# Rapid Identification and Quantitative Analysis of Polycarboxylate Superplasticizers Using ATR-FTIR Spectroscopy Combined with Chemometric Methods 

 and Jiadong Wang ( ${ }^{1}{ }^{1}$<br>${ }^{1}$ National and Provincial Joint Engineering Laboratory of Road \& Bridge Disaster Prevention and Control, Lanzhou Jiaotong University, Lanzhou 730070, China<br>${ }^{2}$ Gansu Highway Construction Management Group Co., Ltd., Lanzhou, China<br>${ }^{3}$ Key Laboratory of Highway Net Monitoring in Gansu Province, Gansu Hengda Road and Bridge Group Co., Ltd., Lanzhou, China<br>Correspondence should be addressed to Bo Li; 8857807@qq.com

Received 31 October 2020; Revised 18 December 2020; Accepted 2 January 2021; Published 21 January 2021
Academic Editor: Dongyu Niu
Copyright © 2021 Zhiwei Li et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.


#### Abstract

Although the quality inspection method of polycarboxylate superplasticizers (PCE) based on macroperformance is still widely used, it has the drawbacks of time-consuming and low precision. This study aims to develop a practicable alternative method for quality inspection of PCE. For this, spectra collection, feature extraction, and cluster analysis were performed up on the PCE samples to demonstrate the feasibility of the method. Also, a new similarity calculation method was introduced in this work. Results show that the solid PCE sample for spectrum collection can be prepared using the simple heating method. High-quality spectra can be rapidly collected by infrared spectrometer combined with ATR accessory. Meanwhile, the accuracy of classification and clustering is high, suggesting that the feature extraction method based on principal component analysis (PCA) is effective. In addition, compared with conventional similarity calculation methods of cosine angle and correlation coefficient, the new similarity calculation method achieves better classification results and better generalization ability. This work provides a method of quantitative analysis and rapid identification of PCE for the construction site.


## 1. Introduction

As an important concrete admixture, the water-reducing agent has been widely used in improving concrete performance. They can improve the compressive strength by reducing the amount of water required [1]. Lignosulfonate was the first plasticizer used in concrete, which is a derivative of lignin in pulp industry. Then, poly melamine sulfonates and poly naphthalene sulfonates were developed and applied. However, these plasticizers have limited water-reducing effect and may bring potential problems to the concrete hardening process [2]. The new generation is polycarboxylate superplasticizer (PCE). These superplasticizers are widely used in the construction site and studied in the academic field. Take

China as an example, the consumption of PCE in China was 7.23 million tons in 2017, which was about $77.6 \%$ of all waterreducing agents used in the year [3]. Therefore, choosing PCE accurately and reasonably is a major concern for industry personnel.

Authenticity is a very important quality criterion for PCEs because there is a big difference among different types of PCE products [4, 5]. However, determination of authenticity for PCEs is traditionally a time-consuming and laborious process, typically using dispersion effect, adsorption amount, and setting times to characterize PCEs [6, 7]. Although these methods can reflect the macroperformance and ensure the engineering performance of PCEs to some extent, they have some drawbacks.

On the one hand, the relevant methods require a huge amount of PCE sample and high operating skills of the test personnel, which may bring potential impact on accuracy [8]. On the other hand, macroperformance test results mainly reflect construction control index, but cannot characterize the relationship between molecular structure and performance of PCE. For example, different PCEs with same or similar macroperformance may exhibit distinct-different in-site performance, with regard of durability, strength, and workability of concrete due to molecular composition of PCE, resulting in structure diseases after long-term use [9].

Moreover, the molecular structure of PCE can be readily adjusted to provide desirable features in order to improve the performance of concrete, such as adaptability to different temperatures and compatibility with different types of cements [10]. Just because the molecular structure is different, the physical and chemical composition of PCE is unique [11]. This uniqueness is similar to human fingerprint, which can be called fingerprint information of PCE. In addition, due to the fingerprint information of PCE, it is feasible to identify PCE in theory.

