

Evaluating chemometric strategies and machine learning approaches for a miniaturized near-infrared spectrometer in plastic waste classification

Claudio Marchesi¹, Monika Rani¹, Stefania Federici¹, Matteo Lancini², Laura E. Depero¹

¹ Department of Mechanical and Industrial Engineering, University of Brescia & UdR INSTM of Brescia, via Branze 38, 25123 Brescia, Italy

² Department of Medical and Surgical Specialties, Radiological Sciences, and Public Health, University of Brescia, viale Europa 11, 25123 Brescia, Italy

ABSTRACT

Optimizing the sorting of plastic waste plays a crucial role in improving the recycling process. In this contribution, we report on a comparative study of multiple machine learning and chemometric approaches to categorize a data set derived from the analysis of plastic waste performed with a handheld spectrometer working in the Near-Infrared (NIR) spectral range. Conducting a cost-effective NIR study requires identifying appropriate techniques to improve commodity identification and categorization. Chemometric techniques, such as Principal Component Analysis (PCA) and Partial Least Squares - Discriminant Analysis (PLS - DA), and machine learning techniques such as Support- Vector Machines (SVM), fine tree, bagged tree, and ensemble learning were compared. Various pre-treatments were tested on the collected NIR spectra. In particular, Standard Normal Variate (SNV) and Savitzky-Golay derivatives as signal pre-processing tools were compared with feature selection techniques such as multiple Gaussian Curve Fit based on Radial Basis Functions (RBF). Furthermore, results were combined into a single predictor by using a likelihood-based aggregation formula. Predictive performances of the tested models were compared in terms of classification parameters such as Non-Error Rate (NER) and Sensitivity (Sn) with the analysis of the confusion matrices, giving a broad overview and a rational means for the selection of the approach in the analysis of NIR data for plastic waste sorting.

Section: RESEARCH PAPER

Keywords: Plastic waste sorting; Near-Infrared Spectroscopy (NIRS); circular economy; chemometrics; machine learning

Citation: Claudio Marchesi, Monika Rani, Stefania Federici, Matteo Lancini, Laura E. Depero, Evaluating chemometric strategies and machine learning approaches for a miniaturized near-infrared spectrometer in plastic waste classification, Acta IMEKO, vol. 12, no. 2, article 40, June 2023, identifier: IMEKO-ACTA-12 (2023)-02-40

Section Editor: Leonardo Iannucci, Politecnico di Torino, Italy

Received March 31, 2023; **In final form** June 19, 2023; **Published** June 2023

Copyright: This is an open-access article distributed under the terms of the Creative Commons Attribution 3.0 License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: The research was funded by the project PON "R&I" 2014-2020: SIRIMAP—Sistemi di Rilevamento dell'Inquinamento Marino da Plastiche e successivo recupero-riciclo (No. ARS01_01183, CUP D86C18000520008) and partially based upon work from COST Action CA20101 Plastics monitoring detection Remediation recovery - PRIORITY, supported by COST (European Cooperation in Science and Technology), www.cost.eu.

Corresponding authors: Stefania Federici, e-mail: stefania.federici@unibs.it; Matteo Lancini, e-mail: matteo.lancini@unibs.it

1. INTRODUCTION

In the perspective of whole-system economic sustainability, the enormous volume of urban plastic waste and the constant increase in human plastic consumption require a high level of waste valorisation. By the numbers, global plastic production reached 367 million tons in 2021, with Europe accounting for 16 % of the total [1]. 9 % of plastic was recycled, 12 % was incinerated, and 79 % ended up in landfills or natural compartments [2]. The recycling of polymer waste has significant environmental advantages owing to the replacement of primary

manufacturing, and waste sorting optimization plays a critical role in the development of the recycling process [3], [4]. Recycling is a technique for plastic product end-of-life waste management [5]. Basically, two types of recycling processes can be distinguished: mechanical and chemical processes [3], [6]. In both, sorting is the most critical stage in the recycling process, and this is true regardless of how effective the recycling program is [3], [4]. The use of automated sorting equipment makes the process more efficient [7]. Usually, these devices rely on vibrational spectroscopic techniques [8]-[11], and camera

systems for the polymer identification of clear and coloured products [5], [12]. Other techniques are based on ultraviolet (UV) spectroscopy [13], [14], X-ray [15], and hyperspectral imaging [16]-[18]. Over the years, this strategy has increased the purity of the output plastic, achieving a high percentage of recyclates in the production of secondary materials. However, these systems reach their limits with mixed plastics that require additional sorting elsewhere and can affect the quality of the recyclate if not appropriately allocated. A positive cost-benefit analysis is only possible if the separated polymer fractions have a high purity grade and satisfy the market demand for high-quality recyclates. Therefore, post-consumer recycling consists of many essential steps: collection, sorting, cleaning, size reduction and separation, and/or compatibilization to reduce polymer contamination [5]. In this scenario, the prospect of combining a well-established polymer identification technology with a small, portable, low-cost, real-time spectrometer for local and intermittent semi-automatic sorting is highly desirable, accompanied by robust data analysis [19], [20]. In recent years, chemometric analysis of non-destructive spectroscopic data has been widely investigated as an automated method for improving plastic sorting systems [21]-[24]. This improvement has been driven by the need to reduce the environmental impact [25]. Recently, machine learning has attracted considerable attention in plastic waste recognition using spectroscopic techniques [26]-[32]. In this study, we compared machine learning and chemometric techniques for classifying plastic waste data acquired with a portable Near-Infrared (NIR) spectrometer (see Figure 1 for the scheme of the work). Comparisons were made between chemometric approaches, Principal Component Analysis (PCA) and Partial Least Squares – Discriminant Analysis (PLS-DA), and machine learning techniques, Support-Vector Machines (SVM), Fine Tree, Bagged Tree, and Ensemble Learning. A comparison was also made in terms of pre-processing: traditional techniques, such as Standard Normal Variate (SNV) and Savitzky-Golay derivatives were examined in contrast to feature reduction techniques, such as multiple Gaussian Curve Fit based on Radial Basis Functions (RBF). The predictive performances of the tested models were compared in terms of classification parameters, such as Non-Error Rate (NER) and Sensitivity (S_n) with the analysis of confusion matrices, providing a comprehensive overview and a rational means of selecting the approach for the analysis of NIR data for plastic waste sorting.

