyanins and Color Characteristics of Pinot Noir Wines during Different Vinification Processes". In: *Journal of Agricultural and Food Chemistry* 45.6, pp. 2003–2008. ISSN: 0021-8561. DOI: 10.1021/jf960836e.
- García, Pedro A. et al. (2007). "Measurement of the relationship between perceived and computed color differences". In: *Journal of the Optical Society of America. A, Optics, image science, and vision* 24.7, pp. 1823–1829. ISSN: 1084-7529. DOI: 10.1364/josaa.24.001823.
- García-Beneytez, Eva, Félix Cabello, and Eugenio Revilla (2003). "Analysis of Grape and Wine Anthocyanins by HPLC-MS". In: *Journal of Agricultural and Food Chemistry* 51.19, pp. 5622–5629. ISSN: 0021-8561. DOI: 10.1021/jf0302207.
- Gey, Manfred (2015). *Instrumentelle Analytik und Bioanalytik: Biosubstanzen, Trennmethode, Strukturanalytik, Applikationen*. 3. Auflage. Springer-Lehrbuch. Berlin and Heidelberg: Springer Spektrum. ISBN: 978-3-662-46255-3. DOI: 10.1007/978-3-622-46255-3.
- Girard, Benoit et al. (1997). "Influence of Vinification Treatments on Aroma Constituents and Sensory Descriptors of Pinot noir Wines". In: *American Journal of Enology and Viticulture* 48.2, pp. 198–206. ISSN: 0002-9254. DOI: 10.5344/ajev.1997.48.2.198.
- Giusti, M. M., L. E. Rodríguez-Saona, and R. E. Wrolstad (1999). "Molar absorptivity and color characteristics of acylated and non-acylated pelargonidin-based anthocyanins". In: *Journal of Agricultural and Food Chemistry* 47.11, pp. 4631–4637. ISSN: 0021-8561. DOI: 10.1021/jf981271k.
- Gómez-Míguez, Manuela et al. (2006). "Influence of different phenolic copigments on the color of malvidin 3-glucoside". In: *Journal of Agricultural and Food Chemistry* 54.15, pp. 5422–5429. ISSN: 0021-8561. DOI: 10.1021/jf0604586.
- Gómez-Plaza, E. et al. (2001). "Phenolic Compounds and Color Stability of Red Wines: Effect of Skin Maceration Time". In: *American Journal of Enology and Viticulture* 52.3, pp. 266–270. ISSN: 0002-9254. DOI: 10.5344/ajev.2001.52.3.266.
- Gordillo, Belén et al. (2012). "Comprehensive colorimetric study of anthocyanic copigmentation in model solutions. Effects of pH and molar ratio". In: *Journal of Agricultural and Food Chemistry* 60.11, pp. 2896–2905. ISSN: 0021-8561. DOI: 10.1021/jf2046202.
- Guild J. (1931). "The colorimetric properties of the spectrum". In: *Philosophical Transactions of the Royal Society of London. Series A, Containing Papers of a Mathematical or Physical Character* 230.681-693, pp. 149–187. ISSN: 0264-3952. DOI: 10.1098/rsta.1932.0005.
- Hagerman, Ann E. and Larry G. Butler (1978). "Protein precipitation method for the quantitative determination of tannins". In: *Journal of Agricultural and Food Chemistry* 26.4, pp. 809–812. ISSN: 0021-8561. DOI: 10.1021/jf60218a027.
- Harrell, Frank E. (2001). *Regression Modeling Strategies: With Applications to Linear Models, Logistic Regression, and Survival Analysis*. Springer eBook Collection Mathematics and Statistics. New York, NY: Springer. ISBN: 978-1-4757-3462-1. DOI: 10.1007/978-1-4757-3462-1.
- Heggie, D., R. H. Wardman, and M. R. Luo (1996). "A comparison of the colour differences computed using the CIE94, CMC (l:c) and BFD (l:c) formulae". In: *Journal of the Society of Dyers and Colourists* 112.10, pp. 264–269. ISSN: 0037-9859. DOI: 10.1111/j.1478-4408.1996.tb01755.x.
- Hensel, Marcel (2024). *Reference Data sets for the Blanc de Noir Check Web Application*. URL: [https://github.com/Der-Hensel/Blanc\\_de\\_Noir\\_check](https://github.com/Der-Hensel/Blanc_de_Noir_check).

- Heras-Roger, Jesus, Carlos Díaz-Romero, and Jacinto Darias-Martín (2016). "What Gives a Wine Its Strong Red Color? Main Correlations Affecting Copigmentation". In: *Journal of Agricultural and Food Chemistry* 64.34, pp. 6567–6574. ISSN: 0021-8561. DOI: 10.1021/acs.jafc.6b02221.
- Hinderer, Florian (2020). *UV/Vis-Absorptions- und Fluoreszenz-Spektroskopie: Einführung in die spektroskopische Analyse mit UV- und sichtbarer Strahlung*. essentials. Wiesbaden and Heidelberg: Springer Spektrum. ISBN: 978-3-658-25441-4.
- Hsich, E. (2011). "Identifying Important Risk Factors for Survival in Patient With Systolic Heart Failure Using Random Survival Forests". In: *Circulation: Cardiovascular Quality and Outcomes* 4.1, pp. 39–45. DOI: 10.1161/CIRCOUTCOMES.110.939371.
- Huertas, R. et al. (2003). "Color variability for a wine sample poured into a standard glass wine sampler". In: *Color Research & Application* 28.6, pp. 473–479. ISSN: 0361-2317. DOI: 10.1002/co1.10200.
- Issa-Issa, Hanán et al. (2021). "Quality, Nutritional, Volatile and Sensory Profiles and Consumer Acceptance of Fondillón, a Sustainable European Protected Wine". In: *Agronomy* 11.9, p. 1701. DOI: 10.3390/agronomy11091701.
- Jackson, Ron S. (2008). *Wine science: Principles and applications*. 3rd ed. Amsterdam: Elsevier Acad. Press. ISBN: 978-0-12-373646-8.
- Judd, D. B. and G. Wyszecki (1963). *Color in Business, Science, and Industry*. Wiley. URL: [https://books.google.de/books?id=n8\\_vAAAAAAAJ](https://books.google.de/books?id=n8_vAAAAAAAJ).
- Jurd, Leonard (1964). "Reactions Involved in Sulfite Bleaching of Anthocyanins". In: *Journal of Food Science* 29.1, pp. 16–19. DOI: 10.1111/j.1365-2621.1964.tb01685.x.
- Kingma, Diederik P. and Jimmy Ba (2014). *Adam: A Method for Stochastic Optimization*. URL: <http://arxiv.org/pdf/1412.6980>.
- Kopjar, M. and V. Piližota (2011). "Prevention of thermal degradation of anthocyanins in blackberry juice with addition of different sugars Prevención de degradación termal de antocianinas en zumo de mora con adición de diferentes azúcares". In: *CyTA - Journal of Food* 9.3, pp. 237–242. ISSN: 1947-6337. DOI: 10.1080/19476337.2010.522735. URL: [https://web.archive.org/web/20171021081303id\\_/http://www.tandfonline.com/doi/pdf/10.1080/19476337.2010.522735](https://web.archive.org/web/20171021081303id_/http://www.tandfonline.com/doi/pdf/10.1080/19476337.2010.522735).
- Kortüm, Gustav (1962). *Kolorimetrie · Photometrie und Spektrometrie: Eine Anleitung zur Ausführung von Absorptions-, Emissions-, Fluoreszenz-, Streuungs-, Trübungs- und Reflexionsmessungen*. Vierte Neubearbeitete und Erweiterte Auflage. Vol. 2. Anleitungen für die Chemische Laboratoriumspraxis. Berlin and Heidelberg: Springer. ISBN: 978-3-642-87212-9. DOI: 10.1007/978-3-642-87211-2.
- Kovac, Vladimir. et al. (1992). "Effect of several enological practices on the content of catechins and proanthocyanidins of red wines". In: *Journal of Agricultural and Food Chemistry* 40.10, pp. 1953–1957. ISSN: 0021-8561. DOI: 10.1021/jf00022a045.
- Kruse, Rudolf et al. (2015). *Computational intelligence: Eine methodische Einführung in künstliche neuronale Netze, evolutionäre Algorithmen, Fuzzy-Systeme und Bayes-Netze*. 2., überarbeitete und erweiterte Auflage. Lehrbuch. Wiesbaden: Springer Vieweg. ISBN: 978-3-658-10904-2. URL: <http://www.computational-intelligence.eu>.
- Kuehni, R. G. (1976). "Color-tolerance data and the tentative CIE 1976 L a b formula". In: *Journal of the Optical Society of America* 66.5, pp. 497–500. ISSN: 0030-3941. DOI: 10.1364/josa.66.000497.
- Kumar, Sunny, Kanika Agrawal, and Nelshan Mandan (2020). "Red Wine Quality Prediction Using Machine Learning Techniques". In: *2020 International Conference on Computer Communication and Informatics*. Piscataway, NJ: IEEE, pp. 1–6. ISBN: 978-1-7281-4514-3. DOI: 10.1109/ICCCI48352.2020.9104095.