Recently, fingerprint identification technology has been widely applied in material analysis [12, 13]. Several techniques have been used to describe and determine the quality of the test sample, including colorimetric methods, Gel Permeation Chromatography (GPC), Nuclear Magnetic Resonance Hydrogen Spectrum (NMR), and Gas Chromatography Mass spectrometry. Although these methods can identify the unique fingerprint information of samples, they have the disadvantages of time-consuming, expensive, or cumbersome operation. Compared with the conventional methods, Fourier transform infrared (FTIR) spectroscopy has become one of the most useful analytical means because of its merits including rapid detection, easy operation, and cost saving. Combined with attenuated total reflection (ATR) accessory, a small amount of samples can be analyzed in a short time without sample preparation. ATR-FTIR spectroscopy measures the sample at a depth of $0.5-5 \mu \mathrm{~m}$ from the surface of the internal reflection element. ATRFTIR schematic diagram is shown in Figure 1, which usually employs a diamond, and collects information from the layer of the sample adjacent to the surface of the internal reflection element. In particular, the development of chemometric technology linked with ATR-FTIR spectral data provides an effective way to determine the physical and chemical properties of materials. Mahsa Mohammadi proposed a new analytical method using ATR-FTIR spectroscopy associated with chemometric methods for addressing regression and classification tasks for crude oil analysis based on API gravity values [14]. Hannah Tiernan applied ATR-FTIR spectroscopy to determining protein content and protein secondary structure composition [15]. Li Zhihui employed ATR-FTIR spectra coupling with SVM classifier to identify graded asphalt [16]. Xiting Sun proposed a rapid and expedient method to determine asphalt properties, including content, softening point, and penetration [17]. Christian Schramm studied the curing process of the impregnated and dried cotton samples by using ATR-FTIR spectroscopy [18].


Figure 1: The schematic diagram of ATR-FTIR technology.

Currently, the combination of ATR-FTIR and chemometrics has been widely used in the evaluation of petroleum, flocculant, and wood. However, there is no work that has treated the properties of PCE by direct spectroscopic methods according to literature index. This work uses ATRFTIR combined with chemometrics to identify and quantitatively analyze PCE. The method of obtaining high-quality spectrum of PCE is firstly explored. The feature extraction and classification were carried out by chemometrics. In addition, different PCEs were compared by the similarity calculation formula, and acceptance threshold was established for each PCE. The results provide a rapid, accurate, and nondestructive method for quality control of PCE.

## 2. Materials and Methods

2.1. Materials. The selected PCEs for this study are commonly used by contractors on construction projects in Gansu Province, China. Four kinds of PCEs are selected from four different manufacturers, denoted here as $\mathrm{A}, \mathrm{B}, \mathrm{C}$, and D . The corresponding batches are denoted here as A 1 , A2, A3, A4, B1, B2, B3, B4, C1, C2, C3, C4, D1, D2, D3, and D4. The acronym A1 represents manufacturer A and batch 1, as indicated in Table 1. All PCE samples are sealed in the sample tank before testing to avoid long time connection with air and stored in a shady and cool place.
2.2. Sample Preparation. All the PCE collected in this research are liquid samples, and the samples contain water. Water can interfere with the quality of infrared spectrum, so it is crucial to eliminate the impact of water. Some researchers achieve this goal by organic solvent extraction. In Standard ASTM C494/C494M-13, a drying method is adopted, that is, the diluted PCE sample is weighed and dried at $105 \pm 3^{\circ} \mathrm{C}$ for $17 \pm 1 / 4 \mathrm{~h}$ [19]. After cooling, infrared spectra of dried residue are studied by the KBr method. However, this method is complex and time-consuming, so it is not suitable for field test. Casale et al. III exploded the feasibility of reducing the drying time to 10 h [5]. Part of the study content, especially the spectral sampling method, is different from this study. Therefore, it is necessary to explore the feasibility of reducing drying time in this study.

This study refers to the test procedure outlined by the ASTM Standard above. In order to study the feasibility of reducing drying time, weigh $3 \pm 0.1 \mathrm{~g}$ PCE sample, put it in an evaporation dish, and dry it in an oven at $105 \pm 3^{\circ} \mathrm{C}$ for 30 , $60,90,120,150$, and 180 minutes until the weight does not

Table 1: Detailed information of PCEs used in this paper.