2. MATERIALS AND METHODS

2.1. Samples collection

The first batch of plastic samples was collected in the Selection Division of the Montello SpA recovery and recycling plant (Bergamo, Italy), which accepts post-consumer plastic in the form of municipal waste for recycling [20]. Subsequently, the dataset was expanded to include new samples from municipal waste collected before ending up in landfills. A total of 325 samples from a variety of polymer classes were used in this study. Specifically, the products studied were: 75 samples of poly(ethylene terephthalate) (PET), 100 samples of polyethylene (PE), 75 samples of polypropylene (PP), and 75 samples of poly(styrene) (PS). The assortment included bottles, containers, and packaging of various sizes, shapes, and colours.

2.2. NIR analysis

Plastic samples were analysed using the MicroNIR On-site spectrometer (Viavi Solutions Inc., CA, United States) in reflectance mode without pre-treatment of the samples. The instrument is a palm-sized, portable spectrometer weighing approximately 250 g and measuring less than 200 mm in length and 50 mm in diameter. The instrument is equipped with a Linear Variable Filter (LVF), coupled to a linear detector array, which operates in the wavelength range 950-1650 nm. Control settings for spectral data acquisition were set to 10 milliseconds integration time and 50 scans, resulting in a short measurement time of 0.25 seconds. A point-and-shoot technique was used to perform 5 replicates for each sample to reduce the effects caused by sample non-uniformity. A total of 1625 spectra were acquired, and acquisition was performed using MicroNIR™ Pro v3.0 software (Viavi Solutions Inc., CA, United States).

2.3. Spectral pre-processing and chemometrics

Pre-processing NIR spectral data has become an essential aspect of chemometric modelling. The goal is to eliminate physical events from the spectra to improve subsequent multivariate regression, classification model, or exploratory analysis [33]. In this study, the spectra were retrieved in a single matrix of 1625×125 (samples \times wavenumbers) and pre-processing was applied using the Savitzky-Golay second derivative method with seven data points and a second order polynomial followed by Standard Normal Variate (SNV). The second derivative was applied to correct the drift effect [34], [35]

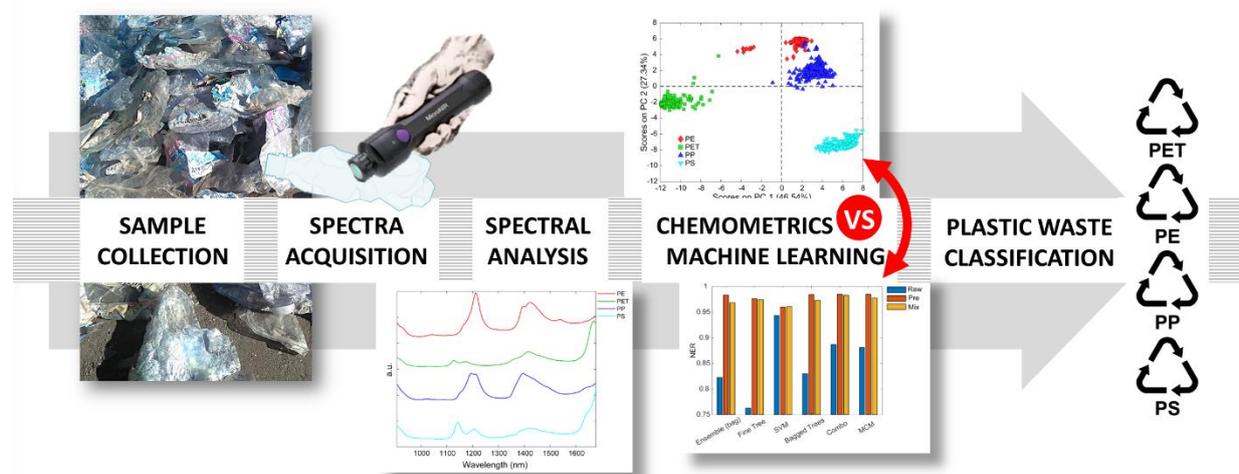


Figure 1. Scheme of the work.

in the NIR spectra, while SNV corrects the baseline shift [36]. SNV was calculated as follows [36]:

$$X_{\text{corr}} = \frac{X_{\text{org}} - a_0}{a_1} \quad (1)$$

where X_{corr} is the spectrum corrected, X_{org} is the raw spectrum collected by the instrument, a_0 is the value of the mean of the spectrum to be corrected, and a_1 is the standard deviation.

In addition, normalization was performed by mean centering. Different chemometric methods were used for the correct evaluation of the data of all analysed samples. PCA was initially applied as an exploratory analysis to investigate the data structure and was performed on 1625 NIR spectra from all polymer classes. Then, PLS-DA was applied as a supervised pattern recognition tool to separate the different commodities. Prior to using PLS-DA, data were split into a training set and a test set using a MATLAB proprietary function. The process was repeated 500 times, generating a different training and test set each time (75 % of the samples belonged to the training set and 25 % to the test set). All chemometric analyses were performed with MATLAB 2021b (The MathWorks, Inc, Natick, MA, USA) using the PLS-Toolbox (Eigenvector Research, Inc. Manson, Washington, USA).