- Lambert, Stephanie G. et al. (2011). "Copigmentation between malvidin-3-glucoside and some wine constituents and its importance to colour expression in red wine". In: *Food Chemistry* 125.1, pp. 106–115. ISSN: 0308-8146. DOI: 10.1016/j.foodchem.2010.08.045. URL: <https://www.sciencedirect.com/science/article/pii/S0308814610010484>.
- Lawless, Harry T. and Hildegarde Heymann (2010). *Sensory evaluation of food: Principles and practices*. 2. ed. Food science text series. New York, NY and Heidelberg: Springer. ISBN: 978-1-4419-6487-8. URL: <http://www.loc.gov/catdir/enhancements/fy1316/2010932599-b.html>.
- Löffler-Mang, Martin (2012). *Optische Sensorik: Lasertechnik, Experimente, Light Barriers*. 1. Aufl. Studium Fertigung. Wiesbaden: Vieweg + Teubner. ISBN: 978-3-8348-1449-4.
- Lübbe, Eva (2013). *Farbempfindung, Farbbeschreibung und Farbmessung: Eine Formel für die Farbsättigung*. SpringerLink Bücher. Wiesbaden: Springer Vieweg. ISBN: 9783834822284. DOI: 10.1007/978-3-8348-2228-4.
- Luo, M. R., G. Cui, and B. Rigg (2001). "The development of the CIE 2000 colour-difference formula: CIEDE2000". In: *Color Research & Application* 26.5, pp. 340–350. ISSN: 0361-2317. DOI: 10.1002/co1.1049.
- Luo, M. R. and B. Rigg (1986). "Chromaticity-discrimination ellipses for surface colours". In: *Color Research & Application* 11.1, pp. 25–42. ISSN: 0361-2317. DOI: 10.1002/co1.5080110107.
- MacAdam, David L. (1942). "Visual Sensitivities to Color Differences in Daylight\*". In: *Journal of the Optical Society of America* 32.5, p. 247. ISSN: 0030-3941. DOI: 10.1364/JOSA.32.000247.
- Maier, Hans Gerhard (1974). *Lebensmittelanalytik*. 2. Aufl. Vol. 342. UTB. Darmstadt: Steinkopff. ISBN: 3798503788.
- Martínez, J. A. et al. (2001). "Note. Visual and Instrumental Color Evaluation in Red Wines". In: *Food Science and Technology International* 7.5, pp. 439–444. ISSN: 1082-0132. DOI: 10.1106/VFAT-5REN-1WK2-5JGQ.
- Matissek, Reinhard, Markus Fischer, and Gabriele Steiner (2018). *Lebensmittelanalytik*. Berlin, Heidelberg: Springer Berlin Heidelberg. ISBN: 978-3-662-55721-1. DOI: 10.1007/978-3-662-55722-8.
- McDonald, R. and K. J. Smith (1995). "CIE94—a new colour-difference formula\*". In: *Journal of the Society of Dyers and Colourists* 111.12, pp. 376–379. ISSN: 0037-9859. DOI: 10.1111/j.1478-4408.1995.tb01688.x.
- McLaren, K. (1980). "CIELAB Hue-Angle Anomalies at Low Tristimulus Ratios". In: *Color Research & Application* 5.3, pp. 139–143. ISSN: 0361-2317. DOI: 10.1002/co1.5080050306.
- McLaren, K. and B. Rigg (1976). "XII—The SDC Recommended Colour-difference Formula: Change to CIELAB". In: *Journal of the Society of Dyers and Colourists* 92.9, pp. 337–338. ISSN: 0037-9859. DOI: 10.1111/j.1478-4408.1976.tb03300.x.
- Melgosa, Manuel (2013). *CIE/ISO new standard: CIEDE2000*. Leeds. URL: [https://color.org/events/colorimetry/Melgosa\\_CIEDE2000\\_Workshop-July4.pdf](https://color.org/events/colorimetry/Melgosa_CIEDE2000_Workshop-July4.pdf).
- Melgosa, Manuel, Rafael Huertas, and Roy S. Berns (2008). "Performance of recent advanced color-difference formulas using the standardized residual sum of squares index". In: *Journal of the Optical Society of America. A, Optics, image science, and vision* 25.7, pp. 1828–1834. ISSN: 1084-7529. DOI: 10.1364/josaa.25.001828.
- Miniati, E., P. Damiani, and G. Mazza (1992). "Copigmentation and self association of anthocyanins in food model systems". In: *Italian Journal of Food Science* 4.2, pp. 109–116. ISSN: 1120-1770. URL: <https://eurekamag.com/research/007/161/007161334.php>.