| PCE | Manufacturer | Batch | Water-reducing rate (\%) | Bleeding rate (\%) | Setting time difference (min) |  |  | ssive <br> th <br> (\%) |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  |  |  |  |  | Initial setting | Final setting | 7 d | 28 d |
| A1 | A | 1 | 31 | 9 | -10 | +35 | 194 | 168 |
| A2 |  | 2 | 32 | 18 | +65 | +5 | 194 | 167 |
| A3 |  | 3 | 29 | 9 | -25 | -65 | 186 | 164 |
| A4 |  | 4 | 28 | 12 | -20 | -55 | 165 | 142 |
| B1 | B | 1 | 35 | 52 | +120 | +120 | 197 | 173 |
| B2 |  | 2 | 27 | 12 | +100 | - | 163 | 149 |
| B3 |  | 3 | 26 | 22 | +105 | - | 153 | 147 |
| B4 |  | 4 | 26 | 35 | +120 | - | 192 | 174 |
| C1 | C | 1 | 27 | 27 | +95 | +70 | 198 | 169 |
| C2 |  | 2 | 29 | 42 | +100 | +60 | 190 | 165 |
| C3 |  | 3 | 32 | 12 | +95 | +105 | 194 | 168 |
| C4 |  | 4 | 28 | 15 | +95 | +40 | 190 | 169 |
| D1 | D | 1 | 35 | 50 | +120 | +120 | 178 | 173 |
| D2 |  | 2 | 27 | 60 | +120 | - | 160 | 146 |
| D3 |  | 3 | 28 | 20 | +80 | +70 | 161 | 155 |
| D4 |  | 4 | 27 | 15 | +100 | - | 158 | 147 |



Figure 2: The sample preparation process and ATR-FTIR spectrum collection.
change. The time is recorded and used as drying time. The flow chart of sample preparation is shown in Figure 2.
2.3. FTIR Measurements. ATR-FTIR spectra were obtained using Thermo Fisher Scientific ${ }^{\circledR}$ NicoletiS5 Fourier Transform Infrared Spectrometer combined with iD5 ATR accessory, with an instrument parameter of 32 scans in $4000-400 \mathrm{~cm}^{-1}$ range and a resolution of $4 \mathrm{~cm}^{-1}$. To collect infrared spectra, the dried residue of PCE was directly put on the diamond ATR crystal, and the torque knob was tightened to ensure that the sample is in close contact with the diamond ATR crystal. Analyses were carried out at room temperature $\left(25^{\circ} \mathrm{C}\right)$. The background collected before every sample was measured. The diamond ATR crystal was cleaned by kerosene. In order to solve the problem that PCE samples are small, a reasonable method is to repeat the number of spectral for each sample. Based on the word reported by Yang et al., three scans were performed and saved in a database for each PCE in each batch [4]. Please note that each scan saved is the average of
four scans, and 48 scans were utilized for chemometric analysis.
2.4. Data Pretreatment and Outlier Detection. Before using the chemometric method for the next calculation, raw data obtained should be preprocessed. Baseline correction was performed by using OMINIC software. Then, the data were smoothed using the Savitzky-Golay method. Also, in order to increase the calculation accuracy, it is necessary to detect and eliminate outliers. To this end, according to the method provided by Mahsa Mohamadi's paper, the PCA model and Hotelling $\mathrm{T}^{2}$ statistical plot are used to detect outliers [20]. Outliers can be eliminated from the original dataset, and the original dataset were loaded in the Unscrambler V-10.5 software (CAMO Analytics) for outlier detection by using the Hotelling $\mathrm{T}^{2}$ statistical method. Data processing and PCA were further performed with MATLAB software (MathWorks).
2.5. Multivariate Analysis: Chemometrics. Infrared spectrum data is a high-dimensional vector composed of
absorbance corresponding to different wave numbers. Along with the development of spectral detection technology, the characteristic variables of spectral data become larger and larger. One of the major drawbacks of high-dimensional data analysis is the curse of dimensionality. However, a related fact is that hyperspace is essentially sparse, causing the phenomenon of empty space [21]. A natural way to mitigate this problem is to significantly reduce the dimension of original data. It has been shown that principal component analysis (PCA) is an effective multivariate statistical approach for both dimension reduction and information retention. Therefore, all ATR-FTIR spectra data were analyzed by principal component analysis (PCA) to determine the characteristic band. In this way, the redundant information can be filtered effectively. Then, the band selection method is verified by support vector machine (SVM) and cluster analysis [22]. As a new algorithm for classification, SVM has been widely used to nonlinear separation problems [23]. SVM represents the samples as points in space. The samples are separated by plane or line vectors after mapping. The new sample data is first mapped to the same space, and the category is determined based on which side of the vector they fall. A set of 108 samples were used to construct the calibration model, and 36 samples were used for the prediction model.