2.4. Machine learning and pre-processing

Various machine learning algorithms were applied for classification purposes; SVM, Fine Tree, Ensemble Learning, and Bagged Tree. In addition, a likelihood-based aggregation procedure (here called Combo) was used to integrate the data into a single predictor, and the same procedure was applied with a Monte Carlo Method (MCM) to make a perturbation on raw data, to improve the generalization performance. The chosen hyperparameters are the following: for Fine Tree Gini's diversity index (gdi) was used as split criterion with 100 maximum number of splits; SVM was performed with a linear kernel function with kernel scale equal to 3. Lastly, Ensemble Learning was performed with the Bagged Tree method with 30 cycles of learning. To test the reliability of the system, 200 random extractions were performed for splitting the training and testing set. Again, 75 % of the samples were used for training and the rest for testing. Machine learning methods were performed on three different datasets: the raw data collected as specified in the previous paragraph (2.2), data reduced using the Gaussian RBF curve fit [37], and a dataset obtained combining raw and pre-processed data.

Each curve of the dataset has been fitted using a combination of 12 gaussian functions and a linear interpolation with a second-degree function, thus reducing the dataset dimension to 12 RBF centres and 12 sigma values. The procedure is as follows:

1. The second order derivative is computed and fed to find detection algorithm for the initial guesses of the RBF centres (here the MATLAB function "Findpeaks" was used with a limitation of 12 peaks maximum and excluding the first and last 20 samples of the spectrum).
2. A linear regression with a second-degree equation is used to remove offset and second-order trends.
3. The RBF centres are used as initial guess to an optimization procedure based on a Sequential Quadratic Programming constrained minimization function [38]. The cost function ε used is reported in (2) where A_i is the frequency of the i -th sample, y_i is its raw values, and μ_j ,

σ_j , and A_j are the centre, sigma and amplitude of the j -th RBF function respectively.

$$\varepsilon = \sum_i \sum_j \varepsilon_{i,j} \quad (2)$$

$$\varepsilon_{i,j} = \begin{cases} 0, & \sigma_j < 0 \\ A_j e^{-\frac{(f_i - \mu_j)^2}{2\sigma_j^2}}, & \sigma_j \geq 0 \end{cases}$$

4. The centres and sigmas found are collected as features of the new dataset.

The condition posed in (2) on the positive value allows to reduce dynamically the number of RBF functions actually used, while the interpolation removes trends that could hide peaks.

A third dataset combining the two previous dataset (raw and RBF Gaussian fit) is also created simply joining the two tables.

All calculations were performed using MATLAB and Statistics Toolbox release 2021b (The MathWorks, Inc, Natick, MA, USA). Automation of the procedure was implemented using MATLAB functions created in-house.

In Figure 2 the data analysis approach starting from raw data is reported, both for chemometrics and machine learning modelling.

3. RESULTS ANN DISCUSSION

3.1. NIR spectra

The main advantage of NIR spectroscopy is that it is a fast-response analytical technique capable of collecting spectra without prior processing and predicting physical and chemical properties from a single spectrum [39]. The absorption bands in the NIR region are caused by overtones and/or combination bands of primarily carbon-hydrogen vibrations and oxygen-hydrogen vibrations. Correct band assignment is difficult since it may be caused by various combinations of fundamental vibrations. Also, overtone vibrations are highly overlapping [40]. Representative NIR reflectance spectra of the four polymers (PE, PET, PP, and PS) are shown in Figure 3.

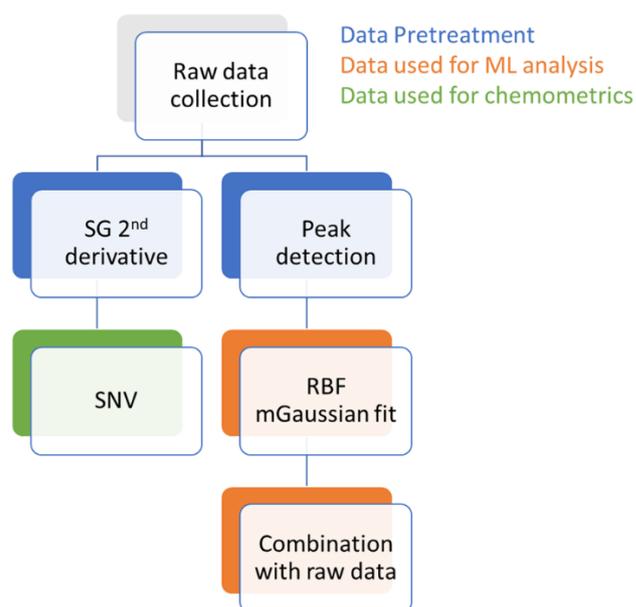


Figure 2. Data analysis approach.

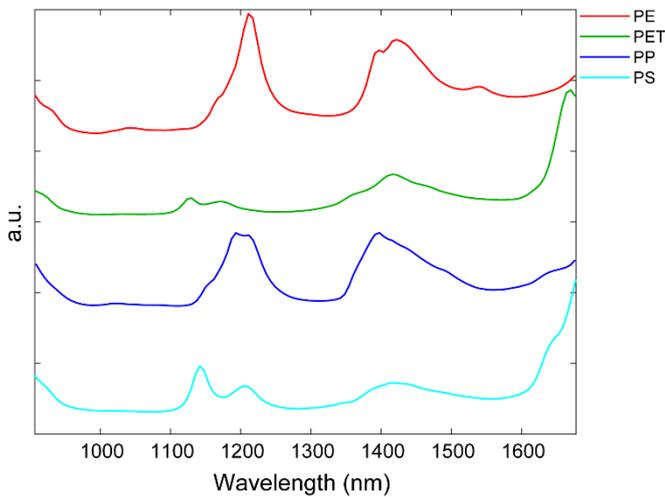


Figure 3. Representative near-infrared (NIR) spectra of the four classes of polymers.

