Cluster-based identification algorithm for in-line recycled concrete aggregates characterization using Laser-Induced Breakdown Spectroscopy (LIBS)

Chang, Cheng; Maio, Francesco Di; Rem, Peter; Gebremariam, Abraham T.; Mehari, Fanuel; Xia, Han

DOI
10.1016/j.resconrec.2022.106507

Publication date
2022

Document Version
Final published version

Published in
Resources, Conservation and Recycling

Resources, Conservation & Recycling 185 (2022) 106507

Contents lists available at ScienceDirect

Resources, Conservation & Recycling

journal homepage: www.elsevier.com/locate/resconrec

Full length article

Cluster-based identification algorithm for in-line recycled concrete aggregates characterization using Laser-Induced Breakdown Spectroscopy (LIBS)

Cheng Chang [a,*], Francesco Di Maio [a], Peter Rem [a], Abraham T. Gebremariam [a], Fanuel Mehari [b], Han Xia [c]

[a] Resource & recycling, Department of Engineering Structures, Faculty of Civil Engineering and Geosciences, Delft University of Technology, Stevinweg 1, 2628 CN, Delft, The Netherlands
[b] SPECTRAL Industries BV, Olof Palmestraat 14, 2616 LR, Delft, The Netherlands
[c] ASML Netherlands BV, De Run 6665, 5504 DT, Veldhoven, The Netherlands

ARTICLE INFO

Keywords:
End-of-Life (EoL) concrete
Laser-induced breakdown spectroscopy (LIBS)
Principal component analysis (PCA)
Classification
Chi-square distribution

ABSTRACT

To upcycle End-of-Life (EoL) concrete from demolished buildings, it is essential to efficiently identify the different materials that may contaminate it. The precise identification and classification of materials and contaminants are vital processes for in-line quality inspection of recycled concrete aggregates transported on a conveyor belt. In this study, a total of eight potential contaminants are considered as target contaminant materials in the streams made of coarse and fine aggregates resulting from the upcycling of EoL concrete. These contaminants degrade the quality of the aggregates even at low concentrations, so it is essential to identify the presence of such contaminants along with the main products of recycling which are recycled coarse aggregates (RCA) and recycled fine aggregates (RFA). An efficient method is proposed to identify and classify EoL concrete waste along with RCA and RFA in motion on conveyor belts via laser-induced breakdown spectroscopy (LIBS) coupled with a cluster-based identification algorithm. The model is verified with an accuracy of 0.97, a precision (weighted average) of 0.98, a recall (weighted average) of 0.97, and an F1-score (weighted average) of 0.98 for the validation set, under the optimal conditions. This study suggests that LIBS may be well suited for fast and in-line analysis of recycled concrete aggregates in industrial applications. This approach presents an innovative approach for the quality characterization of secondary materials produced from EoL concrete being transported on conveyor belts, and therefore can be of great value for the processing and high-end utilization of EoL concrete.

1. Introduction

Concrete has long been one of the most popular manufactured construction materials. In the conventional production process, the concrete is usually made using cement and natural aggregates that have well-defined and predictable properties. Therefore, it is possible to foresee the mechanical and durability properties of the produced concrete. In contrast, when concrete is made using recycled aggregates, it is impossible to predict the resulting concrete's mechanical and durability properties because recycled aggregates have variable properties. That is why it is challenging to upcycle End-of-Life (EoL) concrete and close the material loop. Because recycled aggregates are often blended with other construction waste materials, it commonly serves for low-level construction, for example, embankment, sub-base, and leveling of roads (Vegas et al., 2015).

A significant amount of construction work carried out in the 1950s during the post-World War II economic boom is reaching life expectancy in the next few decades, which will lead to a rapid increase in construction and demolition wastes (C&DW), particularly in Europe. A large amount of C&DW cannot be efficiently recycled and is even dumped directly in landfills, causing environmental pollution (Kabirifar et al., 2021; Nanda and Berruti, 2021). Meanwhile, it is expected that the demand for concrete will rise in the coming years, particularly in developing countries (Bonifazi et al., 2018). The gap between supply and demand for concrete will lead to the consumption of large amounts of resources, and the over-mining of raw materials for concrete also adds

* Corresponding author at: Mr. Cheng Chang, TUDelft: Technische Universiteit Delft, Netherlands.
E-mail addresses: chang-cheng@outlook.com, C.Chang-1@tudelft.nl (C. Chang).

https://doi.org/10.1016/j.resconrec.2022.106507
Received 8 February 2022; Received in revised form 22 June 2022; Accepted 30 June 2022
Available online 9 July 2022

to the damage to the environment. EoL concrete accounts for the vast majority of C&DW (Lotfi and Rem, 2016), and the most viable solution for EoL concrete is recycling or upcycling. The conventional linear approach to recycling needs to be upgraded to a circular process, that is, secondary raw materials are obtained from EoL concrete for a green and sustainable solution (Cossu and Williams, 2015). At present, one of the most popular methods for high-grade concrete recycling is the wet process, which produces clean concrete aggregates by washing coarse aggregates but also produces sludge that needs to be disposed of (Zhang et al., 2019). In addition, an innovation project called C2CA (Concrete to Cement and Aggregate, www.c2ca.eu), funded by the European Commission (EC), proposes a dry alternative to the existing wet process by offering an innovative solution called Advanced Dry Recovery (ADR) (Gebremariam et al., 2020). This solution significantly reduces the cost of processing the coarse fraction of high-grade recycled EoL concrete. The complete recycling of EoL concrete can close the building and demolition lifecycle and is of great benefit to the environment in terms of reducing the depletion of natural resources, noise pollution, energy consumption, and dust and gases emissions (Di Maria et al., 2016).

EoL concrete is a material with a highly variable composition. Its composition results from the original application and recipe of primary concrete, how the materials are connected to the building, and the care and measures are taken when the structure is disassembled and dismantled. The main challenge is safeguarding the quality of the secondary aggregates resulting from the recycling of the EoL concrete. It is challenging to keep the demolished concrete as pure material, and it is usually mixed with other building materials such as bricks, cement paste, foam, glass, gypsum, mineral fibers, plastics, and wood, all of which are considered waste and can have impacts on the quality of the resulting recycled concrete. This implies that special technical and organizational means are required to ensure that the recycled concrete has the same quality as primary concrete, despite the problems mentioned above.

