ISSN: 2572-5971

Research Article

Advances in Nutrition & Food Science

Determination of Water Content in Vanilla Pods and Powder Using Near-Infrared Spectroscopy

A Barreau[1,2*], E Mezzatesta[1], S Charvet[2], P Faury[2], I Bombarda[1] and N Dupuy[1]

[1]*IMBE, Aix Marseille Univ, Avignon Univ, CNRS, IRD, Marseille, France*

[2]*Vanalyse, 14 Boulevard du Travail, 13013 Marseille, France*

[*]**Corresponding Author**
A. Barreau, IMBE, Aix Marseille Univ, Avignon Univ, CNRS, IRD, Marseille and Vanalyse, 14 Boulevard du Travail, 13013 Marseille, France.

Submitted: 2025, Dec 30; **Accepted**: 2026, Jan 08; **Published**: 2026, Feb 03

Citation: Barreau, A., Mezzatesta, E., Charvet, S., Faury, P., Bombarda, I., et al. (2026). Determination of Water Content in Vanilla Pods and Powder Using Near-Infrared Spectroscopy. *Adv Nutr Food Sci, 11*(1), 01-08.

Abstract

The quality of vanilla depends heavily on its water content, a key parameter that influences its stability, preservation, and sensory properties. Conventional moisture determination by loss-on-drying, while accurate, is slow and destructive. In this study, a rapid, non-destructive approach based on near-infrared (NIR) spectroscopy combined with chemometric modeling was developed to predict the moisture content of vanilla samples. A total of 123 samples were analyzed by NIR spectroscopy and thermogravimetric analysis (reference values). Principal component analysis revealed a strong correlation between spectral variance and water content in the 7400-4000 cm^{-1} region. A partial least squares (PLS) regression model was optimized and validated, showing optimal performance with four latent variables ($R^2 = 0.99$, RMSEC = 0.98). Prediction on an independent set of 20 samples yielded a Q^2 of 0.99 and an RMSEP of 0.82, confirming the robustness and generalization of the model. Repeatability tests further demonstrated high stability and low variability of the predicted values. The method demonstrates strong potential for routine industrial quality control. Expanding the calibration set could support broader applicability across vanilla varieties.

Keywords: Chemometrics, NIR, Vanilla, Water Content, Quality Control

1. Introduction

Vanilla, and more specifically its major constituent vanillin, is an important raw material for the food industry, as well as for various sectors such as perfumery, cosmetics, chocolate making, and pharmaceutical industry [1,2]. Its complex aromatic profile and high added value make it a major economic resource for producing countries. The relative scarcity of this product is largely due to strict biological and agronomic constraints: flowering occurs only once a year, and pollination must be carried out manually within an extremely narrow time window [3-5]. The quality of vanilla beans is determined by several physicochemical parameters, including commercial grade, aromatic intensity, the content of phenolic compounds such as vanillin, and water content [6-8]. Among these criteria, as well as any other natural product, water content is a key factor for the stability, preservation, and sensory properties of vanilla. Excess moisture promotes microbial growth (molds, yeasts, bacteria), which can lead to accelerated degradation, reduced aromatic intensity, and a significant decrease in shelf life [9,10]. Conversely, excessive drying can alter color, induce the loss of volatile compounds, and degrade overall sensory characteristics [11-13]. Requirements regarding moisture vary across markets: some applications require pods with rather low water content (20–25%), while others favor more supple and visually attractive pods with higher water content (30–35%) [14,15]. Several analytical techniques are currently available for determining the water content, but many of them require complex sample preparation steps, the use of chemical reagents, or are time-consuming, including loss-on-drying using a thermogravimetric balance, which is the technique used in this study to obtain reference data [16]. In this context, the development of rapid, non-destructive, and sufficiently robust methods capable of competing with reference methods is essential. Near-infrared spectroscopy (NIR) is a particularly attractive analytical approach due to its rapid data acquisition, non-destructive nature, and ability to simultaneously provide multivariate information [17-19]. Since stretching and combination vibrations associated with water molecules exhibit a strong response in this spectral region, NIR is particularly well suited for quantifying moisture in complex matrices [19,20]. The integration of this technique linked to chemometric methods, particularly Partial Least Squares (PLS) regression, enables the development of robust predictive models. The present study aims to develop a chemometric model based on NIR spectroscopy for the rapid and accurate prediction of water content in vanilla samples. After optimizing the calibration model, its performance and robustness will be evaluated on different sample batches. The goal is to provide a reliable analytical tool suitable for routine quality control of spices within the food industry.