The main absorbance band for PET was found at 1660 nm, which is related to the 1st overtone of C-H stretching [41], with other two peaks at about 1130 nm and 1415 nm. For PE the peak around 1211 nm is related to 2nd overtone of methylene C-H group, while the peak at about 1217 nm is related to the C-H stretch [42]. Peaks at 1391 nm and 1168 nm, correspond, respectively, to C-H combination band and 2nd overtone of CH₂ symmetric stretch. Regarding PP, the 2nd overtone of the asymmetric methyl C-H stretch is around 1193 nm, while the asymmetric methylene C-H stretch occurs at about 1211 nm [43]. The two peaks at 1391 nm and 1397 nm are related to methyl and methylene (C-H) combination. Lastly, for PS the peak at 1205 nm corresponds to the 2nd overtone of the aromatic C-H stretch; the stretching vibrational mode of C-H which occurs around 1639 nm, and the 1st overtone of aromatic C-H stretch overlaps with C-H combination band, which occurs at about 1391 nm [42]. To allow comparison between the raw spectra and the same spectra after applying the Savitzky-Golay 2nd derivative and SNV, Figure 4 shows the representative spectra of the four commodities after pre-processing.

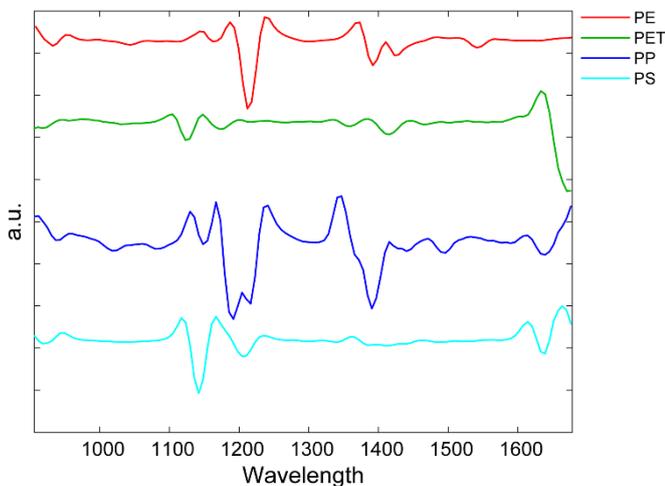


Figure 4. NIR spectra of the four classes of the polymers after the typical pre-processing for chemometric analysis: Savitzky-Golay 2nd derivative and Standard Normal Variate.

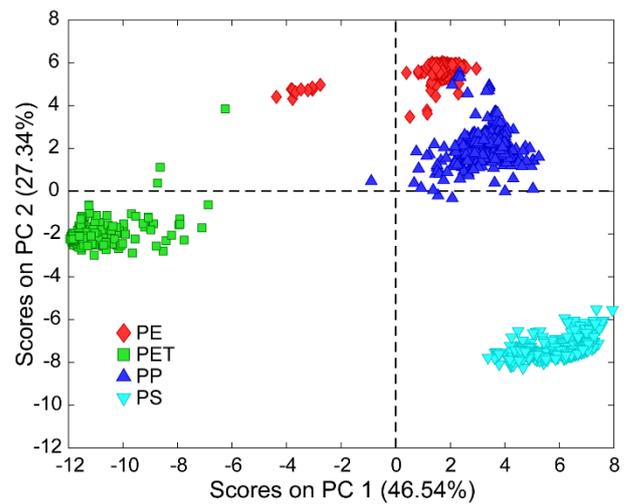


Figure 5. Results of PCA performed with spectral data of different commodities. The score plot of PC1 vs PC2 is presented.

3.2. Principal component analysis

The PCA calculation was performed after the pre-processing described above for the entire spectral range. For data structure analysis, PCA is a useful chemometric method. The goal of PCA is to extract the information stored in many variables into a smaller number of variables, called Principal Components [44]. Figure 5 shows the score plot of the first two components (73.88 % of the total explained variability), in which a clear separation between the polymer classes can be seen. Along PC1 PET is distinguished from the other commodities. PET samples show very negative score values, while the other samples show positive score values. On the other hand, along PC2, PS is clearly separated from the other plastics.

A clear separation between PP and PE can be noticed in the score plot of PC1 vs PC3 in Figure 6, where PC3 accounts for 15.83 % of the total information and explains the difference of PP from the other class of polymers.

3.3. Partial least squares discriminant analysis

Following the exploratory PCA analysis, a supervised classification technique was used to distinguish the different plastic groups. In PLS-DA, a classification objective is added to

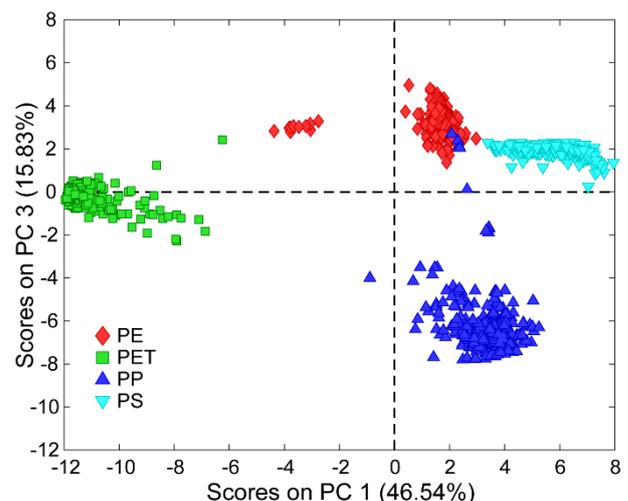


Figure 6. Results of PCA performed with spectral data of different commodities. The score plot of PC1 vs PC3 is presented.