Recycled aggregates are a promising alternative to “Virgin Aggregates”. And one of the main challenging problems affecting the quality of recycled concrete is the presence of different contaminant particles (i.e., bricks, gypsum, wood, plastic, etc.) (Bonifazi et al., 2018) that can severely reduce the strength of the resulting concrete (Silva et al., 2014). When embedded in concrete, organic substances such as wood are unstable when subjected to dry-wet and freeze-thaw cycles (Hansen, 1992). Water-soluble sulfates present in substances such as gypsum can react and may cause expansive reactions (Alexander and Mindess, 2005). In general, the use of crushed waste glass as coarse aggregates leads to a decrease in the mechanical properties of concrete, primarily due to its irregular shape, poor surface characteristics, and high friability(Harrison et al., 2020; Silva et al., 2014). The density of glass is similar to that of stone and bricks, thus complicating its separation, and in addition, non-crystalline metastable silica may undergo alkali-silica reactions (Hansen, 1992). Therefore, when contaminants normally present in EoL concrete waste are absent or below the limits demanded by market standards, the recycled aggregate may be considered "clean" (Lotfi et al., 2014; Lotfi and Rem, 2016; Serranti et al., 2015) so that EoL concrete can be recycled into clean aggregates to close the materials’ loop in the construction sector. To upcycle EoL concrete, contaminants must be identified, monitored, and minimized. It is essential to identify pollutants in secondary materials produced from EoL concrete to signal exceptions in input quality and recycling process conditions and guarantee clean recycled aggregate products, which requires the establishment of an effective classification and quality control system. It is crucial to exploit efficient, reliable, non-destructive, cost-effective sensing technologies to identify contaminants automatically. Also, to facilitate the broader use of recycled aggregates as construction material, it is essential to create transparency on the quality of recycled aggregates through the value chain.

For the recycling process, an important step is the rapid identification of contaminants in EoL concrete waste. Under cumbersome industrial circumstances, this task can be challenging, particularly at high conveyor belt speeds. Nevertheless, given the significance of improving the quality of secondary materials produced from EoL concrete and reducing the contaminants therein, different technologies and procedural systems have been developed to offer unique and feasible approaches. A hyperspectral imaging (HSI) system in the near-infrared range (Serranti et al., 2012) was applied for quality control to recognize the recycled aggregates from different contaminants (Serranti et al., 2015). However, HSI is still not robust enough under harsh industrial conditions. A classification method based on the integration of the laser-induced breakdown spectroscopy (LIBS) spectral emissions (Lotfi et al., 2015; Xia and Bakker, 2014) was proposed for in-line quality inspection, the success of which relies on the quality of the training set and the possibly remaining false positives.

Recently, the use of LIBS has gained more attention in the field of resource recovery. As a simple, rapid, and efficient analytical technique without sampling requirements, LIBS only samples tiny fractions from a target material’s surface by generating a high power density beam using an ultra-short pulse laser (Cremers and Radziemski, 2006; Xia and Bakker, 2014). As the sampled material is ablated, a plasma is formed, resulting in the emission of an observable spectrum. A spectrometer can detect or analyze to acquire information on the composition of the molecules and atoms of the raw material (Lasheras et al., 2011; Xia H, 2021). Additionally, the advantages of LIBS include the removal of impurities from the sample surface by laser ablation, which decreases their influence on the results; the low cost of analyzing samples compared to other traditional analytical techniques (Yan et al., 2021); the relative simplicity and ease of use of the instrument (Hussain and Gondal, 2013); and the ability to analyze a large number of samples simultaneously in a short time and to detect a wide range of elements (Fernandes Andrade et al., 2021). Consequently, LIBS has been widely applied in the areas of elemental detection (Godoi et al., 2011; Hussain and Gondal, 2013), substance identification (Gondal and Siddiqui, 2007; Völker et al., 2020), and material classification (Castro and Pereira-Filho, 2016; Gottlieb et al., 2017).

Furthermore, there are many studies on combining LIBS and various algorithms for identification and classification, including principal component analysis (PCA) (Junjuri and Gundawar, 2020), scaled conjugate gradient (SCG) (Yang et al., 2020), classification and regression tree (CART) (Moncayo et al., 2015), k nearest neighbor (kNN) (Costa et al., 2017), soft independent modeling of class analogy (SIMCA) (Pease and Tchakerian, 2014), linear discriminant analysis (LDA) (Gaudiuso et al., 2018), partial least squares for discriminant analysis (PLS-DA)(Xia and Bakker, 2014), support vector machine (SVM) (Li et al., 2018), factorial discriminant analysis (FDA) (Baskali-Bouregaa et al., 2020), artificial neural networks (ANN) (Junjuri et al., 2020), and convolutional neural network (CNN) (He et al., 2020). Nevertheless, it is still necessary to increase the precision and sensitivity of this technique. To make algorithms based on LIBS widely available in terms of efficiency and detection limits, several methodological improvements remain to be made.

In this study, an EoL concrete waste identification system based on LIBS was developed, which targeted the precise and automated identification of contaminants. The system emulated the actual industrial situation as much as possible, with each material passing underneath LIBS through a conveyor belt. The LIBS single-shot spectra of each constituent of EoL concrete were collected. Based on these spectral data, a cluster-based classification algorithm was used to create separate spectral databases for each material, allowing for precise identification of the constituents according to a single-shot spectrum. In addition, the effects of different data pre-processing methods and parameters were investigated.

2. Chemometric Methods

In this research, the chemometric methods combining principal

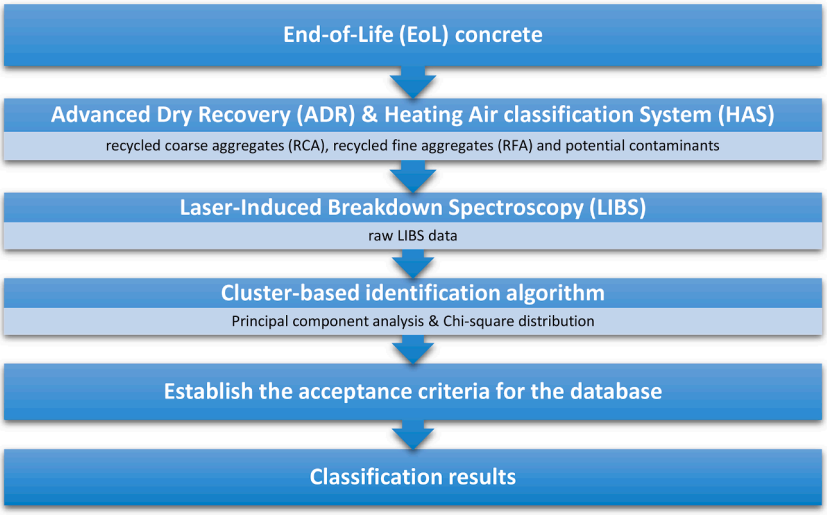

Fig. 1. Scheme for establishing the cluster-based identification model.

component analysis and chi-square distribution are used as a classification model (Fig. 1) for evaluating single-shot spectral data. Their rationales are introduced in detail herewith.