2. Materials and Methods

2.1. Sampling

One hundred twenty-three vanilla samples were used in this study. Among them, 22 were in powdered form and 101 were whole beans (including 2 from the species *V. tahitensis* and 99 from *V. planifolia*).

2.2. Grinding

For pod-form samples, an electric WSG30E grinder from WARING was used (operating at a rotation speed of 19,000 rpm for approximately 2 × 10 seconds). The samples used for NIR spectroscopy and for moisture loss analysis by thermogravimetric balance came from the same batches and were ground simultaneously, ensuring a homogeneous composition and eliminating potential sources of variability. Both types of analyses were carried out as quickly as possible to minimize the influence of time or environmental conditions on the sample state. Approximately 20 g of each sample were used for grinding: 10 g were used to obtain reference values by thermogravimetric analysis, and the remaining 10 g of the ground sample were used for the acquisition of the NIR spectra.

2.3. Water Content by Weight Loss Through Thermo balance

A moisture analyzer from OHAUS was used. Thermogravimetric moisture determination is easy and inexpensive to implement. The sample is heated to 105 °C using an infrared heat source, and the recorded mass loss is interpreted as evaporated water. This takes approximately 40 minutes.

2.4. Acquisition of NIR Spectra

The diffuse reflectance spectra of the samples were measured using an MPA II FT-NIR spectrometer (Bruker, USA). An integrating sphere was used to collect the spectra, and the NIR spectrometer measurement parameters were as follows: a sample cell with a 30 mm quartz window was filled with the samples. The spectra were recorded using OPUS software (version 7.5, Bruker Optik GmbH, Germany). The number of scans was set to 100, with a resolution of 4 cm^{-1}, using an internal reference for automatic background noise acquisition, all in an air-conditioned room. Spectra were recorded from 3800 to 12,500 cm^{-1}. Each sample was measured three times, and the averaged spectra was used for data analysis.

2.5. Chemometrics

Data analysis was performed using The Unscrambler X 10.3 software. Before applying PLS regression, two spectral preprocessing techniques were implemented: standard normal variate (SNV) normalization and the Savitzky–Golay first derivative (FD). The SNV method corrects variations that do not originate from the chemical composition of the samples. It is particularly effective in eliminating effects caused by instrumental fluctuations or measurement conditions that could compromise the results. This correction ensures that the variations observed in each spectrum reflect only the true chemical composition of the sample, independent of external influences [21]. The first derivative obtained through the Savitzky–Golay method further enhances the spectra by reducing noise through smoothing while emphasizing relevant information [22]. A principal component analysis (PCA) was also carried out on the entire dataset. This exploratory approach aims to uncover the internal structure of the data and to condense the information into a limited number of independent components that best explain the variance [23]. Partial Least Squares (PLS) regression is a supervised method used to model the relationship between two sets of variables—in this case, the NIR absorbance spectra and the moisture content of the samples. This technique relies on extracting latent variables that summarize the information contained in the spectra while maximizing their association with the target variable [24,25]. The development of a PLS model involves three main phases: the calibration phase, during which a model is built using approximately two-thirds of the samples by linking the spectra to the known values of the target variable. The validation phase, which is used to determine and validate the optimal number of latent variables required for reliable prediction. The prediction phase, in which this optimal number of latent variables is applied to the remaining samples to assess the model's robustness and predictive performance.