In addition, similarity calculation was used for calculating the spectral difference and similarity threshold for each PCE was determined. Spectral similarity analysis is an effective method to characterize spectral difference [24]. According to the similarity results, the quality fluctuation of PCE can be determined or the types of PCE can be distinguished. In the process of infrastructure construction, IR spectrum can be considered when the sample is completely known, and the material composition and quality are required to confirm. However, based on the favorable comparison between the infrared spectrum and original sample spectrum, the interpretation of the results is qualitative. The New Jersey Department of Transportation (NJDOT) uses correlation coefficient-based quantitative assessment of infrared scanning to accept or reject field sample. The NJDOT determined an acceptable correlation coefficient threshold of $97.5 \%$ [5]. However, the objective fact is that the thresholds for each material may be different. In addition, the quality control method for quantitative evaluation of concrete admixtures using the correlation coefficient has not been established. Therefore, it is necessary to establish the correlation between the sample to be tested and the original sample and to establish the acceptable threshold. This helps to visually verify whether the PCE received at the construction site is acceptable using the quantitative method. Most researchers believe that similarity is an effective parameter to distinguish different spectra [25-27]. After literature retrieval, there are many methods to calculate the similarity, among which the angle cosine formula and
correlation coefficient formula are more commonly used, which can be evaluated by

$$
\begin{align*}
\cos \alpha & =\frac{\sum_{i=1}^{n} a_{i} b_{i}}{\sqrt{\sum_{i=1}^{n} a_{i}^{2}} \sqrt{\sum_{i=1}^{n} b_{i}^{2}}}, \\
S & =\frac{\sum_{i=1}^{n}\left(a_{i}-\bar{a}\right)\left(b_{i}-\bar{b}\right)}{\sqrt{\sum_{i=1}^{n}\left(a_{i}-\bar{a}\right)^{2} \sum_{i=1}^{n}\left(b_{i}-\bar{b}\right)^{2}}}, \tag{1}
\end{align*}
$$

where $\cos \alpha$ means angle cosine value, $a_{i}$ means absorbance values of the measured PCE sample, $b_{i}$ means absorbance values of the standard PCE sample, $\bar{a}$ means average of absorbance value of the measured PCE sample, and $\bar{b}$ means average of absorbance value of the standard PCE sample. The detailed flowchart is illustrated in Figure 3.

## 3. Results and Discussion

3.1. Drying Time. ASTM C494/C494M-13 Sect.18.1.1 requires that dry conditions are $17 \pm 1 / 4 \mathrm{~h}$ at $105 \pm 3^{\circ} \mathrm{C}$ [19]. This method belongs to the existing uniformity and equivalence tests, so the quality of PCEs can be unchanged by high temperature in theory. However, considering that this method needs to dilute the PCE first and the drying time is too long, this study explored the feasibility of shortening the drying time.

Figure 4 illustrates typical variation curves of weight of PCEs with drying time. The drying time at the temperature of $105 \pm 3^{\circ} \mathrm{C}$ was $0,30,60,90,120,150$, and 180 min , respectively. As shown in Figure 4, the weight of four different PCEs decreases with the increase of drying time. The observation can be attributed to the evaporation of water in PCEs. Furthermore, after evaporation for 30 minutes, the weight of PCEs decreases sharply and then decreases slowly with drying time. At 150 minutes, the weight of four PCEs was kept stable. Therefore, the drying time of all samples is determined to be 150 minutes.
3.2. Spectral Characteristics of PCE. The ATR-FTIR spectra of evaporated residue of different PCEs are shown in Figure 5. PCEs are composed of different types of hydrocarbons and oxygen containing compounds. As shown in Figure 5, the spectra of all PCE samples are basically the same. The intensity and position of some absorption peaks are slightly different, which indicates that the composition of different PCEs is different [28, 29]. The peak location provides the structural information of the sample, and the peak intensity corresponds to the information of molecular concentration in the sample. In addition, the width of the peak is easily affected by the pH value and hydrogen bond [30]. The characteristic absorption peaks of all PCE spectral range from 3700 to $700 \mathrm{~cm}^{-1}$ and are especially concentrated between 1700 and $700 \mathrm{~cm}^{-1}$, that is, the fingerprint region. The band observed at $3400-3200 \mathrm{~cm}^{-1}$ corresponds


Figure 3: The flowchart of the chemometric analysis for determination and classification of PCEs.


Figure 4: Variation in weight of PCEs at different drying times: temperature $=105^{\circ} \mathrm{C}$.
to $\mathrm{O}-\mathrm{H}$ stretching vibrations in PCE. The absorption peaks around 2865 and $2950 \mathrm{~cm}^{-1}$ are attributed to aliphatic C-H symmetric and asymmetric stretching vibrations, respectively. Moreover, the band at 1635-1750 indicates the existing of $\mathrm{C}=\mathrm{O}$ stretching vibrations. The band around 1130-1070 corresponds to C-O stretching vibrations. The spectral characteristics measured are consistent with the results of literature, indicating that the sample preparation method is reasonable [31].