- Mirabel, M. et al. (1999). "Copigmentation in Model Wine Solutions: Occurrence and Relation to Wine Aging". In: *American Journal of Enology and Viticulture* 50.2, pp. 211–218. ISSN: 0002-9254. DOI: 10.5344/ajev.1999.50.2.211.
- More, K. A. (2005). "INSTRUMENTATION | Spectrometers". In: *Encyclopedia of modern optics*. Ed. by Robert D. Guenther. ScienceDirect. S.l. and Amsterdam: ScienceDirect Online-Anbieter and Academic Press, pp. 324–336. ISBN: 978-0-12-369395-2. DOI: 10.1016/B0-12-369395-0/00837-X. URL: <https://www.sciencedirect.com/science/article/pii/B012369395000837X>.
- Morrot, G., F. Brochet, and D. Dubourdieu (2001). "The color of odors". In: *Brain and Language* 79.2, pp. 309–320. ISSN: 0093-934X. DOI: 10.1006/brln.2001.2493. URL: <https://www.sciencedirect.com/science/article/pii/S0093934X01924939>.
- Nguyen, Thi H. and Dominik Durner (2023). "Sensory evaluation of wine aroma: Should color-driven descriptors be used?" In: *Food Quality and Preference* 107, p. 104844. ISSN: 0950-3293. DOI: 10.1016/j.foodqual.2023.104844.
- Oberholster, A. et al. (2015). "Barrel maturation, oak alternatives and micro-oxygenation: influence on red wine aging and quality". In: *Food Chemistry* 173, pp. 1250–1258. ISSN: 0308-8146. DOI: 10.1016/j.foodchem.2014.10.043. URL: <https://www.sciencedirect.com/science/article/pii/S0308814614016057>.
- Oglesby, S. (1995). "The effectiveness of CIE94 compared with the CMC equation". In: *Journal of the Society of Dyers and Colourists* 111.12, pp. 380–381. ISSN: 0037-9859. DOI: 10.1111/j.1478-4408.1995.tb01689.x.
- OIV (2021). *Compendium of international methods of wine and must analysis: Volume 1*. Edition 2021. Paris: OIV. ISBN: 978-2-85038-034-1.
- Pangborn, Rose M., Harold W. Berg, and Brenda Hansen (1963). "The Influence of Color on Discrimination of Sweetness in Dry Table-Wine". In: *The American Journal of Psychology* 76.3, p. 492. ISSN: 00029556. DOI: 10.2307/1419795.
- Parmer, Chris (2024). *Dash Styleguide*. URL: <https://codepen.io/chriddyp/pen/bWLwgP>.
- Parpinello, Giuseppina Paola et al. (2009). "Relationship among sensory descriptors, consumer preference and color parameters of Italian Novello red wines". In: *Food Research International* 42.10, pp. 1389–1395. ISSN: 0963-9969. DOI: 10.1016/j.foodres.2009.07.005. URL: <https://www.sciencedirect.com/science/article/pii/S0963996909001884>.
- Pedregosa F et al. (2011). "Scikit-learn: Machine Learning in Python". In: *Journal of Machine Learning Research* 12.85, pp. 2825–2830. URL: <http://jmlr.org/papers/v12/pedregosa11a.html>.
- Pfahl, Leonard et al. (2021). "Effect of Barrel-to-Barrel Variation on Color and Phenolic Composition of a Red Wine". In: *Foods* 10.7. ISSN: 2304-8158. DOI: 10.3390/foods10071669. URL: <https://www.mdpi.com/2304-8158/10/7/1669>.
- Plotly Technologies Inc. (2015). *Plotly: Collaborative data science Publisher: Plotly Technologies Inc.* Montréal, QC. URL: <https://plot.ly>.
- Renner, Hannah, Elke Richling, and Dominik Durner (2022). "Influence of Vibration on Volatile Compounds, Color, SO<sub>2</sub>, and CO<sub>2</sub> of Riesling Sparkling Wine and White Wine". In: *American Journal of Enology and Viticulture* 73.4, pp. 267–276. ISSN: 0002-9254. DOI: 10.5344/ajev.2022.22007. URL: <https://www.ajevonline.org/content/73/4/267>.
- Reynolds, Andrew et al. (2001). "Influence of Fermentation Temperature on Composition and Sensory Properties of Semillon and Shiraz Wines". In: *American Journal of Enology and Viticulture* 52.3, pp. 235–240. ISSN: 0002-9254. DOI: 10.5344/ajev.2001.52.3.235.

- Ribéreau-Gayon, Pascal (2006). *Handbook of enology: Volume 2: The chemistry of wine stabilization and treatments*. 2nd ed. Chichester, West Sussex, England: John Wiley. ISBN: 978-0-470-01038-9. URL: <https://ebookcentral.proquest.com/lib/kxp/detail.action?docID=257673>.
- Riedel, Erwin and Christoph Janiak (2022). *Anorganische Chemie*. De Gruyter. ISBN: 9783110694444. DOI: 10.1515/9783110694444.
- Rigg, B. (1995). "Colour–difference formulae–recent developments". In: *Journal of the Society of Dyers and Colourists* 111.9, pp. 267–271. ISSN: 0037-9859. DOI: 10.1111/J.1478-4408.1995.TB01738.X.
- Rustioni, Laura et al. (2012). "Copigmentation and anti-copigmentation in grape extracts studied by spectrophotometry and post-column-reaction HPLC". In: *Food Chemistry* 132.4. 6th International Conference on Water in Food, pp. 2194–2201. ISSN: 0308-8146. DOI: <https://doi.org/10.1016/j.foodchem.2011.12.058>. URL: <https://www.sciencedirect.com/science/article/pii/S0308814611018504>.
- Es-Safi, Nour-Eddine, Véronique Cheynier, and Michel Moutounet (2002). "Role of Aldehydic Derivatives in the Condensation of Phenolic Compounds with Emphasis on the Sensorial Properties of Fruit-Derived Foods". In: *Journal of Agricultural and Food Chemistry* 50.20, pp. 5571–5585. ISSN: 0021-8561. DOI: 10.1021/jf025503y.
- Schanda, Janos, ed. (2007). *Colorimetry: Understanding th CIE system*. Hoboken, NJ: Wiley. ISBN: 978-0-470-04904-4. URL: <http://www.loc.gov/catdir/enhancements/fy0739/2007026256-d.html>.
- Schwarz, Michael et al. (2005). "Effect of copigments and grape cultivar on the color of red wines fermented after the addition of copigments". In: *Journal of Agricultural and Food Chemistry* 53.21, pp. 8372–8381. ISSN: 0021-8561. DOI: 10.1021/jf051005o.
- Shankar, Maya U., Carmel A. Levitan, and Charles Spence (2010). "Grape expectations: the role of cognitive influences in color-flavor interactions". In: *Consciousness and Cognition* 19.1, pp. 380–390. ISSN: 1053-8100. DOI: 10.1016/j.concog.2009.08.008. URL: <https://www.sciencedirect.com/science/article/pii/S1053810009001342>.
- Shaw, Bipul, Ankur Kumar Suman, and Biswarup Chakraborty (2020). "Wine Quality Analysis Using Machine Learning". In: *Emerging Technology in Modelling and Graphics*. Springer, Singapore, pp. 239–247. DOI: 10.1007/978-981-13-7403-6<sub>23</sub>. URL: [https://link.springer.com/chapter/10.1007/978-981-13-7403-6\\_23](https://link.springer.com/chapter/10.1007/978-981-13-7403-6_23).
- Shimadzu (2023). *How does a photomultiplier tube work?* URL: <https://www.ssi.shimadzu.com/service-support/faq/uv-vis/instrument-design/9/index.html> (visited on 2024).
- Singleton, V. L., Eugene Trousdale, and John Zaya (1979). "Oxidation of Wines. I. Young White Wines Periodically Exposed to Air". In: *American Journal of Enology and Viticulture* 30.1, pp. 49–54. ISSN: 0002-9254. DOI: 10.5344/ajev.1979.30.1.49.
- Sonnet, Daniel (2022). *Neuronale Netze kompakt: Vom Perceptron zum Deep Learning*. IT kompakt. Wiesbaden and Heidelberg: Springer Vieweg. ISBN: 978-3-658-29081-8. URL: <http://www.springer.com/>.
- Specovius, Joachim (2010). *Grundkurs Leistungselektronik: Bauelemente, Schaltungen und Systeme*. 4., aktualisierte und erweiterte Auflage. Springer eBook Collection Computer Science & Engineering. Wiesbaden: Vieweg+Teubner. ISBN: 978-3-8348-9712-1. DOI: 10.1007/978-3-8348-9712-1.
- Starmer, Josh (2022). *The StatQuest illustrated guide to machine learning*. USA: Josh Starmer. ISBN: 979-8986924007.