The performance of the PLS model was evaluated using the correlation coefficients obtained during calibration and testing (R^2 and Q^2), as well as the root mean square errors of calibration (RMSEC) and prediction (RMSEP).

3. Results and Discussions

3.1. NIR Spectra of Samples

The FT-NIR spectra of three samples with different water contents are shown on Figure 1. These spectra cover a wide range of wavenumbers, from approximately 12 500 cm^{-1} to 3 800 cm^{-1}. In near-infrared spectroscopy, harmonic and combination bands are typically observed. The interpretation of the vibrational bands was carried out in accordance with the literature [19, 26-28]. Bands around 4 300 cm^{-1} are characteristic of combinations of CH stretching vibrations from CH_3 and CH_2 groups with other vibrations; bands around 5 800 cm^{-1} correspond to the first overtone of CH stretching vibrations from CH_3, CH_2 and CH=CH groups; and bands around 8 300 cm^{-1} represent the second overtone of CH stretching vibrations from CH_3 and CH_2 groups. Water is a natural component of any plant-based or food product, and the NIR spectra reveal the presence of water in the analyzed samples, particularly through the two bands around 5 200 cm⁻¹ (corresponding to combination bands of OH stretching and bending vibrations) and 7 000 cm⁻¹ (corresponding to the first overtone of symmetric and asymmetric OH stretching vibrations). These features are characteristic of water and demonstrate that NIR spectroscopy is a reliable method for determining water content.