In addition to the absorption peaks described above, there are some nonobvious peaks in infrared spectra of PCEs. In order to determine which absorption peak can represent the difference of different PCEs, all spectra need to be further processed and analyzed by the chemometric method.
3.3. Data Preprocessing and Outlier Detection. Figure 6(a) is the overview of PCA using the Unscrambler software. In the score chart, $85 \%$ of total variance was presented by the first two principal components (PC-1 and PC-2). Outliers can also be detected by Hotelling's $\mathrm{T}^{2}$ statistic plot, as shown in Figure 6(c). It is noteworthy to see that no outlier is detected. Therefore, all spectral data can be utilized for further calculation.
3.4. Spectral Feature Selection. A total of 48 spectral data of 4 different PCEs were imported into MATLAB software to establish the PCA model. Table 2 shows the results of principal component analysis. It can be seen from Table 2 that there were three principal components: $\mathrm{PC} 1, \mathrm{PC} 2$, and PC3, with an eigenvalue of 0.2 or more. The eigenvalue of the first principal component was large with 0.542 , as shown in Figure 7, and the contribution rate of variance of the first four principal components is $98.47 \%$. Therefore, it can be concluded that most of the information of the original data has been retained by the first four principal components [24].

PCA results consist of the score plot and loading plot. Among them, the score plot represents the relationship between different samples, while the loading plot represents the relationship between variables. The loading plot can help to determine which variables carry useful information. Figure 8 shows the distribution of the load factor with the wave number. The relationship between spectra variables becomes clear. The absorption peak around $1730 \mathrm{~cm}^{-1}$ has a higher influence on the PCA model in the negative direction of $X$-axis, which corresponds to $\mathrm{C}=\mathrm{O}$ stretching vibrations, indicating this absorption peak of different PCEs is obviously different. Similarly, the wave number around 3287, 2859, 1573, 1466, 1409, 1342, 1278, 1241, 1093, 961, and $839 \mathrm{~cm}^{-1}$ also have high influence on the PCE model. In addition, compared with PC1, PC2, and PC3, the loading factor distribution of PC4 is more balanced. This is because the variance of PC 4 is less than that of $\mathrm{PC} 1, \mathrm{PC} 2$, and PC 3 , that is to say, PC4 contains less difference among sample classes. Therefore, PC4 basically does not contain valid variables.

In addition, according to the rotation component matrix, the absorbance at the wave number of $3287,2859,1730$, $1573,1466,1409,1342,1278,1241,1093,961$, and 839 has a large load. It is noteworthy that these 12 bands correspond to the corresponding absorption peaks of ATR-FTIR spectrum of PCE, indicating that the difference between PCEs is mainly near these 12 peaks, as shown in Table 3. It should be noted that the O-H hydroxyl absorption peak around $3287 \mathrm{~cm}^{-1}$ fluctuates greatly and is not suitable for modeling. Therefore, the 11 absorption peak wavelengths can be selected for further calculation, as shown in Table 3.
3.5. Support Vector Machine (SVM) Model. The present work proposed a model of classification based on SVM. In order to expand the sample, sample preparation was performed, and 36 spectral data were collected for each PCE sample. A total of 144 spectral data were collected. Among them, a set of 108 samples were used to construct the calibration model, and 36


Figure 5: Full-wave band and local band spectral of PCE without any pretreatment.


Figure 6: (a) Score plot. (b) Residual plot on ATR-FTIR of PCEs. (c) Hotelling's T2 statistic plot.

Table 2: Results of principal component analysis.

| Number of principal components | Eigenvalue | Variance contribution (\%) | Cumulative variance contribution (\%) |
| :--- | :---: | :---: | :---: |
| 1 | 0.6521 | 0.5426 | 0.5426 |
| 2 | 0.2556 | 0.2127 | 0.7553 |
| 3 | 0.2413 | 0.2008 | 0.9561 |
| 4 | 0.0343 | 0.0286 | 0.9847 |
| 5 | 0.0076 | 0.0063 | 0.9910 |
| 6 | 0.0063 | 0.0053 | 0.9963 |
| 7 | 0.0013 | 0.0011 | 0.9974 |
| 8 | 0.0008 | 0.0007 | 0.9981 |
| 9 | 0.0004 | 0.0004 | 0.9984 |
| 10 | 0.0004 | 0.0003 | 0.9988 |
| 11 | 0.0003 | 0.0003 | 0.9990 |
| 12 | 0.0002 | 0.0002 | 0.9992 |