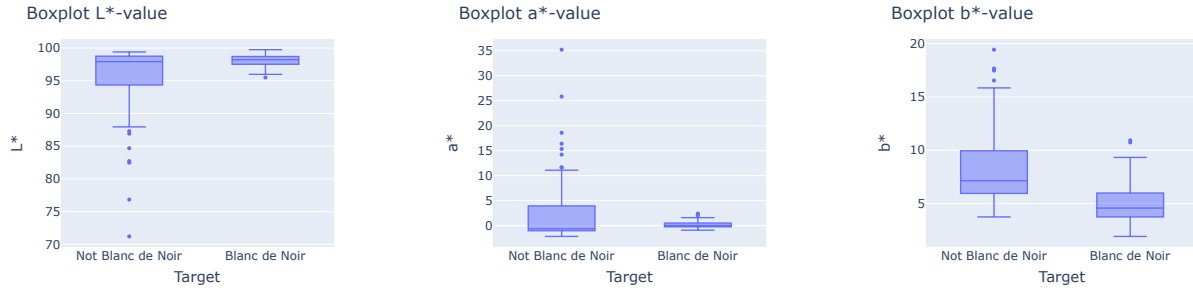
- Stiles, W. S. and J. M. Burch (1959). "N.P.L. Colour-matching Investigation: Final Report (1958)". In: *Optica Acta: International Journal of Optics* 6.1, pp. 1–26. ISSN: 0030-3909. DOI: 10.1080/713826267.
- TensorFlow Developers (2024). *TensorFlow*. DOI: 10.5281/ZENODO.4724125.
- Tofalo, Rosanna, Giovanna Suzzi, and Giorgia Perpetuini (2021). "Discovering the Influence of Microorganisms on Wine Color". In: *Frontiers in microbiology* 12, p. 790935. ISSN: 1664-302X. DOI: 10.3389/fmicb.2021.790935.
- Wang, Qian Janice and Charles Spence (2019). "Drinking through rosé-coloured glasses: Influence of wine colour on the perception of aroma and flavour in wine experts and novices". In: *Food research international (Ottawa, Ont.)* 126, p. 108678. DOI: 10.1016/j.foodres.2019.108678. URL: <https://www.sciencedirect.com/science/article/pii/S0963996919305642>.
- Waterhouse, Andrew L. (2002). "Wine phenolics". In: *Annals of the New York Academy of Sciences* 957.1, pp. 21–36. ISSN: 1749-6632. DOI: 10.1111/j.1749-6632.2002.tb02903.x.
- Wenzel, K., H. H. Dittrich, and M. Heimfarth (2015). "Die Zusammensetzung der Anthocyane in den Beeren verschiedener Rebsorten: 65 Pages / VITIS - Journal of Grapevine Research, Vol. 26 No. 2 (1987): VITIS". In: DOI: 10.5073/vitis.1987.26.65-78.
- Williams, Dudley H. (1975). *Spektroskopische Methoden zur Strukturaufklärung: Mit 49 Übungsaufg., 42 Tab. 3., überarb. dt. Aufl. Vol. 5. Thieme Taschenlehrbuch der organischen Chemie A, theoretische und allgemeine Gebiete. Stuttgart: Thieme. ISBN: 3134372037.*
- Wright, F. E. (1911). *The Methods of Petrographic-microscopic Research, Their Relative Accuracy and Range of Application*. Carnegie Institution of Washington publication. Carnegie Institution of Washington. URL: <https://books.google.de/books?id=-Hk1HiExWQ4C>.
- Wright, W. D. (1929). "A re-determination of the trichromatic coefficients of the spectral colours". In: *Transactions of the Optical Society* 30.4, pp. 141–164. ISSN: 1475-4878. DOI: 10.1088/1475-4878/30/4/301.
- Zhang, Bo et al. (2015). "Copigmentation of malvidin-3-O-glucoside with five hydroxybenzoic acids in red wine model solutions: experimental and theoretical investigations". In: *Food Chemistry* 170, pp. 226–233. ISSN: 0308-8146. DOI: 10.1016/j.foodchem.2014.08.026.