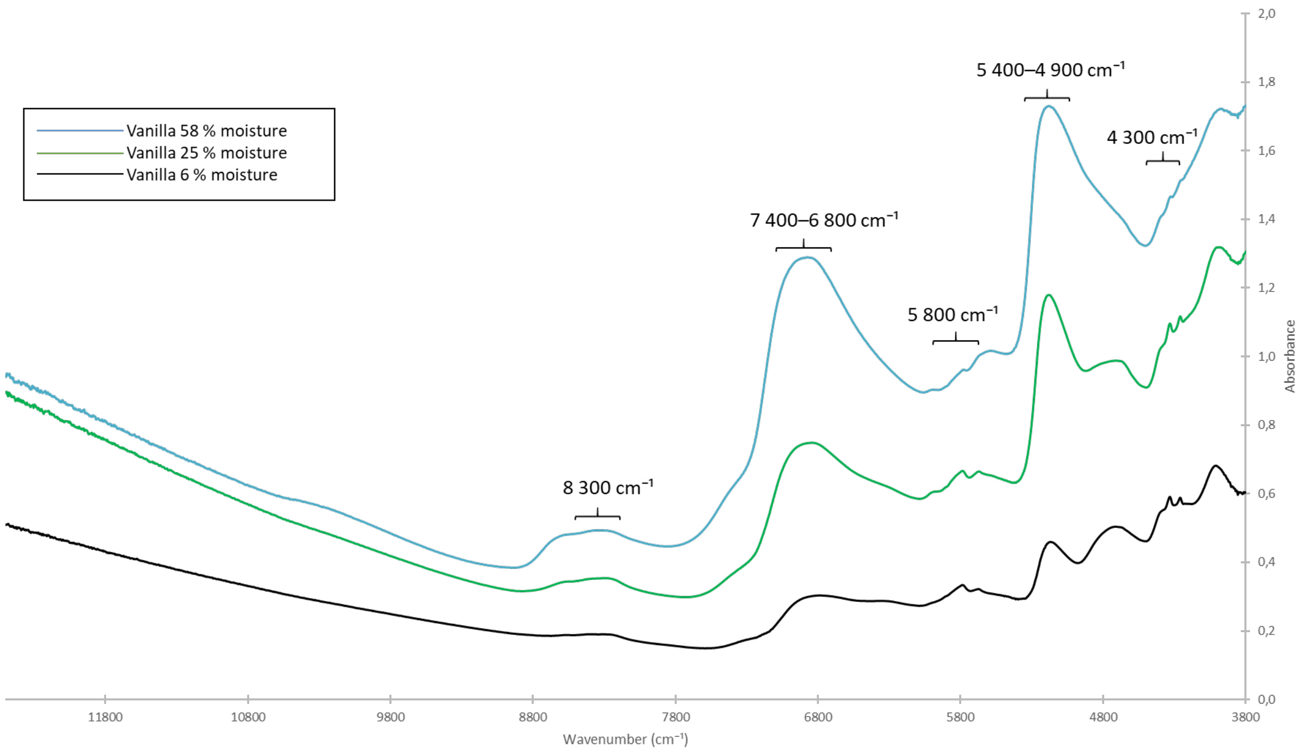

Figure 1: FT-NIR spectra of three samples with different water contents.

As shown in Figure 1, the higher the water content of a sample, the higher the absorbance values in the FT-NIR spectrum, especially within the specific water bands located in the 7 400–6 800 cm⁻¹ and 5 400–4 900 cm⁻¹ regions. In particular, for the sample with a water content of 58%, the strong contribution of the water bands clearly dominates the spectrum, almost completely masking other characteristic bands from the sample's constituents.

To analyze the data in greater detail, a principal component analysis (PCA) was performed, as shown in Figure 2(a), on all samples using reference data obtained from loss-on-drying measurements with a thermogravimetric balance. The PCA focused exclusively on the spectral region from 7 400 to 4 000 cm⁻¹ to remove the upper part of the spectral range, which does not provide significant information, and to better visualize the distribution of samples based on their water content. A clear distinction between samples according to their water content was observed, corresponding to a separation along the first principal component (PC1). This component alone explains 78% of the total variance in the data. It highlights a trend of increasing water content, ranging from samples with less than 10% moisture to those with high levels exceeding 40%. A more detailed analysis of Figure 2(b) highlights the first principal component associated with these spectra, showing that the two NIR bands characteristic of water are strongly correlated with the positive side of this component. This grouping includes the samples with the highest water contents. However, although these water bands are particularly prominent, making them an interesting spectral region for moisture modeling, focusing exclusively on these bands would lead to a substantial loss of useful information, because their broadness tends to dominate the spectrum and extend in a way that masks other bands.

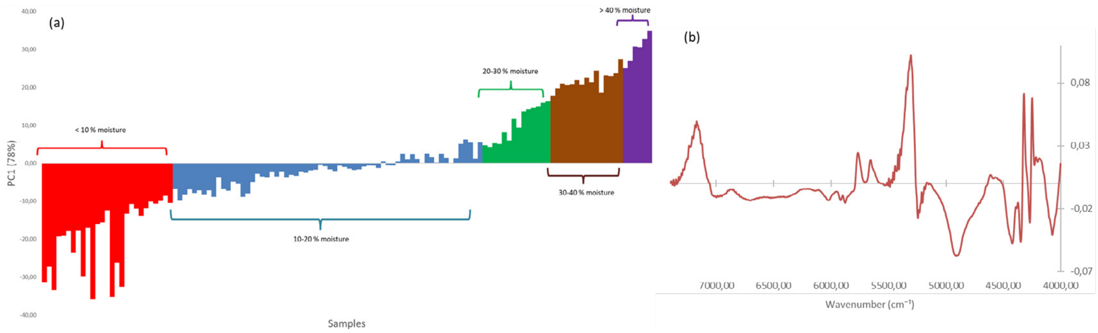

Figure 2: PCA from NIR spectra in the 7400–4000 cm⁻¹ regions of vanilla samples with different moisture contents obtained by weight loss through thermobalance (a), First principal component resulting from the PCA (b).