2.1. Principal Component Analysis

As an unsupervised dimensionality reduction method, PCA used for data visualization and pattern detection of raw data, is the most widely used multivariate data analysis algorithm in the LIBS community (Pořízka et al., 2018). Although the thousands of dimensions of the raw spectral data preserve all the information simultaneously, much noise is also retained along with it, resulting in data redundancy and leading to an increased computational effort. Therefore, the high-dimensional raw spectral data needs to be dimensionally reduced.

A database $x[s] = (x_1, x_2, \ldots, x_N)[s]; (s = 1, 2, \ldots, S)$ of S emission spectra for a specific material X is generated, where $x_i (i = 1, \ldots, N)$ is the intensity of plasma emission at a wavelength $\lambda_i (i = 1, 2, \ldots, N)$, N is the number of spectral wavelengths recorded by the spectrometer. Thus, each spectrum can be considered as a point in an N-dimensional space. In this case, the thousands of emission spectra of material X form a cloud in this space that resembles a multi-dimensional ellipsoid. Different materials appear as different clouds of points. For a new spectrum of an unknown material, the classification challenge is to locate the cloud to which it belongs or mark it as unrecognizable if it is too far away from any documented cloud in the database.

Due to a significant amount of variation in particle properties or plasma formation conditions, each spectrum $x[s]$ of material X differs from the centroid $\bar{x}$ of the cluster, then the spectral cloud as a whole represents a multi-dimensional distribution with the centroid $\bar{x}$ as the mean. There is always the possibility to scale (transform) and rotate the axes of the coordinate system, aiming for a simpler multi-dimensional normal distribution of the points in the cluster. Notably, it is always possible to have a new, rotated orthonormal coordinate system with axes aligned along N unit vectors: $e_k = (e_{k1}, e_{k2}, \ldots, e_{kN}); (k = 1, 2, \ldots, N)$. Then, in this new coordinate system, the multi-dimensional normal distribution is equivalent to N independent one-dimensional normal distributions, one for each new axis. Consequently, the spectrum of material X in the database has been transformed into the new coordinate system:

$$\xi[s] = (\xi_1, \xi_2, \ldots, \xi_N)[s] = (x[s]\cdot e_1, x[s]\cdot e_2, \ldots, x[s]\cdot e_N) \tag{1}$$

Then the center point or average of the spectra of material X is:

$$\bar{\xi} = (\overline{\xi_1}, \overline{\xi_2}, \ldots, \overline{\xi_N}) = (\bar{x}\cdot e_1, \bar{x}\cdot e_2, \ldots, \bar{x}\cdot e_N) \tag{2}$$

And the deviations of the spectra concerning the center point in the new system are:

$$\begin{aligned}\Delta\xi[s] &= (\Delta\xi_1, \Delta\xi_2, \ldots, \Delta\xi_N)[s] \\ &= ((x[s] - \bar{x})\cdot e_1, (x[s] - \bar{x})\cdot e_2, \ldots, (x[s] - \bar{x})\cdot e_N)\end{aligned} \tag{3}$$

Thus, the components $\Delta\xi_g[s] (g = 1, 2, \ldots, N)$ and $\Delta\xi_l[s] (l = 1, 2, \ldots, N)$ of the set of spectral deviations along these new axes are mutually uncorrelated with the centroids, shown in Eq. (4),

$$\frac{1}{S}\sum_{1}^{S}\Delta\xi_g[s]\Delta\xi_l[s] = \begin{cases}\overline{\Delta\xi_g^{\,2}} & if\ g = l \\ 0 & if\ g \neq l\end{cases} \tag{4}$$

Moving back to the original coordinate system, Eq. (4) turns into Eq. (5):

$$\frac{1}{S}\sum_{1}^{S}[(x[s] - \bar{x})\cdot e_g][(x[s] - \bar{x})\cdot e_l] = \overline{\Delta\xi_g^{\,2}}\delta_{gl} \tag{5}$$

Eq. (5) is used to find the appropriate set of new unit vectors e_k. The new axes are chosen in such an order that the variances $\overline{\Delta\xi_g^{\,2}}$ of the multi-dimensional normal distribution along the new axes go from high to low values, so that the first one e_1 coincides with the maximum variance $\overline{\Delta\xi_1^{\,2}}$, etc. It is worth noting that not all of these N dimensions are essential for the subsequent categorization process. Only a much lower number n of dimensions needs to be considered. This entails that information from parts of the emission spectra that do not have a large impact (are zero or have little variation except for noise) is omitted. In contrast, the potentially interesting information is presented in the preceding dimensions of the new coordinate system. The significant information is filtered out by projecting the raw spectral data into a low-dimensional space. It is worth mentioning that the value of n will have a substantial impact on the classification accuracy. Therefore, after PCA, the spectral database of S emission spectra for material X will record a number $n[X]$, a set of unit vectors $e_m[X]$ $(m = 1, 2, \ldots, n[X])$, a set of vectors of principal components $(\xi_1, \xi_2, \ldots, \xi_{n[X]})[s]$, and a center point or average $(\overline{\xi_1}, \overline{\xi_2}, \ldots, \overline{\xi_{n[X]}})$ along with variances $\overline{\Delta\xi_m^{\,2}}$, to describe the multi-dimensional normal distribution of the S spectra in the database.

2.2. Chi-square distribution

If $Z_1, Z_2, \ldots, Z_j$ are j independent standard normal distribution $N(0,1)$ random variables, then the sum of their squares $W_j = Z_1^{\,2} + Z_2^{\,2} + \cdots + Z_j^{\,2}$ is said to have a chi-square (χ^2) distribution with j degrees of

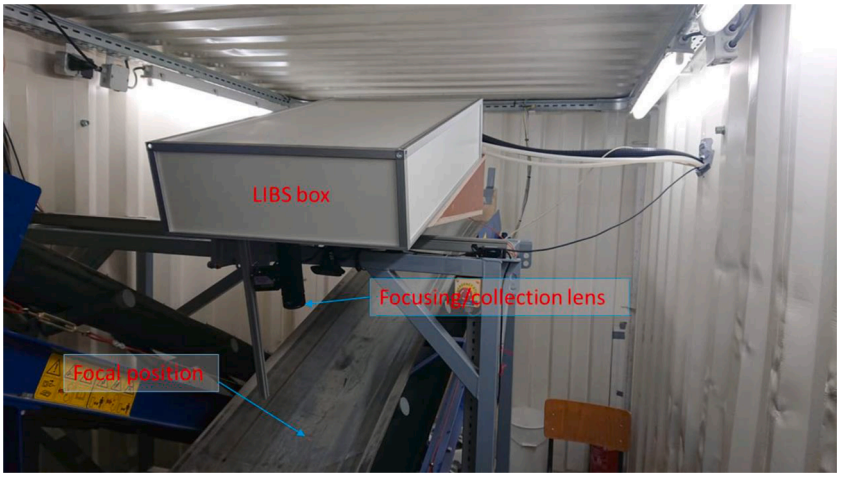

Fig. 2. The LIBS system

freedom, which is often expressed as $W_j \sim \chi^2(j)$ or $W_j \sim \chi_j^2$ (Lancaster and Seneta, 2005). The obtained principal components after data processing can be regarded as chi-square distributions, and then a cluster-based identification algorithm will be established accordingly.