# I Appendix – Basic Framework of the Web Application in Full display

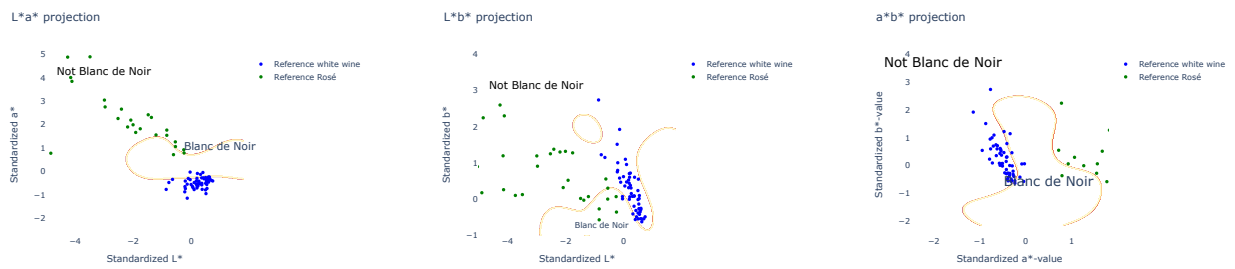
Welcome to Blanc de Noir check

Please insert the CIE L\*a\*b\* parameters in the boxes below

Statistical Distribution



Two-dimensional projection



Three-dimensional projection

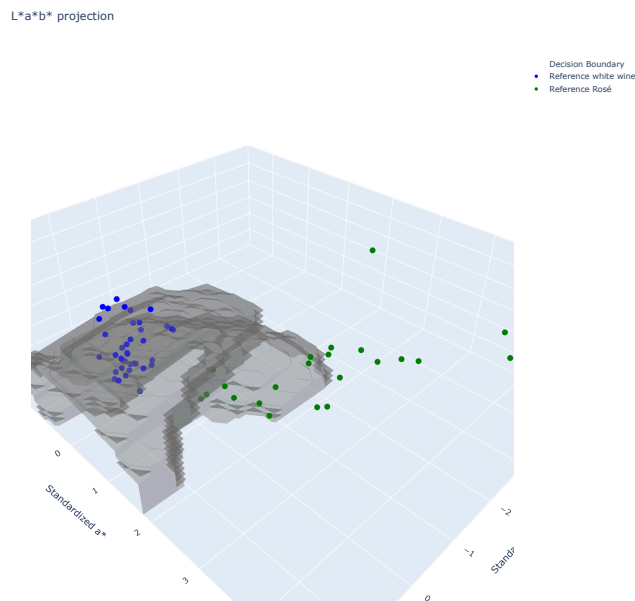


Figure 43: Basic framework of the web application after completely loading.

**II Appendix – Used Red Wines for section 6.2**

ID	Origin	Style	Vintage	Cuvee	Variety 1
RR1	Österreich	Holzfass	2020	nein	Merlot
RR2	Österreich	Barrique	2019	nein	Merlot
RR3	Österreich	Barrique	2019	nein	Merlot
RR4	Österreich	Holzfass	2018	nein	Merlot
RR5	Italien	INOX	2020	ja	Sangiovese
RR6	Italien	Holzfass	2019	nein	Sangiovese
RR7	Italien	Holzfass	2018	nein	Sangiovese
RR8	Italien	Barrique	2017	ja	Sangiovese
RR9	Italien	Barrique	2017	nein	Sangiovese
RR10	Italien	INOX	2021	nein	Sangiovese
RR11	Italien	INOX	2021	ja	Sangiovese
RR12	Italien	Barrique	2021	nein	Sangiovese
RR13	Italien	INOX	2020	ja	Sangiovese
RR14	Italien	INOX	2020	ja	Sangiovese
RR15	Italien	Holzfass	2020	ja	Sangiovese
RR16	Italien	Holzfass	2019	ja	Sangiovese
RR17	Italien	Barrique	2019	ja	Sangiovese
RR18	Italien	Holzfass	2018	ja	Sangiovese
RR19	Italien	INOX	2021	nein	Sangiovese
RR20	Italien	Holzfass	2020	ja	Sangiovese
RR21	Italien	Barrique	2020	ja	Sangiovese
RR22	Italien	INOX	2020	ja	Sangiovese
RR23	Italien	INOX	2020	ja	Sangiovese
RR24	Italien	INOX	2020	ja	Sangiovese
RR25	Italien	Barrique	2018	ja	Sangiovese
RR26	Italien	Holzfass	2016	nein	Sangiovese
RR27	Italien	Barrique	2013	ja	Sangiovese
RR28	Spanien	INOX	2019	ja	Tempranillo
RR29	Spanien	Barrique	2019	nein	Tempranillo
RR30	Spanien	Barrique	2017	ja	Tempranillo
RR31	Spanien	Barrique	2017	ja	Tempranillo
RR32	Spanien	INOX	2021	nein	Tempranillo
RR33	Spanien	INOX	2021	nein	Tempranillo
RR34	Spanien	Barrique	2020	ja	Tempranillo
RR35	Spanien	INOX	2020	nein	Tempranillo
RR36	Spanien	INOX	2020	nein	Tempranillo
RR37	Spanien	INOX	2021	ja	Tempranillo
RR38	Spanien	INOX	2021	ja	Tempranillo
RR39	Chile	Barrique	2021	nein	Cabernet Sauvignon
RR40	Chile	Barrique	2021	ja	Cabernet Sauvignon
RR41	Chile	Barrique	2021	ja	Cabernet Sauvignon
RR42	Chile	INOX	2020	nein	Cabernet Sauvignon
RR43	Chile	Barrique	2020	nein	Cabernet Sauvignon
RR44	Chile	INOX	2020	nein	Cabernet Sauvignon
RR45	Italien	INOX	2021	nein	Cabernet Sauvignon
RR46	Italien	INOX	2020	nein	Sangiovese
RR47	Italien	Barrique	2020	nein	Sangiovese

RR48	Italien	Barrique	2021	ja	Cabernet Sauvignon
RR49	Spanien	Barrique	2019	nein	Tempranillo
RR50	Spanien	INOX	2019	ja	Tempranillo
RR51	Spanien	Holzfass	2019	nein	Tempranillo
RR52	Italien	INOX	2018	ja	Sangiovese
RR53	Spanien	INOX	2019	nein	Tempranillo
RR54	Spanien	INOX	2019	nein	Tempranillo
RR55	Spanien	Holzfass	2018	nein	Tempranillo
RR56	Spanien	INOX	2018	nein	Tempranillo
RR57	Italien	INOX	2021	nein	Merlot
RR58	Italien	INOX	2021	nein	Merlot
RR59	Italien	INOX	2021	ja	Merlot
RR60	Italien	INOX	2021	ja	Merlot
RR61	Italien	INOX	2021	ja	Merlot
RR62	Italien	INOX	o.A.	ja	Merlot
RR63	Spanien	Barrique	2018	nein	Tempranillo
RR64	Spanien	INOX	2018	ja	Tempranillo
RR65	Spanien	Holzfass	2018	nein	Tempranillo
RR66	Spanien	Holzfass	2020	nein	Tempranillo
RR67	Spanien	Holzfass	2020	nein	Tempranillo
RR68	Spanien	INOX	2019	nein	Tempranillo
RR69	Spanien	Holzfass	2019	nein	Tempranillo
RR70	Italien	Barrique	o.A.	ja	Primitivo
RR71	Italien	INOX	o.A.	ja	Primitivo
RR72	Italien	Barrique	2021	ja	Primitivo
RR73	Italien	INOX	2021	ja	Primitivo
RR74	Italien	INOX	2021	ja	Primitivo
RR75	Italien	INOX	2020	nein	Primitivo
RR76	Italien	INOX	2020	nein	Primitivo
RR77	Italien	Barrique	2020	nein	Primitivo
RR78	Italien	Barrique	2019	nein	Primitivo
RR79	Italien	Barrique	2018	nein	Primitivo
RR80	Italien	INOX	2017	nein	Primitivo
RR81	Spanien	Barrique	2017	nein	Tempranillo
RR82	Spanien	Barrique	2017	nein	Tempranillo
RR83	Spanien	INOX	2017	ja	Tempranillo
RR84	Spanien	INOX	2017	nein	Tempranillo
RR85	Italien	INOX	2021	nein	Primitivo
RR86	Italien	INOX	2021	nein	Primitivo
RR87	Italien	INOX	2021	nein	Primitivo
RR88	Italien	INOX	2021	nein	Primitivo
RR89	Italien	INOX	2021	nein	Primitivo
RR90	Italien	INOX	2021	nein	Primitivo
RR91	Italien	INOX	2021	nein	Primitivo
RR92	Italien	INOX	2021	nein	Primitivo
RR93	Spanien	INOX	2019	nein	Tempranillo
RR94	Spanien	INOX	2019	nein	Tempranillo
RR95	Spanien	Holzfass	2019	nein	Tempranillo