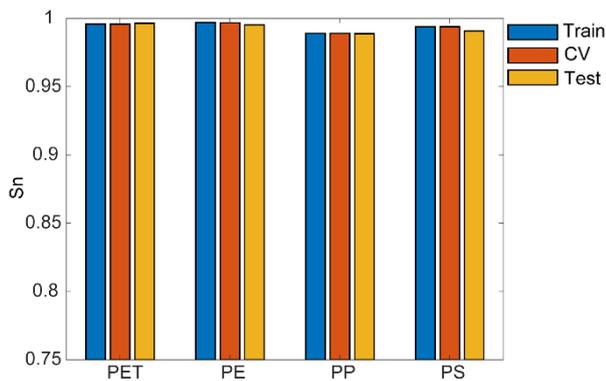


Figure 7. PLS-DA Model. Class sensitivities (Sn) calculated for training set, in cross-validation and for test set.

the PLS regression technique. The response variable is categorical and reflects the class to which the statistical units belong. PLS-DA returns the prediction as a vector with values between 0 and 1 and a length equal to the number of classes in the predictor variables [45], [46]. Each time PLS-DA was performed, the parameters such as NER and sensitivity were calculated in fitting, in cross-validation (CV), and for the test set. The cross-validation procedure was based on venetian blind approach with 5 groups. CV was also used to determine the optimal number of Latent Variables (LVs) for each PLS-DA model. Figure 7 shows all sensitivities for each class, calculated for training set, CV, and for test set. The values are close to 1, indicating a very high classification performance. Moreover, the results are very balanced between training, CV, and test set; therefore, overfitting is completely avoided, and the model can be considered reliable and stable.

Table 1 shows the NER defined as mean class sensitivity [47], calculated for all the training set, cross-validation, and test set. Overall, 99 % of the samples were correctly classified for each of the 500 iterations.

3.4. Machine learning

Due to the complexity and the large number of results, for the machine learning analysis the classification parameters are presented only for the test set. Figure 8 shows the NER of the classes for each computed model and for each treatment of the data. It is noticeable that the models run on raw data have the worst performances. The NER ranges from 0.74 (Fine Tree) to 0.9 (SVM), indicating a high variability in the results. For raw data only SVM can be considered as a satisfactory model for pattern recognition. Lower variability in the results is observed for pre-treated data and for a mixture of pre-treated and raw data, where the NER ranges from 0.96 to 0.99 and from 0.96 to 0.98, respectively. Thus, there is no difference in the results between pre-processed data and the combination of raw and pre-treated data. These results confirm that feature reduction based on the Gaussian curve with RBF gives high performances for pattern recognition in machine learning analysis.

Table 1. PLS-DA Model. Non-Error Rate calculated for training set, in CV and for test set.

	NER
Training	0.99
CV	0.99
Test	0.99

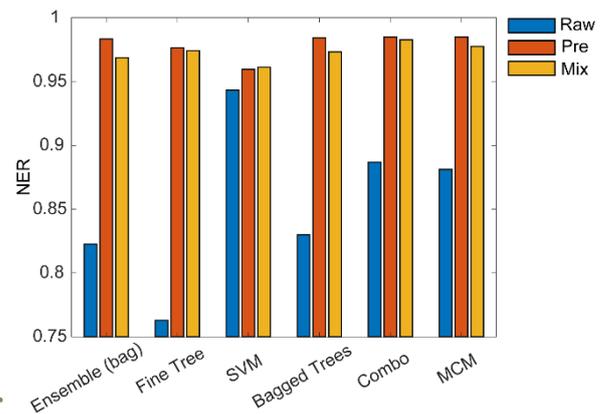


Figure 8. Machine Learning. Comparison of the Non-Error Rate (NER) calculated from the confusion matrices for each model. Results are presented for raw data, pre-treated data, and the combination of raw and pre-treated data.

In general, the model performance is comparable between machine learning and multivariate analysis methods. After random extraction of training and test data repeated 500 and 200 times for chemometrics and machine learning, respectively, the NER calculated for the test set is above 0.95 for both methods. However, the use of chemometrics reduces the computational time, compared to the computationally intensive machine learning algorithms.

4. CONCLUSION

This paper included a side-by-side comparison between conventional chemometric methods and machine learning algorithms for the classification of a dataset obtained from the study of plastic waste with a portable Near-Infrared (NIR) spectrometer. Multivariate methods such as Principal Component Analysis (PCA) and Partial Least Squares - Discriminant Analysis (PLS - DA) were investigated, as well as machine learning methods such as Support Vector Machines (SVM), Fine Tree, Bagged Tree and Ensemble Learning. Results were also compared in terms of data processing: signal pre-processing tools, SNV, and Savitzky-Golay derivatives were compared with feature reduction approaches such as Multiple Gaussian Curve Fit based on Radial Basis Functions (RBF). In addition, the machine learning algorithms were run on raw data, pre-processed data, and the combination of the two approaches. The results from PLS-DA showed very high performances for pattern recognition; in fact, the NER for the training set, in CV, and for the test set are all equal to 0.99. In contrast, for machine learning, the NER for raw data ranges from 0.74 for Fine Tree to 0.90 for SVM, indicating high variability in the results. The results for the pre-processed data show lower variability with NER value ranging from 0.96 to 0.99, which is also valid for the combination of raw data and pre-processed data. This confirms that RBF-based variable reduction is the most crucial point to improve classification performances. Contrarily to some results found in the literature regarding the pre-treatment of data having a negative effect in accuracy using chemometrics [48], the pre-treatment of data is generally an improvement in the detection accuracy using machine learning techniques. We can conclude that the multivariate and machine learning approaches produce comparable results in terms of model performance. The NER estimated for the test set is above 0.95 for both chemometrics and machine learning after randomly extracting the training and

test data and repeating them 500 and 200 times, respectively. On the other hand, chemometrics is characterised by a lower computation time compared to machine learning algorithms and it can therefore be considered more advantageous.