Based on these observations, the information extracted from the 7 400–4 000 cm⁻¹ spectral region was selected for the development of a PLS regression model. This model was specifically constructed to maximize the accuracy of water-content predictions across the different samples studied.

3.2. PLS Model

The PLS model was developed using a set of samples whose reference water content was determined by the loss-on-drying method using a thermogravimetric balance. The calibration set consisted of 82 samples covering a water-content range from 3.06% to 48.51%, providing representative diversity and a solid basis for building a robust model. To construct a PLS model capable of reliably predicting the water content of samples, it is necessary to determine the optimal number of latent variables. The PLS model was trained on a test set composed of 20 samples in order to determine and fix the optimal number of latent variables required for moisture prediction.

Table 1 presents the model's performance for an increasing number of latent variables (from 1 to 15), as well as the prediction performance on the test set, to identify the best compromise between calibration quality and predictive ability on an independent dataset. It was observed that as the number of latent variables increases, both calibration and prediction performances improve. This improvement reaches a satisfactory level starting from 4 latent variables, where the model and the test set show an excellent compromise, with maximum predictive ability and a prediction error below Table 1.

Latent variables (LV)	**Calibration**		**Test set**	
	$R^{2(a)}$	$RMSEC^{(b)}$	$Q^{2(a)}$	$RMSEP^{(c)}$
1	0.93	2.78	0.94	1.72
2	0.99	1.23	0.97	1.08
3	0.99	1.18	0.97	1.16
4	0.99	0.98	0.99	0.72
5	0.99	0.90	0.99	0.67
6	0.99	0.83	0.99	0.75
7	0.99	0.76	0.99	0.71
8	0.99	0.50	0.98	0.83
9	0.99	0.37	0.99	0.76
10	0.99	0.31	0.99	0.73
11	0.99	0.23	0.99	0.69
12	0.99	0.17	0.99	0.66
13	0.99	0.13	0.99	0.70
14	0.99	0.08	0.99	0.71
15	0.99	0.07	0.99	0.72
LV optimal = 4				

(a) Determination coefficient, (b) Root Mean Square Error of Calibration, (c) Root Mean Square Error of Prediction

Table 1: Determination of the optimal number of latent variables.

Although it might seem advantageous to choose a higher number of latent variables since the RMSEP continues to decrease at certain points (LV = 5, LV = 11, etc.), the differences in RMSEP between these points are not significant.

Moreover, beyond LV = 4, the prediction parameters fluctuate, which reflects a typical overfitting phenomenon. Thus, choosing LV = 4 is justified because it is mathematically and statistically preferable to retain the model with the minimum number of latent variables. It ensures a more reliable and generalizable model and prevents overfitting, providing a better balance between performance and robustness when applied to a dataset different from the one used to build the model. The optimization results show that the model reaches a coefficient of determination R^2 of 0.99 and an RMSEC of 0.98 using a total of 4 latent variables. To evaluate the predictive effectiveness of the constructed model, a separate group of samples was projected into the PLS model space to test its robustness.

3.3. Repetition Assessment

To validate the PLS model, a repeatability analysis was carried out. This step is essential to confirm the reliability and robustness of the model, as it assesses its ability to generate stable predictions when the same sample is measured repeatedly under identical conditions. Table 2 below presents the results of this analysis performed on a single sample.

Data set	Reference value %	Average %	Standard deviation %	Relative error %
30 spectra	17.51	18.47	0.62	5.48
10 averaged spectra	17.51	18.47	0.36	5.48

Table 2: Evaluation of the repetition of the prediction for the same sample.

The approach consisted of comparing the results obtained from 30 individual spectra with those obtained from 10 averaged spectra, each calculated from three single spectra. Using the 30 individual spectra, an estimated mean value of 18.47% is obtained, which is close to the reference value measured by thermogravimetric analysis (17.51%). The associated standard deviation (0.62) indicates limited data dispersion, while the relative error of 5.48% demonstrates very good predictive accuracy. When the spectra are grouped and averaged in sets of three, resulting in 10 mean spectra, the predicted value and the relative error remained unchanged. However, the standard deviation decreased (0.36), indicating a reduction in variability due to the smoothing effect of spectral averaging.