RR96	Spanien	INOX	2018	nein	Tempranillo
RR97	Spanien	INOX	2018	nein	Tempranillo
RR98	Frankreich	Barrique	2018	ja	Cabernet Sauvignon
RR99	Libanon	INOX	2016	ja	Cabernet Sauvignon
RR100	Spanien	INOX	2021	ja	Tempranillo
RR101	Spanien	INOX	2021	ja	Tempranillo
RR102	Spanien	INOX	2020	ja	Tempranillo
RR103	Spanien	INOX	2020	nein	Tempranillo
RR104	Spanien	INOX	2019	ja	Tempranillo
RR105	Spanien	Barrique	2018	ja	Tempranillo
RR106	Spanien	Barrique	2016	nein	Tempranillo
RR107	Portugal	INOX	2020	ja	Sangiovese
RR108	Australien	INOX	2021	nein	Cabernet Sauvignon
RR109	Australien	INOX	2021	nein	Cabernet Sauvignon
RR110	Australien	Barrique	2020	nein	Cabernet Sauvignon
RR111	Australien	Barrique	2020	nein	Cabernet Sauvignon
RR112	Australien	Barrique	2020	nein	Cabernet Sauvignon
RR113	Australien	INOX	2020	ja	Cabernet Sauvignon
RR114	Australien	Holzfass	2020	ja	Cabernet Sauvignon
RR115	Australien	Holzfass	2020	ja	Cabernet Sauvignon
RR116	Australien	Holzfass	2020	ja	Cabernet Sauvignon
RR117	Rumänien	INOX	2021	nein	Merlot
RR118	Rumänien	INOX	2020	nein	Merlot
RR119	Bosnien-Herzegowina	INOX	2016	nein	Merlot
RR120	Slowenien	Barrique	2016	nein	Merlot
RR121	Rumänien	Barrique	2016	ja	Merlot
RR122	Ungarn	Barrique	2015	ja	Merlot
RR123	Spanien	INOX	2021	nein	Tempranillo
RR124	Spanien	INOX	2021	ja	Tempranillo
RR125	Spanien	INOX	2020	nein	Tempranillo
RR126	Spanien	INOX	2021	nein	Tempranillo
RR127	Spanien	INOX	2021	nein	Tempranillo
RR128	Spanien	Barrique	2020	ja	Tempranillo
RR129	Spanien	Barrique	2020	nein	Tempranillo
RR130	Portugal	Barrique	2020	ja	Merlot
RR131	Rumänien	INOX	2021	nein	Cabernet Sauvignon
RR132	Kroatien	Barrique	2018	nein	Cabernet Sauvignon
RR133	Moldawien	Barrique	2016	ja	Cabernet Sauvignon
RR134	Moldawien	Barrique	2015	ja	Cabernet Sauvignon
RR135	Kroatien	Barrique	2013	ja	Cabernet Sauvignon
RR136	Spanien	INOX	2021	ja	Tempranillo
RR137	Spanien	INOX	2021	nein	Tempranillo
RR138	Spanien	INOX	2021	ja	Tempranillo
RR139	Spanien	INOX	2021	nein	Tempranillo
RR140	Spanien	INOX	2021	nein	Tempranillo
RR141	Spanien	INOX	2021	nein	Tempranillo
RR142	Spanien	INOX	2019	nein	Tempranillo
RR143	Spanien	Barrique	2018	nein	Tempranillo



RR144	Spanien	Barrique	2015	ja	Tempranillo
RR145	Chile	Barrique	2021	ja	Cabernet Sauvignon
RR146	Chile	Holzfass	2021	ja	Cabernet Sauvignon
RR147	Chile	INOX	2021	nein	Cabernet Sauvignon
RR148	Chile	INOX	2020	ja	Cabernet Sauvignon
RR149	Türkei	Barrique	2020	nein	Tempranillo
RR150	Italien	Barrique	2020	nein	Tempranillo
RR151	Deutschland	INOX	2020	nein	Tempranillo
RR152	Moldawien	Barrique	2019	nein	Tempranillo
RR153	Spanien	Barrique	2018	ja	Tempranillo
RR154	Spanien	Barrique	2018	ja	Tempranillo
RR155	Spanien	Barrique	2017	ja	Tempranillo
RR156	Spanien	Barrique	2017	ja	Tempranillo
RR157	Italien	INOX	2021	nein	Primitivo
RR158	Italien	Barrique	2021	ja	Primitivo
RR159	Spanien	INOX	2016	nein	Tempranillo
RR160	Spanien	Barrique	2015	ja	Tempranillo
RR161	Spanien	Barrique	2015	ja	Tempranillo
RR162	Spanien	INOX	2015	ja	Tempranillo
RR163	Spanien	Barrique	2014	nein	Tempranillo
RR164	Spanien	INOX	2011	ja	Tempranillo
RR165	Australien	INOX	2019	ja	Cabernet Sauvignon
RR166	Australien	Holzfass	2018	ja	Cabernet Sauvignon
RR167	Australien	Holzfass	2018	ja	Cabernet Sauvignon
RR168	Australien	Holzfass	2017	ja	Cabernet Sauvignon
RR169	Australien	Barrique	2016	nein	Cabernet Sauvignon
RR170	Frankreich	Barrique	2019	ja	Cabernet Sauvignon
RR171	Frankreich	Barrique	2019	ja	Merlot
RR172	Frankreich	Barrique	2018	ja	Merlot
RR173	Frankreich	Barrique	2018	ja	Merlot
RR174	Frankreich	Barrique	2017	ja	Merlot
RR175	Frankreich	INOX	2017	ja	Merlot
RR176	Frankreich	Holzfass	2016	ja	Merlot
RR177	Italien	INOX	2021	nein	Primitivo
RR178	Italien	INOX	2021	nein	Primitivo
RR179	Italien	INOX	2021	nein	Primitivo
RR180	Italien	INOX	2020	ja	Primitivo
RR181	Italien	INOX	2020	ja	Primitivo
RR182	Italien	INOX	2018	nein	Primitivo
RR183	Nordmazedonien	Holzfass	2019	nein	Cabernet Sauvignon
RR184	Australien	Barrique	2021	nein	Merlot
RR185	Chile	Barrique	2021	ja	Merlot
RR186	Chile	Barrique	2021	ja	Merlot
RR187	Südafrika	INOX	2021	nein	Merlot
RR188	Chile	INOX	2021	nein	Merlot
RR189	Südafrika	Barrique	2020	ja	Merlot
RR190	United States of America	Barrique	2020	ja	Merlot
RR191	Australien	INOX	2019	nein	Merlot