After transforming the spectrum of material X to the new set of coordinates and restricting to the first $n[X]$ dimensions, $(\xi_1, \xi_2, \ldots, \xi_{n[X]})[s]$ is a realization of the multi-dimensional normal distribution for the principal components of spectra of material X. This means that each component ξ_m is normally distributed with mean $\overline{\xi_m}$ and variance $\overline{\Delta\xi_m{}^2}$. Therefore, after z-score standardization, each value:

$$Z_m = \frac{\xi_m - \overline{\xi_m}}{\sqrt{\overline{\Delta\xi_m{}^2}}} \tag{6}$$

is normally distributed with a mean of 0 and a variance of 1. This implies that if it is not known which material was hit by the laser, then the spectrum $x[s]$ can be accepted as originating from material X in case it is highly probable that the set of values Z_m result from $n[X]$ independent standard normal distributions. According to the chi-square distribution, the value:

$$\chi^2[s] = \sum_1^{n[X]} Z_m{}^2 = \sum_1^{n[X]} \frac{(\xi_m - \overline{\xi_m})^2}{\overline{\Delta\xi_m{}^2}} \tag{7}$$

is used to check whether it is small enough to come from the $\chi_{n[X]}{}^2$-distribution. Each $\chi^2[s]$ can be translated into the probability P-value which is the $p[s]$ of $\chi_{n[X]}{}^2$-distribution. The larger the $\chi^2[s]$, the smaller the P-value $p[s]$, the higher the confidence level. P-values lower than the selected significance level indicate statistical significance. By setting the significance level for material X as acceptance criteria, which can be determined according to the P-value $p[X]$ and the associated value of $\chi^2[X]$ for material X, the fraction of all spectra with P-value $p[s]$ greater than $p[X]$ or $\chi^2[s]$ less than $\chi^2[X]$ will be considered as deriving from particles of material X. A small value of the threshold P-value $p[X]$ indicates that most or nearly all spectra from material X will be accepted, but it is also possible that spectra from other materials will be misclassified as material X. A large value of the threshold P-value $p[X]$ implies that many spectra will be classified as not accepted, so these spectra do not contribute to the quality analysis. The issue is to find a good compromise.

3. Experiment and data pre-processing

3.1. Experimental setup

As shown in Fig. 2, the LIBS system consisted of a laboratory-scale conveyor belt, a compact optical module, and an Nd: YAG nanosecond pulse laser for the present study. The Nd: YAG nanosecond pulse laser (TRLi DPSS Series) emitted at a wavelength of 1064 nm, a pulse width of 8-10 ns, a frequency of 100 Hz, and laser energy of 170 mJ per pulse. With a 300 mm focal length lens, the laser was focused vertically onto the sample surface to produce laser-induced plasma. The focusing lens collected the plasma emission spectra and then coupled them to an optical fiber attached to a spectrometer (SPECTRAL Industries, Iris Echelle spectrometer). A delay time of 1.5 μs was employed for the acquisition of the spectra to avoid interference from continuous laser-induced plasma radiation. The timing of the LIBS experiment was triggered with a digital delay pulse generator (Quantum Composers). The experiments were performed under atmospheric conditions. The speed of the conveyor belt was variable and could reach a maximum speed of 50 cm/s. Samples were moved at a constant speed of 20 cm/s to simulate the transport of materials on a typical feed conveyor belt. At 100 Hz, the laser shots every 2 mm on the sample stream.

3.2. EoL concrete samples

Several samples of demolition wastes were collected from demolition sites in the Netherlands. Due to selective demolition, the resulting EoL concrete was clean. Other demolition wastes such as bricks and glasses were separately handpicked from demolition sites. The coarse and fine recycled aggregates were processed by using C2CA technologies (Gebremariam et al., 2020), where the crushed 0-16 mm was treated with ADR and classified as the recycled coarse aggregates (RCA) (4-16mm), and the fine fraction (0-4mm). The fine fraction of recycled aggregates was further treated with Heating Air classification System (HAS) to produce the recycled fine aggregates (RFA) and recycled cement paste-rich powder. Recycled mineral fibers were collected from demolition sites and mechanically ground. The flat glass was also collected from demolition sites and broken into pieces. Recycled gypsum was also in its ground form, while representative forms of foam, wood, and plastics were used.

3.3. Data pre-processing

In general, appropriate pre-processing methods can improve models' classification results by reducing the spectral fluctuations between various measurements (Zeaiter et al., 2006). This research performed no

spectral background subtraction or additional spectral filtering methods on the raw spectral data. This is to avoid losing the spectral information of the laser-induced plasma emission after 1.5 μs since the laser incidence. There were a total of 2,400 single-shot spectra per material at the wavelength range from 179.4 nm to 1199.4 nm with a total of 11790 intensity values per shot. To reduce the accidental error, the average of every adjacent 5 intensity values (Averaged by 5) and the average of every adjacent 10 intensity values (Averaged by 10) were calculated, compared, and evaluated. In addition, Box-Cox transformation was performed to make the intensity values converge to normal distributions. The spectral dataset for each material was divided into training and validation (ratio 9:1): 240 single-shot spectra were randomly selected from each material and combined into a dataset of 2400 single-shot spectra for validation. The remaining 2160 single-shot spectra per material were used for training to build a standard library for each material.

4. Results and discussion

4.1. Optimization of pre-processing methods

Taking the spectral data of bricks as an example, the spectral data using five pre-processing methods (Box-Cox transformation, Averaged by 5, Averaged by 10, Averaged by 5 & Box-Cox transformation, and Averaged by 10 & Box-Cox transformation) were subjected to PCA and compared with the processing methods of the original spectral data. After PCA dimensionality reduction of the brick spectral data, their cumulative explained variance is calculated. The various pre-processing methods increased the explained variance of the first principal component to different degrees. Among the effects resulting from the single-step pre-processing methods, the Box-Cox transformation method showed the most significant improvement compared to the two averaging methods. The difference between the impacts caused by the two averaging methods was not significant. The superimposed pre-processing methods had a greater influence than the single-step pre-processing methods, but the differences between them were not significant. As for the cumulative explained variance, the first 10 principal components of all pre-processing methods could represent almost all information of the spectra. The cumulative explained variance of the first 50 principal components of all pre-processing methods was greater than 0.999, indicating that these principal components were sufficient to cover most of the brick spectra information. The Box-Cox transformation method had a negative impact compared to the original spectral data, while both averaging techniques improved the impact. The averaging methods combined with the Box-Cox transformation method had a negative effect. However, the two superimposed pre-processing methods did not differ much from each other.