REFERENCES

- [1] Plastics Europe. Plastics Europe website. Online [Accessed 25 June 2023] <http://www.plasticseurope.org>
- [2] R. Geyer, J. R. Jambeck, K. L. Law, Production, use, and fate of all plastics ever made, *Science Advances* vol. 3 (2017) no. 7. DOI: [10.1126/sciadv.1700782](https://doi.org/10.1126/sciadv.1700782)
- [3] S. M. Al-Salem, P. Lettieri, J. Baeyens, Recycling and recovery routes of plastic solid waste (PSW): A review, *Waste Management* vol. 29 (2009) no.10, pp. 2625–2643. DOI: [10.1016/j.wasman.2009.06.004](https://doi.org/10.1016/j.wasman.2009.06.004)
- [4] R. Siddique, J. Khatib, I. Kaur, Use of recycled plastic in concrete: A review, *Waste Management* vol.28 (2008) no.10, pp. 1835-1852. DOI: [10.1016/j.wasman.2007.09.011](https://doi.org/10.1016/j.wasman.2007.09.011)
- [5] J. Hopewell, R. Dvorak, E. Kosior, Plastics recycling: Challenges and opportunities, *Philosophical Transactions of the Royal Society B: Biological Sciences* vol. 364 (2009) no. 1526, pp. 2115-2126. DOI: [10.1098/rstb.2008.0311](https://doi.org/10.1098/rstb.2008.0311)
- [6] K. Ragaert, L. Delva, K. Van Geem, Mechanical, and chemical recycling of solid plastic waste, *Waste Management* vol. 69 (2017), pp. 24-58. DOI: [10.1016/j.wasman.2017.07.044](https://doi.org/10.1016/j.wasman.2017.07.044)
- [7] S. P. Gundupalli, S. Hait, A. A. Thakur, A review on automated sorting of source-separated municipal solid waste for recycling, *Waste Management* vol. 60 (2017), pp. 56-74. DOI: [10.1016/j.wasman.2016.09.015](https://doi.org/10.1016/j.wasman.2016.09.015)
- [8] V. Allen, J. H. Kalivas, R. G. Rodriguez, Post-consumer plastic identification using Raman spectroscopy, *Applied Spectroscopy* vol. 53 (1999) no. 6, pp. 672–681. DOI: [10.1366/000370299194732](https://doi.org/10.1366/000370299194732)
- [9] V. Ludwig, Z. Da Costa Ludwiga, M. M. Rodrigues, V. Anjos, C. Batestin Costa, D. R. Sant'Anna das Dores, V. R. Da Silva, F. Soares, Analysis by Raman and infrared spectroscopy combined with theoretical studies on the identification of plasticizer in PVC films, *Vibrational Spectroscopy* vol. 98 (2018), pp. 134-138. DOI: [10.1016/j.vibspec.2018.08.004](https://doi.org/10.1016/j.vibspec.2018.08.004)
- [10] O. Rozenstein, E. Puckrin, J. Adamowski, Development of a new approach based on midwave infrared spectroscopy for post-consumer black plastic waste sorting in the recycling industry, *Waste Management* vol. 68 (2017), pp. 38–44. DOI: [10.1016/j.wasman.2017.07.023](https://doi.org/10.1016/j.wasman.2017.07.023)
- [11] A. Vázquez-Guardado, M. Money, N. McKinney, D. Chanda, Multi-spectral infrared spectroscopy for robust plastic identification, *Appl. Optics* vol. 54 (2015) no. 24, pp. 7396-7405. DOI: [10.1364/AO.54.007396](https://doi.org/10.1364/AO.54.007396)
- [12] Y. Tachwali, Y. Al-Assaf, Y. & A. R. Al-Ali, Automatic multistage classification system for plastic bottles recycling, *Resources, Conservation and Recycling* vol. 52 (2007) no. 2, pp. 266–285. DOI: [10.1016/j.resconrec.2007.03.008](https://doi.org/10.1016/j.resconrec.2007.03.008)
- [13] E. Maris, A. Aoussat, E. Naffrechoux, D. Froelich, Polymer tracer detection systems with UV fluorescence spectrometry to improve product recyclability, *Minerals Engineering* vol. 29 (2012), pp. 77–88. DOI: [10.1016/j.mineng.2011.09.016](https://doi.org/10.1016/j.mineng.2011.09.016)
- [14] S. M. Safavi, H. Masoumi, S. Mirian, M. Tabrizchi, Sorting of polypropylene resins by color in MSW using visible reflectance spectroscopy, *Waste Management* vol. 30 (2010) no. 11, pp. 2216–2222. DOI: [10.1016/j.wasman.2010.06.023](https://doi.org/10.1016/j.wasman.2010.06.023)
- [15] S. Brunner, P. Fomin, C. Kargel, Automated sorting of polymer flakes: Fluorescence labeling and development of a measurement system prototype, *Waste Management* vol. 38 (2015), pp. 49–60. DOI: [10.1016/j.wasman.2014.12.006](https://doi.org/10.1016/j.wasman.2014.12.006)
- [16] Y. Zheng, J. Bai, J. Xu, X. Li, Y. Zhang, A discrimination model in waste plastics sorting using NIR hyperspectral imaging system, *Waste Management* vol. 72 (2018), pp. 87–98. DOI: [10.1016/j.wasman.2017.10.015](https://doi.org/10.1016/j.wasman.2017.10.015)