Figure 7: The scoring diagram of the first six principal components.
samples were used for the prediction model. By using the SVM classifier, the accuracy rate of total verifying samples stood at up to $100 \%$, as shown in Figure 9. Figure 10 shows the prediction accuracy of four PCE samples. The diagonal part is the proportion accurately predicted, while the other part is the proportion misjudged as other classes. As shown, the four PCEs used in this work were found with an accuracy of $100 \%, 100 \%, 100 \%$, and $100 \%$. It means it was $100 \%$ correctly classified, whereas none were misclassified. Similarly, PCE-A, PCE-B, and PCE-C were $100 \%$ correctly classified. Therefore, the type of PCE can be rapidly identified by using the SVM classifier. In addition, characterized bands selected are effective and suggested for further calculation.
3.6. Systematic Clustering Analysis. Figure 11 shows the results for systematic clustering analysis, and the data used for analysis were the characteristic bands of ATR-FTIR spectra of four PCEs. The abscissa represents the sample number, and the ordinate represents the rescaled distance cluster combine.

It can be seen that the spectra of all samples are classified into four categories, and the result was consistent with the
number of PCE samples used in this study. Among them, one cluster was A1-1, A1-2, A1-3, A2-1, A2-2, A2-3, A3-1, A3-2, A3-3, A4-1, A4-2, and A4-3. The second cluster was B1-1, B12, B1-3, B2-1, A2-2, A2-3, B3-1, B3-2, B3-3, B4-1, B4-2, B4-3, and C3-2. The third cluster was C1-1, C1-2, C1-3, C2-1, C2-2, A2-3, C3-1, C3-3, C4-1, C4-2, and C4-3, while D1-1, D1-2, D1-3, D2-1, D2-2, D2-3, D3-1, D3-2, D3-3, D4-1, D4-2, and D4-3 belong to another cluster. The PCE spectra are clustered well according to their types except that C3-2 is incorrectly clustered. The potential reason is that the spectrum was not operated properly during collection. PCEs have different molecular structure, so the intensity and location of infrared absorption peaks are slightly different. Therefore, the PCEs can also be classified effectively by the cluster method. This further shows that the feature selection method based PCE can effectively highlight the difference between different water-reducing agents.
3.7. Similarity Analysis of ATR-FTIR Spectra of PCE Samples. Figures 12 and 13 show similarity calculation results by using the angle cosine formula and correlation coefficient formula, respectively. To establish the correlation between different PCE samples, the similarities of infrared


Figure 8: Distribution of the load factor with the wave number: (a) PC1, (b) PC2, (c) PC3, and (d) PC4.

Table 3: Rotating principal component factor score.

| No. | Wave number $\left(\mathrm{cm}^{-1}\right)$ | Factor 1 | Factor 2 | Factor 3 | Factor 4 | Upper bound $\left(\mathrm{cm}^{-1}\right)$ | Lower bound $\left(\mathrm{cm}^{-1}\right)$ |
| :--- | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| 1 | 2859 | -0.002 | $\mathbf{0 . 0 1 5}$ | -0.006 | -0.006 | 2755 | 3000 |
| 2 | 1730 | -0.020 | -0.022 | $-\mathbf{0 . 0 6 0}$ | -0.025 | 1681 |  |
| 3 | 1573 | $\mathbf{0 . 0 4 7}$ | 0.005 | 0.046 | -0.006 | 1504 |  |
| 4 | 1466 | 0.005 | $\mathbf{0 . 0 1 5}$ | 0.003 | -0.005 | 1428 | 1652 |
| 5 | 1409 | $\mathbf{0 . 0 2 6}$ | 0.005 | 0.014 | 0.002 | 1385 | 1504 |
| 6 | 1342 | -0.002 | $\mathbf{0 . 0 3 6}$ | 0.013 | 0.001 | 1318 | 1427 |
| 7 | 1278 | -0.002 | $\mathbf{0 . 0 2 2}$ | -0.005 | 0.005 | 1263 | 1223 |
| 8 | 1241 | -0.006 | $\mathbf{0 . 0 1 4}$ | -0.016 | -0.002 | 1067 | 1298 |
| 9 | 1093 | -0.019 | $\mathbf{0 . 1 1 8}$ | 0.062 | 0.036 | 1258 |  |
| 10 | 961 | -0.008 | $\mathbf{0 . 0 3 5}$ | -0.006 | 0.013 | 1132 |  |
| 11 | 839 | -0.021 | $\mathbf{0 . 0 4 1}$ | -0.006 | -0.001 | 882 | 1032 |

spectrum of each PCE sample from all batches were evaluated.