It is worth mentioning that, in contrast to the conventional spectral analysis models, this identification model is not mainly dependent on the wave peaks in the LIBS spectra but the overall distribution of the spectra. Therefore, the wave peaks are not analyzed in detail in this paper.

The training set of bricks was used to build its unique database, and the classification results with different pre-processing methods are compared. Because only the spectra of bricks were used as a training data set, ideally, all spectra should be identified as coming from bricks. However, because a uniform p-value was set (for comparison purposes), some spectra were identified as outliers i.e. not coming from bricks. Although the averaging of the raw spectral data could effectively improve the explained variance, it did not affect the discrimination of the training set. In contrast, the Box-Cox transformation method could slightly increase the classification accuracy of the training set.

Furthermore, the validation set of all materials was used to compare the pre-processing methods. When identifying whether a shot was from bricks, there was little difference between the identification results of each pre-processing method, with the Box-Cox transformation method being slightly better. And when identifying whether a shot was from the outliers, the Box-Cox transformation method improved the identification accuracy significantly. In contrast, the technique that averaged every five intensity values was slightly better than the method that averaged every ten intensity values.

Based on the classification results of the validation set, the impact of

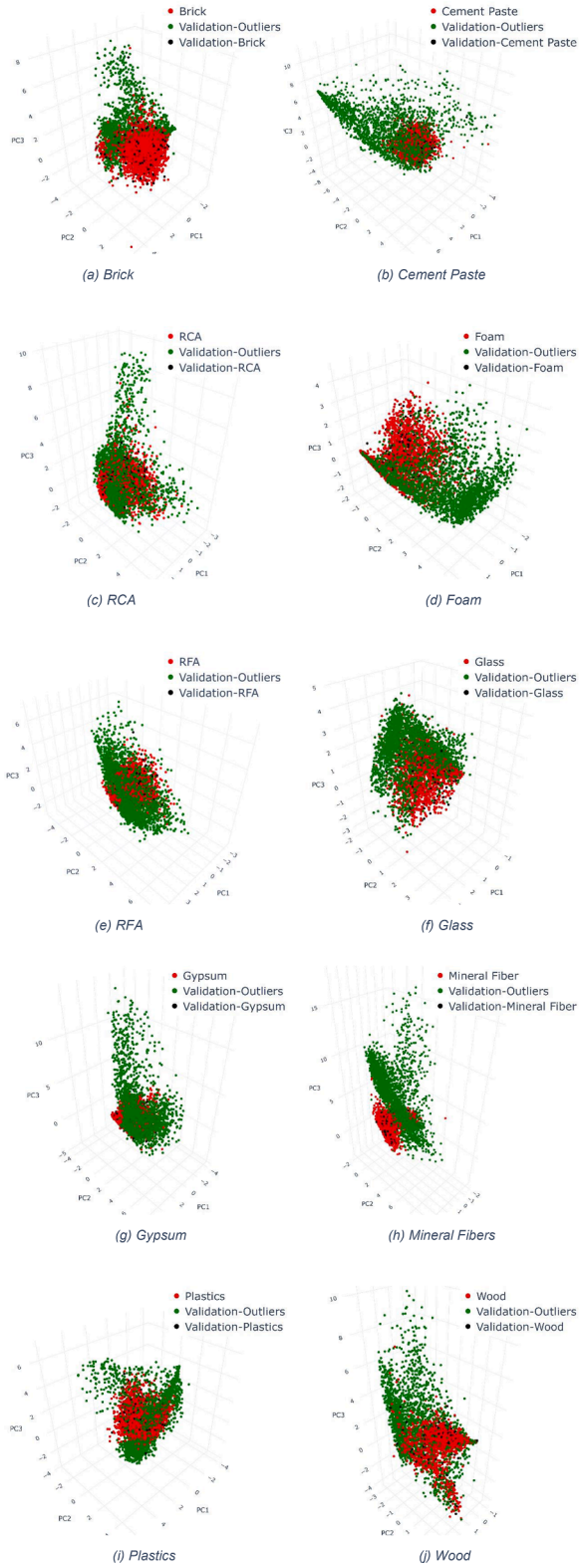

Fig. 3. 3D plots of the first three principal components for each material

each pre-processing method on the final results of the model is evaluated. Using the raw spectral data, the model showed the worst accuracy, while the model using Box-Cox transformation and Averaged by 5 & Box-Cox transformation methods showed the best accuracy, precision (weighted average), recall (weighted average), and F1-score (weighted average) all reaching 0.99. Overall, the Averaged by 5 & Box-Cox transformation pre-processing method was selected to reduce the number of computer operations while achieving better accuracy.

4.2. Optimization of acceptance criteria

As previously mentioned, the number of principal components is the main parameter that affected the final classification accuracy of the model. Combining the training and validation sets, 3D plots of the first three principal components for each material are shown in Fig. 3. Each point represents a single-shot spectrum. Red dots indicate single-shot spectra of certain material in the training set, green dots indicate single-shot spectra of nine materials other than that material in the validation set, and black dots indicate single-shot spectra of that material in the validation set. There were significant differences in the results between the different materials, and it was feasible to differentiate the single-shot spectra from various materials based on the transformed principal components. Different single-shot spectra of the same type of material appeared clustered. For the remaining nine materials, the data points distributed in space were more or less mixed. So each material could create its own exclusive database separately using the red dots and identify and classify other materials accordingly. However, choosing too few principal components may result in a poorly differentiated database, with overlap between different materials. And choosing too many may result in a classification model with a too high threshold that excludes too many points that should have belonged to that material. Thereby, the optimum number of principal components for each material needed to be selected.

In addition, the probability P-value of the chi-square distribution also played an essential role in the accuracy of the final model. Thus, combining the number of principal components and the final P-value was necessary to extract the optimal pairing. After several rounds of attempts, the best matches for different materials are selected.

4.3. Discussion on further optimization of the algorithm

Once the acceptance criteria and databases for each material were created, the validation data set was used to check the accuracy of the entire package of models. When comparing the validation data set with the established material databases, it could be found that some single-shot spectra were accepted by two or even three material databases. Thus, resulting in overlaps for which the belongings of these spectra could not be determined. Among them, the highest number of overlaps was between cement paste and RCA, with the number reaching 50. This was due to the presence of adhering cement paste on the surface of the RCA, which made it difficult to distinguish between the two. To determine the final attribution of the overlapping spectra, their P-values could be made use of. In this case, the P-values of each spectrum obtained in the overlapped material databases should be compared, and the material database corresponding to the maximum P-value is the one to which the spectrum belonged.