RR192	Libanon	INOX	2016	ja	Merlot
RR193	Spanien	Barrique	2021	ja	Merlot
RR194	Spanien	INOX	2021	ja	Merlot
RR195	Spanien	Barrique	2020	nein	Merlot
RR196	Spanien	Barrique	2019	ja	Merlot
RR197	Spanien	Holzfass	2019	ja	Merlot
RR198	Spanien	Holzfass	2016	ja	Merlot
RR199	Spanien	Barrique	2014	ja	Merlot
RR200	Südafrika	INOX	2021	nein	Cabernet Sauvignon
RR201	Südafrika	Barrique	2020	ja	Cabernet Sauvignon
RR202	Südafrika	Holzfass	2020	nein	Cabernet Sauvignon
RR203	Südafrika	INOX	2020	ja	Cabernet Sauvignon
RR204	Südafrika	Barrique	2020	nein	Cabernet Sauvignon
RR205	Südafrika	Holzfass	2019	ja	Cabernet Sauvignon
RR206	Südafrika	Barrique	2019	ja	Cabernet Sauvignon
RR207	Südafrika	Holzfass	2018	ja	Cabernet Sauvignon
RR208	Südafrika	Holzfass	2018	nein	Cabernet Sauvignon
RR209	Südafrika	Barrique	2018	ja	Cabernet Sauvignon
RR210	Italien	Holzfass	2019	nein	Sangiovese
RR211	Italien	INOX	2019	nein	Sangiovese
RR212	Italien	Barrique	2019	ja	Sangiovese
RR213	Italien	INOX	2018	ja	Sangiovese
RR214	Türkei	Barrique	2020	nein	Merlot
RR215	Bulgarien	Barrique	2020	ja	Merlot
RR216	Bulgarien	Barrique	2020	nein	Merlot
RR217	Türkei	Barrique	2019	ja	Merlot
RR218	Türkei	Barrique	2019	ja	Merlot
RR219	Rumänien	Barrique	2019	ja	Merlot
RR220	Tschechische Republik	Holzfass	2018	nein	Merlot
RR221	Rumänien	Barrique	2018	ja	Merlot
RR222	Kroatien	Barrique	2018	nein	Merlot
RR223	Italien	Barrique	2021	ja	Cabernet Sauvignon
RR224	Italien	Holzfass	2020	ja	Cabernet Sauvignon
RR225	Italien	Barrique	2019	nein	Cabernet Sauvignon
RR226	Italien	Barrique	2019	ja	Cabernet Sauvignon
RR227	Italien	Barrique	2019	ja	Cabernet Sauvignon
RR228	Ukraine	INOX	o.A.	nein	Cabernet Sauvignon
RR229	Türkei	Barrique	2020	nein	Cabernet Sauvignon
RR230	Türkei	Barrique	2019	ja	Cabernet Sauvignon
RR231	Türkei	Barrique	2019	ja	Cabernet Sauvignon
RR232	Türkei	Barrique	2019	nein	Cabernet Sauvignon
RR233	Türkei	Barrique	2019	ja	Cabernet Sauvignon
RR234	Rumänien	Barrique	2019	ja	Cabernet Sauvignon
RR235	Tschechische Republik	Holzfass	2018	nein	Cabernet Sauvignon
RR236	Ungarn	Barrique	2018	nein	Cabernet Sauvignon
RR237	Rumänien	Barrique	2017	nein	Cabernet Sauvignon
RR238	Ungarn	Barrique	2016	ja	Cabernet Sauvignon
RR239	Italien	INOX	2020	ja	Merlot

RR240	Italien	INOX	2020	ja	Merlot
RR241	Italien	Barrique	2019	nein	Merlot
RR242	Italien	INOX	2019	ja	Merlot
RR243	Italien	Barrique	2013	ja	Merlot
RR244	Armenien	Barrique	2019	ja	Cabernet Sauvignon
RR245	Moldawien	Barrique	2019	ja	Cabernet Sauvignon
RR246	Kroatien	Holzfass	2019	ja	Cabernet Sauvignon
RR247	Bosnien-Herzegowina	Barrique	2017	ja	Cabernet Sauvignon
RR248	Kroatien	Barrique	2017	ja	Cabernet Sauvignon
RR249	Bulgarien	Barrique	2017	ja	Cabernet Sauvignon
RR250	China, Volksrepublik	Holzfass	o.A.	nein	Cabernet Sauvignon
RR251	China, Volksrepublik	Barrique	o.A.	ja	Cabernet Sauvignon
RR252	China, Volksrepublik	Barrique	o.A.	nein	Cabernet Sauvignon
RR253	Chile	Holzfass	2022	ja	Cabernet Sauvignon
RR254	Chile	Holzfass	2022	ja	Cabernet Sauvignon
RR255	China, Volksrepublik	INOX	2018	ja	Cabernet Sauvignon
RR256	China, Volksrepublik	Barrique	2017	nein	Cabernet Sauvignon
RR257	Spanien	INOX	2021	nein	Cabernet Sauvignon
RR258	Spanien	INOX	2021	ja	Cabernet Sauvignon
RR259	Spanien	Barrique	2021	nein	Cabernet Sauvignon
RR260	Spanien	INOX	2021	nein	Cabernet Sauvignon
RR261	Spanien	INOX	2021	nein	Cabernet Sauvignon
RR262	Spanien	Barrique	2020	nein	Cabernet Sauvignon
RR263	Spanien	Barrique	2019	ja	Cabernet Sauvignon
RR264	Spanien	INOX	2019	ja	Cabernet Sauvignon
RR265	Spanien	Barrique	2018	ja	Cabernet Sauvignon
RR266	Spanien	INOX	2017	nein	Cabernet Sauvignon
RR267	Spanien	Barrique	2013	ja	Cabernet Sauvignon
RR268	Italien	Holzfass	2020	ja	Sangiovese
RR269	Italien	INOX	2020	nein	Sangiovese
RR270	Italien	Holzfass	2019	nein	Sangiovese
RR271	Spanien	INOX	2021	nein	Tempranillo
RR272	Spanien	INOX	2021	nein	Tempranillo
RR273	Spanien	INOX	2019	nein	Tempranillo
RR274	Spanien	Barrique	2018	nein	Tempranillo
RR275	Spanien	Holzfass	2018	nein	Tempranillo
RR276	Spanien	Holzfass	2018	nein	Tempranillo
RR277	Spanien	Holzfass	2017	nein	Tempranillo
RR278	Spanien	Holzfass	2017	nein	Tempranillo
RR279	Spanien	Holzfass	2015	nein	Tempranillo
RR280	Italien	Holzfass	2021	nein	Primitivo
RR281	Italien	INOX	2021	nein	Primitivo
RR282	Italien	INOX	2021	nein	Primitivo
RR283	Italien	Barrique	2021	nein	Primitivo
RR284	Italien	Holzfass	2020	nein	Primitivo
RR285	Italien	Barrique	2020	nein	Primitivo
RR286	Italien	INOX	2020	ja	Primitivo
RR287	Italien	Holzfass	2019	nein	Primitivo