Overall, these observations highlight the remarkable repeatability of the PLS model: the predictions are stable, reliable, and consistent with the reference value. The improved stability provided by spectrum averaging underscores the value of this approach for reducing instrumental noise and enhancing prediction robustness, while maintaining an excellent level of accuracy.

3.4. Moisture Content on Prediction Set

It includes 20 samples, with known water contents ranging from 5.97% to 41.56%. This methodological consistency logically supports the quality of the predictions, as the samples used in this dataset have water-content values within the limits of the PLS model. The predicted values presented in Table 3 are in very good agreement with the reference values, showing small deviations, with performance confirmed by a Q^2 of 0.99 and an RMSEP of 0.82, still using 4 latent variables (LV = 4). These results demonstrate the model's ability to generalize its predictions when measurement methods are consistent between calibration and prediction.

Samples	Reference value %	Predicted value %
Vanilla 1	8.70	8.62
Vanilla 2	17.20	17.65
Vanilla 3	17.44	17.49
Vanilla 4	31.86	33.05
Vanilla 5	33.27	35.24
Vanilla 6	5.97	5.91
Vanilla 7	36.22	36.92
Vanilla 8	10.48	11.34
Vanilla 9	11.79	12.81
Vanilla 10	14.29	13.94
Vanilla 11	13.72	14.00
Vanilla 12	14.63	13.98
Vanilla 13	16.09	15.34
Vanilla 14	41.56	41.29
Vanilla 15	7.56	7.98
Vanilla 16	12.41	11.61
Vanilla 17	15.59	15.47
Vanilla 18	26.48	27.37
Vanilla 19	20.49	19.83
Vanilla 20	19.35	17.77
LV optimal = 4		

Table 3: Prediction on 20 other Sample

4. Conclusion

In conclusion, the PLS model developed to predict the water content of vanilla samples demonstrated good predictive performance, with an R^2 and Q^2 of 0.99, as well as an RMSEC of 0.98 and an RMSEP of 0.82, indicating its robustness and ability to generalize predictions across varied samples. Optimizing the number of latent variables to four allowed an optimal balance between accuracy and the avoidance of overfitting. The repeatability assessment, conducted on a single sample, confirmed the consistency and stability of the predictions. It is important to note that, although the water content measurement by loss-on-drying using a thermogravimetric balance and the NIR spectral acquisition were performed on the same vanilla sample, the exact portion of the ground or powdered material used for the two analyses was not identical. Furthermore, given the heterogeneity of water content values and the fact that the model contained more samples with moisture levels between 10% and 20%, the discrepancies observed between the two methods can be explained, particularly for values above 30%, due to their underrepresentation in the constructed model. Nevertheless, the predictions were consistent and reproducible for most samples. This highlights the potential for improvement, particularly by expanding the calibration database, as the samples used in this study were mostly from the species Vanilla planifolia. However, for Vanilla tahitensis, some vanilla beans can have much higher moisture contents (>70%). Expanding the model could allow the development of a more universal model applicable to all vanilla samples. These results emphasize the value of NIR spectroscopy for quality control, paving the way for industrial applications. The use of this technique provides two major advantages: on the one hand, it is non-destructive to the sample, and on the other hand, it also allows a significant time saving, since a complete analysis takes approximately 15 minutes (acquisition plus prediction), compared to nearly 40 minutes for a measurement using a thermogravimetric balance. It provides a reliable and sufficiently accurate tool to estimate the water content of vanilla, offering a rapid alternative to existing methods.

Funding

This work was supported by Vanalyse SARL, the Association nationale de la recherche et de la technologie (ANRT), Aix-Marseille University (AMU), Institut de Recherche pour le Développement (IRD) and Centre National pour la Recherche Scientifique (CNRS).