Moreover, after evaluating all spectra in the validation set through all material databases, some single-shot spectra were rejected by all material databases. As a result, the belongings of these spectra could not be determined. In this case, an optional method is to compare the P-values. Each of these spectra could obtain a corresponding P-value from each of the ten material databases. The spectrum was then classified into a material database corresponding to the largest P-value by comparing the magnitude of the ten P-values for each spectrum. However, among these spectra, some spectra remained with a P-value of 0 in all ten material databases and could not be classified according to their P-values. Eventually, these spectra were classified as unrecognizable spectra.

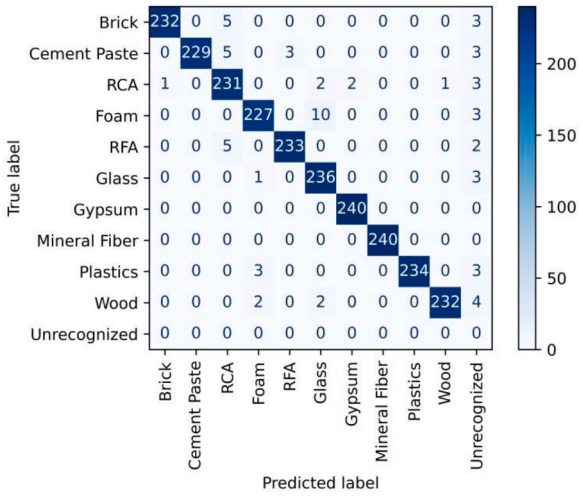

Fig. 4. Confusion matrix of the validation set

Table 1
Classification report of the validation set

	Precision	Recall	F1-score	Support
Brick	1.00	0.97	0.98	240
Cement Paste	1.00	0.95	0.98	240
RCA	0.94	0.96	0.95	240
Foam	0.97	0.95	0.96	240
RFA	0.99	0.97	0.98	240
Glass	0.94	0.98	0.96	240
Gypsum	0.99	1.00	1.00	240
Mineral Fibers	1.00	1.00	1.00	240
Plastics	1.00	0.97	0.99	240
Wood	1.00	0.97	0.98	240
Unrecognized	0.00	0.00	0.00	0
weighted avg	0.98	0.97	0.98	240

The final classification results are shown in Fig. 4. There were still 0-4 single-shot spectra of each material that could not be distinguished. Foam and glass were misidentified the most, with up to 10 single-shot spectra of foam being mistaken for glass, which mainly resulted in a precision of 0.94 for glass. The classification report of the validation set is shown in Table 1. An increase in the value of accuracy, precision, or recall indicated that the model had a better classification performance. Wherein the F1 score is the harmonic mean of the precision and recall, which are mutually constrained. The higher the value of the F1 score is close to 1, the better the model's classification performance is. The accuracy of the whole model reached 0.97, with the precision (weighted average) of 0.98, the recall (weighted average) of 0.97, and the F1-score (weighted average) of 0.98.

The results indicated that the combination of LIBS and cluster-based identification algorithm enabled the precise identification of contaminants in secondary materials produced from EoL concrete. Materials with similar appearance and composition could be distinguished almost completely. The graded materials could be used in different classes of construction work to improve their utilization.

Taking bricks as an example, the raw spectral data of the validation set were classified as bricks, unrecognized, and misclassified spectra, respectively. From Fig. 5, it could be found that the unrecognized spectra were usually caused by the presence of certain peaks much more significant than typical values (pink lines). In contrast, the misclassified spectra had an overall scale much smaller than typical values and were hidden below the typical values. Thus, when translating a spectrum to a point in multi-dimensional space, the point from the unrecognized

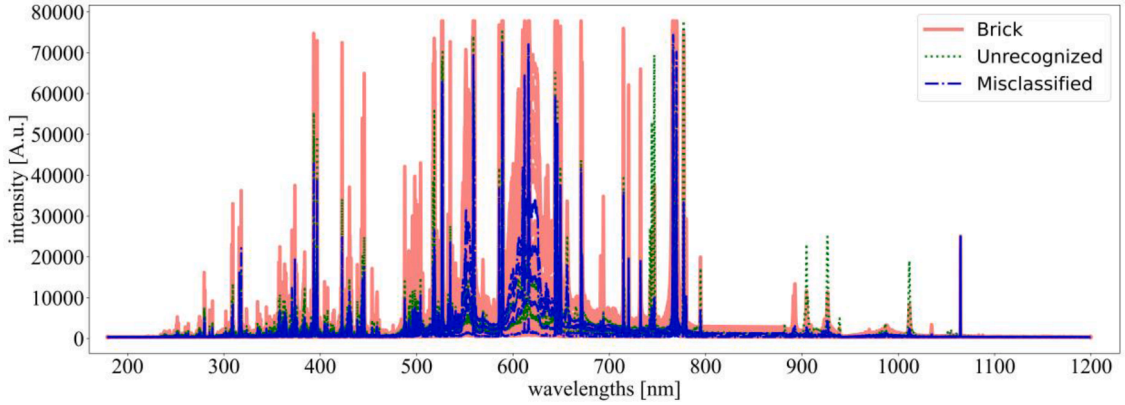

Fig. 5. The raw spectral data of the validation set with classification results

spectrum was usually kept away from the ellipsoid of the brick database, while the point from the misclassified spectrum was contained within the ellipsoid of the brick database. This explains why these spectra were classified as unrecognized spectra and misclassified spectra.

There are also a few recommendations for this identification model. Before using the model for identification, it should be calibrated. Seasonal variations and slow, device-related changes over a long time can produce a drift in the database center point itself. This part of the calibration must be done continuously while the LIBS system is operating.

5. Conclusion

Proposed was a reliable identification technique based on the LIBS spectral emissions for secondary materials produced from EoL concrete in motion. Object material sourced from concrete demolition waste was sampled with a laser in the air. Particular attention was paid to reproducing the working conditions that the feed was experiencing moving on a conveyor belt in recycling practice as closely as possible. An investigation of the method was carried out to analyze the technique's ability to categorize spectra. Firstly different pre-processing methods were used, out of which the Averaged by 5 & Box-Cox transformation method reached the most reliable results. To avoid losing any information, no spectral background subtraction or other sorts of spectral filtering was applied to the raw spectral data. Then, the study of the best match between the number of principal components and P-values for each material was initiated, leading to the creation of a database for each material. The overall accuracy of the model reached 0.97 according to the results of the validation set classification. This approach has excellent accuracy for single-shot LIBS spectra of material in motion compared to conventional qualitative LIBS techniques. Moreover, the proposed methodology does not require the characterization of individual wave peaks appearing in the LIBS spectra. Although the proposed model is sensitive to drift and computationally intensive, it is still worth trying because it is highly reliable in identifying the correct material. Besides, it can be corrected relatively easily for slowly changing conditions. It works better in a reduced dimensional space of variables, reflecting that most of the thousands of spectral data do not contain essential information.