- [17] M. Vidal, A. Gowen, J. M. Amigo, NIR Hyperspectral Imaging for Plastics Classification, *NIR news* vol. 23 (2012) no.1, pp. 13–15. DOI: [10.1255/nirn.1285](https://doi.org/10.1255/nirn.1285)
- [18] M. Moroni, A. Mei, A. Leonardi, E. Lupo, F. La Marca, PET and PVC separation with hyperspectral imagery, *Sensors* vol. 15 (2015) no. 1, pp. 2205–2227. DOI: [10.3390/s150102205](https://doi.org/10.3390/s150102205)
- [19] I. Vollmer, M. J. F. Jenks, M. C. P. Roelands, R. J. White, T. van Harmelen, G. P. van der Laan, F. Meirer, J. T. F. Keurenjes, B. M. Weckhuysen, Beyond Mechanical Recycling: Giving New Life to Plastic Waste, *Angewandte Chemie - International Edition* vol. 59 (2020) no.36, pp. 15402–15423. DOI: [10.1002/anie.201915651](https://doi.org/10.1002/anie.201915651)
- [20] M. Rani, C. Marchesi, S. Federici, G. Rovelli, I. Alessandri, I. Vassalini, S. Ducoi, L. Borgese, A. Zacco, F. Bilo, E. Bontempi, L. E. Depero, Miniaturized Near-Infrared (MicroNIR) Spectrometer in Plastic Waste Sorting, *Materials* vol. 12 (2019) no. 17. DOI: [10.3390/ma12172740](https://doi.org/10.3390/ma12172740)
- [21] R. Junjuri, M. K. Gundawar, A low-cost LIBS detection system combined with chemometrics for rapid identification of plastic waste, *Waste Management* vol. 117 (2021), pp. 48–57. DOI: [10.1016/j.wasman.2020.07.046](https://doi.org/10.1016/j.wasman.2020.07.046)
- [22] V. C. Costa, F. W. B. Aquino, C. M. Paranhos, E. R. Pereira-Filho, Identification and classification of polymer e-waste using laser-induced breakdown spectroscopy (LIBS) and chemometric tools, *Polymer Testing* vol. 59 (2017), pp. 390–395. DOI: [10.1016/j.polymertesting.2017.02.017](https://doi.org/10.1016/j.polymertesting.2017.02.017)
- [23] G. Bonifazi, L. Fiore, R. Gasbarrone, P. Hennebert, P. S. Serranti, Detection of brominated plastics from e-waste by short-wave infrared spectroscopy, *Recycling* vol. 6 (2021) no. 3. DOI: [10.3390/recycling6030054](https://doi.org/10.3390/recycling6030054)
- [24] E. R. K. Neo, Z. Yeo, J. S. C. Low, V. Goodship, K. Debattista, A review on chemometric techniques with infrared, Raman and laser-induced breakdown spectroscopy for sorting plastic waste in the recycling industry, *Resources, Conservation and Recycling* vol. 180 (2022) 106217. DOI: [10.1016/j.resconrec.2022.106217](https://doi.org/10.1016/j.resconrec.2022.106217)
- [25] C. Araujo-Andrade, E. Bugnicourt, L. Philippet, L. Rodriguez-Turienzo, D. Nettleton, L. Hoffmann, M. Schlummer, Review on the photonic techniques suitable for automatic monitoring of the composition of multi-materials wastes in view of their posterior recycling, *Waste Management & Research* vol. 39 (2021) no. 5, pp. 631–651. DOI: [10.1177/0734242X21997908](https://doi.org/10.1177/0734242X21997908)
- [26] V. Da Silva, H. Murphy, J. M. Amigo, C. Stedmon, J. Strand, Classification and Quantification of Microplastics (<100 μm) Using a Focal Plane Array-Fourier Transform Infrared Imaging System and Machine Learning, *Analytical Chemistry* vol. 92 (2020) no. 20, pp. 13724–13733. DOI: [10.1021/acs.analchem.0c01324](https://doi.org/10.1021/acs.analchem.0c01324)
- [27] S. Zhu, H. Chen, M. Wang, X. Guo, Y. Lei, G. Jin, Plastic solid waste identification system based on near infrared spectroscopy in combination with support vector machine, *Advanced Industrial and Engineering Polymer Research* vol. 2 (2019) no.2, pp. 77–81. DOI: [10.1016/j.aiepr.2019.04.001](https://doi.org/10.1016/j.aiepr.2019.04.001)
- [28] A. P. M. Michel, A. E. Morrison, V. L. Preston, C. T. Marx, B. C. Colson, H. K. White, Rapid Identification of Marine Plastic Debris via Spectroscopic Techniques and Machine Learning Classifiers, *Environmental Science & Technologies* vol. 54 (2020) no. 17, pp. 10630–10637. DOI: [10.1021/acs.est.0c02099](https://doi.org/10.1021/acs.est.0c02099)
- [29] Y. Yang, W. Zhang, Z. Wang, Y. Li, Differentiation of Plastics by Combining Raman Spectroscopy and Machine Learning, *Journal of Applied Spectroscopy* vol. 89 (2022) no.4, pp. 790–798. DOI: [10.1007/s10812-022-01426-1](https://doi.org/10.1007/s10812-022-01426-1)