As shown in Figures 12 and 13, the similarities (correlation coefficient value) of all 16 samples were all greater
than 0.98 . In addition, the results show that the similarity of the same PCE, even from different batches, is close to 1 as expected, indicating that the chemical composition is basically the same. However, the PCE cannot be accurately


Figure 9: Comparison of the test set and SVM prediction results (RBF kernel function).


Figure 10: SVM matrix results of individual class in percentage.
matched by the above two formulas. For example, the similarity between D3 and D4 (0.998) from the same manufacturer is the same as C1 and D3 (0.998) from different manufacturers by using the angle cosine formula. Similar results are also found in the calculation results using the correlation coefficient formula. This is because the two formulas have some limitations, that is, they are not sensitive to the concentration. For this reason, other similarity calculation methods are needed to be introduced.

In order to improve the sensitivity to the concentration and increase the discrimination between different samples, a new similarity calculation method is introduced in this work, which is shown as

$$
\begin{equation*}
S=1-\frac{\sum_{i=1}^{n}\left|\left(X_{i}-Y_{i}\right) /\left(X_{i}+Y_{i}\right)\right|}{\left(n\left(\sum_{i=1}^{n} X_{i}+\sum_{i=1}^{n} Y_{i}\right)\right)^{k}} \tag{2}
\end{equation*}
$$

where $S$ means similarity between measured PCE spectral and standard PCE spectral, $X_{i}$ means peak area of the measured PCE sample, $Y_{i}$ means peak area of the standard PCE sample, $n$ means total absorption peak number of the standard sample spectrum, and $k$ means shape parameter, $k=0.3$ in this work, specifically.

In PCA analysis, the characteristic bands of the PCE spectra were extracted, and the bands correspond to the absorption peaks in PCE spectra, indicating the difference between different PCEs. Therefore, the area under the absorption peak corresponding to the characteristic band is substituted into equation (2). On this basis, similarity calculation results are shown in Figure 14. As expected, the similarity between PCEs from the same manufacturer is basically greater than 0.9 , whereas PCEs from different manufacturers is less than 0.8 , which shows that the new similarity calculation formula has better distinguishing ability. When the PCE product received by the construction site changes, the new formula can choose to reject the product according to the change of some absorption peak area.

The similarity analysis should have an upper and lower bound, making it transferable between dataset and methods [32]. As mentioned earlier, NJDOT currently uses 0.975 as the acceptance standard threshold for concrete admixtures. However, the fluctuation of each PCE is different, and the specific threshold should be determined for each PCE. In this work, the three sigma method is used to establish target threshold, as shown in

$$
\begin{equation*}
\mathrm{AT}=\mu-3 \sigma \tag{3}
\end{equation*}
$$



Figure 11: Dendrogram of PCEs based on selected variables.


Figure 12: Similarity calculation results using the angle cosine formula.
where $\mu$ means the average of similarities and $\sigma$ means standard deviation of similarities involved in calculation.

Therefore, the Acceptance Threshold (AT) corresponding to each PCE is given in Table 4. In this case, the AT
of PCE-D is 0.883 , indicating the quality fluctuation of this product is large. In addition, the AT of PCE-A, PCE-B, and PCE-C is $0.965,0.942$, and 0.928 , respectively. Therefore, it is more accurate to determine a specific AT for each PCE.


FIGURE 13: Similarity calculation results using the correlation coefficient formula.


Figure 14: Similarity calculation results utilization.

Table 4: The AT corresponding to each PCE.

| Type of PCE | $\mu$ | $\sigma$ | AT |
| :--- | :---: | :---: | :---: |
| A | 0.983 | 0.006 | 0.965 |
| B | 0.972 | 0.010 | 0.942 |
| C | 0.976 | 0.016 | 0.928 |
| D | 0.952 | 0.023 | 0.883 |