The achieved results demonstrate that the cluster-based classification algorithm is a practical technique for the rapid and online analysis of EoL concrete in motion and can serve as a new method and technique for the industrial selection and quality control of secondary materials produced from EoL concrete. Although only single material streams are sampled to test the quality characterization model in this research, this validates the feasibility of employing the technique to identify contaminants in secondary materials and provides the basis for future tests of mixed product-waste streams. The ultimate goal of the recycled aggregate quality assessment is to provide users with sufficient information about the quality of the product and how to use the material in the best way for a particular application. It is also worth noting that further studies on the levels and grain size distribution of contaminants are needed, which requires finding a good technique to measure and calculate them directly or indirectly.

CRediT authorship contribution statement

Cheng Chang: Conceptualization, Methodology, Software, Data analysis, Writing original draft, Visualization. **Francesco Di Maio:** Conceptualization, Supervision, Writing review & editing. **Peter Rem:** Conceptualization, Supervision, Writing review & editing. **Abraham T. Gebremariam:** Data curation, Writing review & editing. **Fanuel Mehari:** Data curation. Han Xia: Writing review & editing.

Declaration of Competing Interest

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

Data will be made available on request.

Acknowledgment

Funding: This research was supported by European Union's Horizon 2020 funded Project "Cost-Effective Recycling of C&DW in High Added Value Energy Efficient Prefabricated Concrete Components for Massive Retrofitting of our Built Environment" (VEEP, grant agreement No. 723582) and European Union Horizon's 2020 funded Project "Innovative Circular Economy Based solutions demonstrating the Efficient Recovery of valuable material resources from the Generation of representative end-of-life building materials" (ICEBERG, grant agreement No. 869336).

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.resconrec.2022.106507.

Reference

Alexander, M., Mindess, S., 2005. Aggregates in concrete. CRC Press.

Baskali-Bouregaa, N., Milliand, M.-L., Mauffrey, S., Chabert, E., Forrestier, M., Gilon, N., 2020. Tea geographical origin explained by LIBS elemental profile combined to isotopic information. Talanta 211. https://doi.org/10.1016/j.talanta.2019.120674.