Credit Authorship Contribution Statement

Anthony Barreau: Writing – original draft, Investigation, Data curation, Formal analysis, Conceptualization. **Elodie Mezzatesta:** Writing – review & editing, Supervision, Conceptualization. **Sophie Charvet:** Writing – review, Supervision. **Philippe Faury:** Writing – review, Supervision. **Isabelle Bombarda:** Writing – review & editing, Supervision, Conceptualization. **Nathalie Dupuy:** Writing

– review & editing, Supervision, Conceptualization.

Declaration of Interest statement

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

The authors do not have permission to share data.

Declaration of generative AI in scientific writing

During the preparation of this work the author(s) used generative AI in order to improve the readability and language of the manuscript. After using this tool/service, the author(s) reviewed and edited the content as needed and take(s) full responsibility for the content of the published article.

References

1. Anuradha, K., Shyamala, B. N., & Naidu, M. M. (2013). Vanilla-its science of cultivation, curing, chemistry, and nutraceutical properties. *Critical reviews in food science and nutrition, 53*(12), 1250-1276.
2. Olatunde, A., Mohammed, A., Ibrahim, M. A., Tajuddeen, N., & Shuaibu, M. N. (2022). Vanillin: A food additive with multiple biological activities. *European Journal of Medicinal Chemistry Reports, 5,* 100055.
3. Karremans, A. P. (2024). A historical review of the artificial pollination of Vanilla planifolia: the importance of collaborative research in a changing world. *Plants, 13*(22), 3203.
4. Neimark, B., Osterhoudt, S., Blum, L., & Healy, T. (2021). Mob justice and 'The civilized commodity'. *The Journal of Peasant Studies, 48*(4), 734-753.
5. Van Dyk, S., McGlasson, W. B., Williams, M., Spooner-Hart, R., & Holford, P. (2024). Vanilla planifolia: Artificial and Insect Pollination, Floral Guides and Volatiles. *Plants, 13*(21), 2977.
6. Brillouet, J. M., Odoux, E., & Conejero, G. (2010). A set of data on green, ripening and senescent vanilla pod (Vanilla planifolia; Orchidaceae): anatomy, enzymes, phenolics and lipids. *Fruits, 65*(4), 221-235
7. Gassenmeier, K., Riesen, B., & Magyar, B. (2008). Commercial quality and analytical parameters of cured vanilla beans (Vanilla planifolia) from different origins from the 2006–2007 crop. *Flavour and Fragrance Journal, 23*(3), 194-201.
8. Ranadive, A. S. (Ed.). (2018). Quality control of vanilla beans and extracts. *Handbook of vanilla science and technology*, 237-260.
9. Feng, Y., Zhang, T., Yang, J., Liu, W., Yang, Y., Huang, J., ... & Zhou, Q. (2025). Characterization of microbial communities in flavors and fragrances during storage. *Frontiers in Microbiology, 16*, 1516594.
10. Peña-Barrientos, A., Perea-Flores, M. D. J., Martínez-Gutiérrez, H., Patrón-Soberano, O. A., González-Jiménez, F. E., Vega-Cuellar, M. Á., & Dávila-Ortiz, G. (2023). Physicochemical, microbiological, and structural relationship of vanilla beans (Vanilla planifolia, Andrews) during traditional curing process and use of its waste. *Journal of Applied Research on Medicinal and Aromatic Plants, 32*, 100445.
11. Liang, M., Wei, L., Li, B., Fu, Q., Cui, C., Zhang, W., ... & Ren, R. (2025). Effect of Drying Methods on the Drying Kinetics and Volatile Components of Vanilla Beans. *Journal of Food Processing and Preservation, 2025*(1), 1189632.