RR288	Italien	Barrique	2021	nein	Merlot
RR289	Italien	INOX	2021	nein	Merlot
RR290	Italien	INOX	2015	ja	Merlot
RR291	Portugal	Barrique	2021	ja	Cabernet Sauvignon
RR292	Portugal	Barrique	2021	ja	Cabernet Sauvignon
RR293	Portugal	Barrique	2020	ja	Cabernet Sauvignon
RR294	Portugal	Barrique	2020	ja	Cabernet Sauvignon
RR295	Portugal	INOX	2020	ja	Cabernet Sauvignon
RR296	Portugal	Barrique	2019	nein	Cabernet Sauvignon
RR297	Italien	INOX	2020	nein	Sangiovese
RR298	Italien	INOX	2020	nein	Sangiovese
RR299	Italien	INOX	2017	nein	Sangiovese
RR300	Italien	INOX	2016	nein	Sangiovese
RR301	Italien	INOX	2021	nein	Primitivo
RR302	Italien	Barrique	2021	nein	Primitivo
RR303	Italien	INOX	2021	nein	Primitivo
RR304	Italien	Barrique	2021	nein	Primitivo
RR305	Italien	Barrique	2021	nein	Primitivo
RR306	Italien	Barrique	2021	nein	Primitivo
RR307	Italien	Barrique	2021	nein	Primitivo
RR308	Italien	INOX	2021	nein	Primitivo
RR309	Frankreich	INOX	2021	nein	Cabernet Sauvignon
RR310	Frankreich	INOX	2021	nein	Cabernet Sauvignon
RR311	Frankreich	INOX	2021	nein	Cabernet Sauvignon
RR312	Frankreich	INOX	2020	nein	Cabernet Sauvignon
RR313	Frankreich	INOX	2020	nein	Cabernet Sauvignon
RR314	Frankreich	INOX	2020	nein	Cabernet Sauvignon
RR315	Frankreich	INOX	2020	ja	Cabernet Sauvignon
RR316	Spanien	Holzfass	2021	nein	Tempranillo
RR317	Spanien	Barrique	2020	nein	Tempranillo
RR318	Spanien	Barrique	2020	nein	Tempranillo
RR319	Spanien	INOX	2020	nein	Tempranillo
RR320	Spanien	INOX	2020	nein	Tempranillo
RR321	Italien	INOX	2021	nein	Primitivo
RR322	Italien	INOX	2021	nein	Primitivo
RR323	Italien	INOX	2021	nein	Primitivo
RR324	Italien	INOX	2020	nein	Primitivo
RR325	Italien	Barrique	2020	nein	Primitivo
RR326	Italien	INOX	2019	nein	Primitivo
RR327	Italien	INOX	2019	nein	Primitivo
RR328	Italien	Barrique	2018	nein	Primitivo
RR329	Spanien	INOX	2021	nein	Tempranillo
RR330	Chile	INOX	2021	nein	Cabernet Sauvignon
RR331	Chile	INOX	2021	nein	Cabernet Sauvignon
RR332	Chile	Barrique	2021	nein	Cabernet Sauvignon
RR333	Chile	INOX	2021	nein	Cabernet Sauvignon
RR334	Chile	INOX	2021	nein	Cabernet Sauvignon
RR335	Chile	INOX	2021	nein	Cabernet Sauvignon

RR336	Chile	INOX	2021	nein	Cabernet Sauvignon
RR337	Italien	INOX	2020	nein	Sangiovese
RR338	Italien	INOX	2020	ja	Sangiovese
RR339	Italien	Holzfass	2020	ja	Sangiovese
RR340	Italien	Holzfass	2020	nein	Sangiovese
RR341	Italien	Holzfass	2019	nein	Sangiovese
RR342	Italien	Barrique	2019	nein	Sangiovese
RR343	Frankreich	INOX	2021	nein	Merlot
RR344	Frankreich	INOX	2021	nein	Merlot
RR345	Frankreich	INOX	2021	nein	Merlot
RR346	Frankreich	INOX	2020	nein	Merlot
RR347	Italien	Barrique	2019	nein	Sangiovese
RR348	Italien	INOX	2019	nein	Sangiovese
RR349	Italien	Holzfass	2019	nein	Sangiovese
RR350	Italien	Holzfass	2018	ja	Sangiovese
RR351	Italien	Holzfass	2018	nein	Sangiovese
RR352	Italien	INOX	2018	ja	Sangiovese
RR353	Griechenland	Barrique	2019	nein	Cabernet Sauvignon
RR354	Italien	INOX	2021	nein	Primitivo
RR355	Italien	Holzfass	2020	nein	Primitivo
RR356	Frankreich	Barrique	2019	ja	Cabernet Sauvignon
RR357	Frankreich	Barrique	2019	ja	Merlot
RR358	Frankreich	INOX	2019	ja	Cabernet Sauvignon
RR359	Frankreich	Holzfass	2019	ja	Merlot
RR360	Frankreich	INOX	2019	ja	Cabernet Sauvignon
RR361	Frankreich	Holzfass	2018	ja	Cabernet Sauvignon
RR362	Italien	Barrique	2018	nein	Sangiovese
RR363	Italien	Holzfass	2017	ja	Sangiovese
RR364	Italien	Holzfass	2016	nein	Sangiovese
RR365	United States of America	Barrique	2019	ja	Cabernet Sauvignon
RR366	United States of America	Barrique	2019	ja	Cabernet Sauvignon