Bonifazi, G., Palmieri, R., Serranti, S., 2018. Evaluation of attached mortar on recycled concrete aggregates by hyperspectral imaging. Construction and Building Materials 169, 835–842. https://doi.org/10.1016/j.conbuildmat.2018.03.048.
Castro, J.P., Pereira-Filho, E.R., 2016. Twelve different types of data normalization for the proposition of classification, univariate and multivariate regression models for the direct analyses of alloys by laser-induced breakdown spectroscopy (LIBS). Journal of Analytical Atomic Spectrometry 31, 2005–2014. https://doi.org/10.1039/c6ja00224b.
Cossu, R., Williams, I.D., 2015. Urban mining: Concepts, terminology, challenges. Waste Management 45, 1–3. https://doi.org/10.1016/J.WASMAN.2015.09.040.
Costa, V.C., Aquino, F.W.B., Paranhos, C.M., Pereira-Filho, E.R., 2017. Identification and classification of polymer e-waste using laser-induced breakdown spectroscopy (LIBS) and chemometric tools. Polymer Testing 59. https://doi.org/10.1016/j.polymertesting.2017.02.017.
Cremers, D.A., Radziemski, L.J., 2006. History and fundamentals of LIBS. In: Miziolek, A. W., Schechter, I., Palleschi, V (Eds.), Laser Induced Breakdown Spectroscopy. Cambridge University Press, Cambridge, pp. 1–39. https://doi.org/10.1017/CBO9780511541261.002.
Di Maria, F., Bianconi, F., Micale, C., Baglioni, S., Marionni, M., 2016. Quality assessment for recycling aggregates from construction and demolition waste: An image-based approach for particle size estimation. Waste Management 48, 344–352. https://doi.org/10.1016/J.WASMAN.2015.12.005.
Fernandes Andrade, D., Pereira-Filho, E.R., Amarasiriwardena, D., 2021. Current trends in laser-induced breakdown spectroscopy: a tutorial review. Applied Spectroscopy Reviews. https://doi.org/10.1080/05704928.2020.1739063.
Gaudiuso, R., Ewusi-Annan, E., Melikechi, N., Sun, X., Liu, B., Campesato, L.F., Merghoub, T., 2018. Using LIBS to diagnose melanoma in biomedical fluids deposited on solid substrates: Limits of direct spectral analysis and capability of machine learning. Spectrochimica Acta Part B: Atomic Spectroscopy 146. https://doi.org/10.1016/j.sab.2018.05.010.
Gebremariam, A.T., di Maio, F., Vahidi, A., Rem, P., 2020. Innovative technologies for recycling End-of-Life concrete waste in the built environment. Resources, Conservation and Recycling 163. https://doi.org/10.1016/j.resconrec.2020.104911.
Godoi, Q., Leme, F.O., Trevizan, L.C., Pereira Filho, E.R., Rufini, I.A., Santos, D., Krug, F. J., 2011. Laser-induced breakdown spectroscopy and chemometrics for classification of toys relying on toxic elements. Spectrochimica Acta - Part B Atomic Spectroscopy 66, 138–143. https://doi.org/10.1016/j.sab.2011.01.001.
Gondal, M.A., Siddiqui, M.N., 2007. Identification of different kinds of plastics using laser-induced breakdown spectroscopy for waste management. Journal of Environmental Science and Health - Part A Toxic/Hazardous Substances and Environmental Engineering 42, 1989–1997. https://doi.org/10.1080/10934520701628973.
Gottlieb, C., Millar, S., Grothe, S., Wilsch, G., 2017. 2D evaluation of spectral LIBS data derived from heterogeneous materials using cluster algorithm. Spectrochimica Acta Part B: Atomic Spectroscopy 134. https://doi.org/10.1016/j.sab.2017.06.005.
Hansen, T.C., 1992. Recycling of demolished concrete and masonry. CRC Press.
Harrison, E., Berenjian, A., Seifan, M., 2020. Recycling of waste glass as aggregate in cement-based materials. Environmental Science and Ecotechnology 4, 100064. https://doi.org/10.1016/J.ESE.2020.100064.
He, Y., Zhao, Y., Zhang, C., Li, Y., Bao, Y., Liu, F., 2020. Discrimination of grape seeds using laser-induced breakdown spectroscopy in combination with region selection and supervised classification methods. Foods 9. https://doi.org/10.3390/foods9020199.
Hussain, T., Gondal, M.A., 2013. Laser induced breakdown spectroscopy (LIBS) as a rapid tool for material analysis. In: Journal of Physics: Conference Series. Institute of Physics Publishing. https://doi.org/10.1088/1742-6596/439/1/012050.
Junjuri, R., Gundawar, M.K., 2020. A low-cost LIBS detection system combined with chemometrics for rapid identification of plastic waste. Waste Management 117. https://doi.org/10.1016/j.wasman.2020.07.046.
Junjuri, R., Prakash Gummadi, A., Kumar Gundawar, M., 2020. Single-shot compact spectrometer based standoff LIBS configuration for explosive detection using artificial neural networks. Optik 204. https://doi.org/10.1016/j.ijleo.2019.163946.
Kabirifar, K., Mojtahedi, M., Changxin Wang, C., Tam, V.W.Y., 2021. Effective construction and demolition waste management assessment through waste management hierarchy; a case of Australian large construction companies. Journal of Cleaner Production 312, 127790. https://doi.org/10.1016/J.JCLEPRO.2021.127790.
Lancaster, H.O., Seneta, E., 2005. Chi-Square Distribution. Encyclopedia of Biostatistics. John Wiley & Sons, Ltd, Chichester, UK.. https://doi.org/10.1002/0470011815.b2a15018
Lasheras, R.J., Bello-Gálvez, C., Rodríguez-Celis, E.M., Anzano, J., 2011. Discrimination of organic solid materials by LIBS using methods of correlation and normalized coordinates. Journal of Hazardous Materials 192. https://doi.org/10.1016/j.jhazmat.2011.05.074.
Li, X., Yang, S., Fan, R., Yu, X., Chen, D., 2018. Discrimination of soft tissues using laser-induced breakdown spectroscopy in combination with k nearest neighbors (kNN) and support vector machine (SVM) classifiers. Optics & Laser Technology 102. https://doi.org/10.1016/j.optlastec.2018.01.028.
Lotfi, S., di Maio, F., Xia, H., Serranti, S., Palmieri, R., Bonifazi, G., 2015. Assessment of the contaminants level in recycled aggregates and alternative new technologies for contaminants recognition and removal. EMABM 2015: Proceedings of the 15th Euroseminar on Microscopy Applied to Building Materials, Delft, The Netherlands, 17-19 June 2015.
Lotfi, S., Rem, P., 2016. Recycling of End of Life Concrete Fines into Hardened Cement and Clean Sand. Journal of Environmental Protection 07. https://doi.org/10.4236/jep.2016.76083.
Moncayo, S., Manzoor, S., Navarro-Villoslada, F., Caceres, J.O., 2015. Evaluation of supervised chemometric methods for sample classification by Laser Induced Breakdown Spectroscopy. Chemometrics and Intelligent Laboratory Systems 146. https://doi.org/10.1016/j.chemolab.2015.06.004.
Nanda, S., Berruti, F., 2021. Municipal solid waste management and landfilling technologies: a review. Environmental Chemistry Letters 19. https://doi.org/10.1007/s10311-020-01100-y.
Pease, P., Tchakerian, V., 2014. Source provenance of carbonate grains in the Wahiba Sand Sea, Oman, using a new LIBS method. Aeolian Research 15. https://doi.org/10.1016/j.aeolia.2014.06.001.
Pořízka, P., Klus, J., Képeš, E., Prochazka, D., Hahn, D.W., Kaiser, J., 2018. On the utilization of principal component analysis in laser-induced breakdown spectroscopy data analysis, a review. Spectrochimica Acta - Part B Atomic Spectroscopy. https://doi.org/10.1016/j.sab.2018.05.030.
Serranti, S., Gargiulo, A., Bonifazi, G., 2012. Classification of polyolefins from building and construction waste using NIR hyperspectral imaging system. Resources, Conservation and Recycling 61, 52–58. https://doi.org/10.1016/J.RESCONREC.2012.01.007.
Silva, R.v., de Brito, J., Dhir, R.K., 2014. Properties and composition of recycled aggregates from construction and demolition waste suitable for concrete production. Construction and Building Materials 65, 201–217. https://doi.org/10.1016/J.CONBUILDMAT.2014.04.117.
Vegas, I., Broos, K., Nielsen, P., Lambertz, O., Lisbona, A., 2015. Upgrading the quality of mixed recycled aggregates from construction and demolition waste by using near-infrared sorting technology. Construction and Building Materials 75, 121–128. https://doi.org/10.1016/J.CONBUILDMAT.2014.09.109.
Völker, T., Millar, S., Strangfeld, C., Wilsch, G., 2020. Identification of type of cement through laser-induced breakdown spectroscopy. Construction and Building Materials 258. https://doi.org/10.1016/j.conbuildmat.2020.120345.
Xia H, 2021. Sensor-based quality inspection of secondary resources Laser-induced breakdown spectroscopy. https://doi.org/10.4233/uuid:d6fa8b1f-0f9b-4ed3-aedc-b4e93c83b2fc.
Xia, H., Bakker, M.C.M., 2014. Reliable classification of moving waste materials with LIBS in concrete recycling. Talanta 120, 239–247. https://doi.org/10.1016/j.talanta.2013.11.082.
Yan, B., Liang, R., Li, B., Tao, J., Chen, G., Cheng, Z., Zhu, Z., Li, X., 2021. Fast identification and characterization of residual wastes via laser-induced breakdown spectroscopy and machine learning. Resources, Conservation and Recycling 174. https://doi.org/10.1016/j.resconrec.2021.105851.
Yang, Y., Li, C., Liu, S., Min, H., Yan, C., Yang, M., Yu, J., 2020. Classification and identification of brands of iron ores using laser-induced breakdown spectroscopy combined with principal component analysis and artificial neural networks. Analytical Methods 12, 1316–1323. https://doi.org/10.1039/c9ay02443c.
Zeaiter, M., Roger, J.M., Bellon-Maurel, V., 2006. Dynamic orthogonal projection. A new method to maintain the on-line robustness of multivariate calibrations. Application to NIR-based monitoring of wine fermentations, in: Chemometrics and Intelligent Laboratory Systems. pp. 227–235. https://doi.org/10.1016/j.chemolab.2005.06.011.
Zhang, C., Hu, M., Dong, L., Gebremariam, A., Mirand-Xicotencatl, B., di Maio, F., Tukker, A., 2019. Eco-efficiency assessment of technological innovations in high-grade concrete recycling. Resources, Conservation and Recycling 149, 649–663. https://doi.org/10.1016/J.RESCONREC.2019.06.023.