12. Prusinowska, R., & Smigielski, K. (2015). Losses of essential oils and antioxidants during the drying of herbs and spices. A review. *Nauki Inżynierskie i Technologie*, (2 (17).
13. Van Dyk, S., McGlasson, W. B., Williams, M., & Gair, C. (2010). Influence of curing procedures on sensory quality of vanilla beans. *Fruits, 65*(6), 387-399.
14. Wahyuningsih, R., & Fitriasari, B. (2022). Development of Vanilla Agribusiness and Its Export Opportunities To Support Triple Export Program (Gratitude) on Lombok Island. T*raektoriâ Nauki= Path of Science, 8*(1), 5019-5023.
15. Purwanto, Y. A., Widodo, S., & Iriani, E. S. (2024). Rapid assessment of vanilla (Vanilla planifolia) quality parameters using portable near-infrared spectroscopy combined with random forest. *Journal of Food Composition and Analysis, 133*, 106346.
16. Zambrano, M. V., Dutta, B., Mercer, D. G., MacLean, H. L., & Touchie, M. F. (2019). Assessment of moisture content measurement methods of dried food products in small-scale operations in developing countries: A review. *Trends in Food Science & Technology, 88,* 484-496.
17. Cozzolino, D. (2021). The ability of near infrared (NIR) spectroscopy to predict functional properties in foods: Challenges and opportunities. *Molecules, 26*(22), 6981.
18. Fodor, M., Matkovits, A., Benes, E. L., & Jókai, Z. (2024). The role of near-infrared spectroscopy in food quality assurance: A review of the past two decades. *Foods, 13*(21), 3501.
19. Manley, M. (2014). Near-infrared spectroscopy and hyperspectral imaging: non-destructive analysis of biological materials. *Chemical Society Reviews, 43*(24), 8200-8214.
20. Moll, V., Beć, K. B., Grabska, J., & Huck, C. W. (2022). Investigation of water interaction with polymer matrices by near-infrared (NIR) spectroscopy. *Molecules, 27*(18), 5882.
21. Rinnan, Å., Van Den Berg, F., & Engelsen, S. B. (2009). Review of the most common pre-processing techniques for near-infrared spectra. *TrAC Trends in Analytical Chemistry, 28*(10), 1201-1222.
22. Savitzky, A., & Golay, M. J. (1964). Smoothing and differentiation of data by simplified least squares procedures. *Analytical chemistry, 36*(8), 1627-1639.
23. Ivosev, G., Burton, L., & Bonner, R. (2008). Dimensionality reduction and visualization in principal component analysis. *Analytical chemistry, 80*(13), 4933-4944.
24. Diniz, P. H. G. D., Pistonesi, M. F., & Araújo, M. C. U. (2015). Using i SPA-PLS and NIR spectroscopy for the determination of total polyphenols and moisture in commercial tea samples. *Analytical Methods, 7*(8), 3379-3384.
25. Yuan, X., Xu, W., Wang, Y., Yang, C., & Gui, W. (2024). A

deep residual PLS for data-driven quality prediction modeling in industrial process. *IEEE/CAA Journal of Automatica Sinica, 11*(8), 1777-1785.
26. Galtier, O., Dupuy, N., Le Dréau, Y., Ollivier, D., Pinatel, C., Kister, J., & Artaud, J. (2007). Geographic origins and compositions of virgin olive oils determined by chemometric analysis of NIR spectra. *Analytica chimica acta, 595*(1-2), 136-144.
27. García-González, D. L., Baeten, V., Pierna, J. A. F., & Tena, N. (2013). Infrared, raman, and fluorescence spectroscopies: Methodologies and applications. In *Handbook of olive oil: Analysis and properties* (pp. 335-393). Boston, MA: Springer US.
28. Hourant, P., Baeten, V., Morales, M. T., Meurens, M., & Aparicio, R. (2000). Oil and fat classification by selected bands of near-infrared spectroscopy. *Applied spectroscopy, 54*(8), 1168-1174.