- [30] B. Carrera, V. L. Piñol, J. B. Mata, K. Kim, A machine learning based classification models for plastic recycling using different wavelength range spectrums, *Journal of Cleaner Production* vol.374 (2022) 133883.
DOI: [10.1016/j.jclepro.2022.133883](https://doi.org/10.1016/j.jclepro.2022.133883)
- [31] D. Covarrubias-Martínez, H. Lobato-Morales, J. M. Ramírez-Cortés, G. A. Álvarez-Botero, Classification of plastic materials using machine-learning algorithms and microwave resonant sensor, *Journal of Electromagnetic Waves and Applications* vol. 36 (2022) no. 12, pp. 1760–1775.
DOI: [10.1080/09205071.2022.2043192](https://doi.org/10.1080/09205071.2022.2043192)
- [32] S. Zinchik, S. Jiang, S. Friis, F. Long, L. Høgstvedt, V. M. Zavala, E. Bar-Ziv, Accurate Characterization of Mixed Plastic Waste Using Machine Learning and Fast Infrared Spectroscopy, *ACS Sustainable Chemistry & Engineering* vol. 9 (2021) no. 42, pp. 14143-14151.
DOI: [10.1021/acssuschemeng.1c04281](https://doi.org/10.1021/acssuschemeng.1c04281)
- [33] A. Rinnan, F. van den Berg, S. B. Engelsen, Review of the most common pre-processing techniques for near-infrared spectra, *TrAC Trends in Analytical Chemistry* vol. 28 (2009) no. 10, pp. 1201–1222.
DOI: [10.1016/j.trac.2009.07.007](https://doi.org/10.1016/j.trac.2009.07.007)
- [34] P. Oliveri, C. Malegori, R. Simonetti, M. Casale, The impact of signal pre-processing on the final interpretation of analytical outcomes – A tutorial, *Analytica Chimica Acta* vol. 1058 (2019), pp. 9–17.
DOI: [10.1016/j.aca.2018.10.055](https://doi.org/10.1016/j.aca.2018.10.055)
- [35] V. M. Taavitsainen, Denoising and Signal-to-Noise Ratio Enhancement: Derivatives. in *Comprehensive Chemometrics*, eds. Brown, S. D., Tauler, R. & Walczak, B., pp. 57–66, Elsevier, 2009, ISBN 978-0-444-52701-1.
DOI: [10.1016/B978-0-444-52701-1.00101-0](https://doi.org/10.1016/B978-0-444-52701-1.00101-0)
- [36] R. J. Barnes, M. S. Dhanoa, S. J. Lister, Standard normal variate transformation and de-trending of near-infrared diffuse reflectance spectra, *Applied Spectroscopy* vol. 43 (1989) no. 5, pp. 772–777.
DOI: [10.1366/000370289420220](https://doi.org/10.1366/000370289420220)
- [37] M. M. Li, The development of a nonlinear curve fitter using RBF neural networks with hybrid neurons. *Lecture Notes in Computer Science*, part of 13th International Symposium on Neural Networks, ISNN 2016, St. Petersburg, Russia, July 6-8, 2016, *Proceedings* vol. 9719, 2016, pp. 434–443.
DOI: [10.1007/978-3-319-40663-3_50](https://doi.org/10.1007/978-3-319-40663-3_50)
- [38] J. Frédéric Bonnans, J. Charles Gilbert, C. Lemaréchal, C. A. Sagastizábal, *Numerical Optimization: Theoretical and Practical Aspects*, Springer Berlin Heidelberg, 2006, ISBN 978-3-540-35447-5.
DOI: [10.1007/978-3-540-35447-5](https://doi.org/10.1007/978-3-540-35447-5)
- [39] M. Blanco, M. I. Villarroya, NIR spectroscopy: a rapid-response analytical tool, *TrAC Trends in Analytical Chemistry* vol. 21 (2002) no. 4, pp. 240–250.
DOI: [10.1016/S0165-9936\(02\)00404-1](https://doi.org/10.1016/S0165-9936(02)00404-1)
- [40] J. Workman Jr., L. Weyer, *Practical Guide and Spectral Atlas for Interpretive Near-Infrared Spectroscopy*, CRC Press, Boca Raton, 2006, ISBN 9781439875254.
DOI: [10.1201/b11894](https://doi.org/10.1201/b11894)
- [41] E. W. Crandall, A. N. Jagtap, The near-infrared spectra of polymers. *Journal of Applied Polymer Science* vol. 21 (1977), pp. 449–454.
DOI: [10.1002/app.1977.070210211](https://doi.org/10.1002/app.1977.070210211)
- [42] J. Workman Jr., *The Handbook of Organic Compounds, NIR, IR, R, and UV-Vis Spectra Featuring Polymers and Surfactants*, Academic Press, 2001, ISBN 13: 9780127635613.
DOI: [10.1016/B978-0-12-763560-6.X5000-4](https://doi.org/10.1016/B978-0-12-763560-6.X5000-4)
- [43] S. Zhu, Z. Song, S. Shi, M. Wang, M. G. Jin, Fusion of Near-Infrared and Raman Spectroscopy for In-Line Measurement of Component Content of Molten Polymer Blends, *Sensors* vol. 19 (2019) no.16, 3643.
DOI: [10.3390/s19163463](https://doi.org/10.3390/s19163463)
- [44] D. Ballabio, A MATLAB toolbox for Principal Component Analysis and unsupervised exploration of data structure, *Chemometrics and Intelligent Laboratory Systems* vol. 149 (2015), pp. 1–9.
DOI: [10.1016/j.chemolab.2015.10.003](https://doi.org/10.1016/j.chemolab.2015.10.003)
- [45] D. Ballabio, V. Consonni, Classification tools in chemistry. Part 1: Linear models. PLS-DA, *Analytical Methods* vol. 5 (2013), pp. 3790–3798.
DOI: [10.1039/C3AY40582F](https://doi.org/10.1039/C3AY40582F)
- [46] R. G. Brereton, G.R. Lloyd, Partial least squares discriminant analysis: Taking the magic away, *Journal of Chemometrics* vol. 28 (2014), pp. 213–225.
DOI: [10.1002/cem.2609](https://doi.org/10.1002/cem.2609)
- [47] D. Ballabio, F. Grisoni, R. Todeschini, Multivariate comparison of classification performance measures, *Chemometrics and Intelligent Laboratory Systems* vol. 174 (2018), pp. 33–44.
DOI: [10.1016/j.chemolab.2017.12.004](https://doi.org/10.1016/j.chemolab.2017.12.004)
- [48] P. Mishra, D. N. Rutledge, J. M. Roger, K. Wali, H. A. Khan, Chemometric pre-processing can negatively affect the performance of near-infrared spectroscopy models for fruit quality prediction, *Talanta* vol. 229 (2021) 122303.
DOI: [10.1016/j.talanta.2021.122303](https://doi.org/10.1016/j.talanta.2021.122303)