## 4. Conclusions

This study developed a practicable alternative method for quality inspection of PCE. Specifically, the methods of spectra collection, feature extraction, and cluster analysis were discussed in this study. Moreover, the specific operation steps of PCE quality inspection were also proposed. Based on the experimental result and analysis, the following conclusions can be drawn:
(i) FTIR analysis indicated that a simple heating method can be used to prepare samples for collecting high-quality PCE spectra.
(ii) The feature extraction method based on PCA can effectively reserve significant information bands while reducing the load of calculation. The SVM and cluster analysis results further confirm the validity of the feature extraction method and optimization method.
(iii) Similarity calculation results indicated that the new similarity calculation method not only has better classification performance but also has good generalization ability compared with the conventional method.
(iv) The acceptance threshold determination method based on the $3 \sigma$ principle determined the specific threshold for each PCE rather than unified threshold, which is much more rational and in accordance with the actual conditions.
(v) It should be noted that it is not enough to establish the acceptance threshold only by 4 batches. In practical application, the correlation between sample number and acceptance threshold should be further explored. In addition, concrete admixtures are not limited to PCE. Future research studies should further verify the feasibility of applying this method to air entraining agent, retarder, and other concrete admixtures.

## Data Availability

The PCE spectra data used to support the findings of this study are available from the corresponding author upon request.

## Conflicts of Interest

The authors declare that they have no conflicts of interest regarding the publication of this paper.

## Acknowledgments

This work was supported by the National Natural Science Foundation of China (51408287, 51668038, and 51868042), Distinguished Young Scholars Fund of Gansu Province (1606RJDA318), Natural Science Foundation of Gansu Province (1506RJZA064), Industry Support and Guidance Project by University and College in Gansu Province (2020C-13), and Foundation of a Hundred Youth Talents Training Program of Lanzhou Jiaotong University.

## References

[1] A. Tramaux, N. Azéma, Y. El Bitouri et al., "Synthesis of phosphonated comb-like copolymers and evaluation of their dispersion efficiency on $\mathrm{CaCO}_{3}$ suspensions part II: effect of macromolecular structure and ionic strength," Powder Technology, vol. 334, pp. 163-172, 2018.
[2] A. Colombo, M. R. Geiker, H. Justnes, R. A. Lauten, and K. De Weerdt, "On the effect of calcium lignosulfonate on the rheology and setting time of cement paste," Cement and Concrete Research, vol. 100, pp. 435-444, 2017.
[3] J. Liu, C. Yu, X. Shu, Q. Ran, and Y. Yang, "Recent advance of chemical admixtures in concrete," Cement and Concrete Research, vol. 124, pp. 105-834, 2019.
[4] Y. Hong, J. Irudayaraj, and M. M. Paradkar, "Discriminant analysis of edible oils and fats by FTIR, FT-NIR and FTRaman spectroscopy," Food Chemistry, vol. 93, no. 1, pp. 25-32, 2005.
[5] A. J. Casale, J. Doukakis, H. Najm, and K. Davis, "Quantitative assessment of infrared analysis of concrete admixtures," International Journal of Concrete Structures and Materials, vol. 7, no. 3, pp. 203-214, 2013.
[6] S. Tong, Z. Yuqi, and W. Qiang, "Recent advances in chemical admixtures for improving the workability of alkali-activated slag-based material systems," Construction and Building Materials, vol. 124, 2020.
[7] Y. Ma, C. Shi, L. Lei et al., "Research progress on polycarboxylate based superplasticizers with tolerance to clays - a review," Construction and Building Materials, vol. 255, 2020.
[8] H. Tian, X. Kong, Y. Cui, Q. Wang, and D. Wang, "Effects of polycarboxylate superplasticizers on fluidity and early hydration in sulfoaluminate cement system," Construction and Building Materials, vol. 228, 2019.
[9] S. Sha, M. Wang, C. Shi, and Y. Xiao, "Influence of the structures of polycarboxylate superplasticizer on its performance in cement-based materials-a review," Construction and Building Materials, vol. 233, 2020.
[10] X. Lin, H. Pang, D. Wei, M. Lu, and B. Liao, "Effect of the cross-linker structure of cross-linked polycarboxylate, superplasticizers on the behavior of cementitious mixtures," Colloids and Surfaces A: Physicochemical and Engineering Aspects, vol. 608, 2020.
[11] F. Autelitano, "The odour fingerprint of bitumen," Road Materials \& Pavement Design, vol. 18, no. Suppl.2, pp. 1-11, 2017.
[12] F. S. Hashemi-Nasab and H. Parastar, "Pattern recognition analysis of gas chromatographic and infrared spectroscopic fingerprints of crude oil for source identification," Microchemical Journal, vol. 153, pp. 104-326, 2020.
[13] N. Shekari, M. Vosough, and K. Tabar Heidar, "Chromatographic fingerprinting through chemometric techniques for
herbal slimming pills: a way of adulterant identification," Forensic Science International, vol. 286, pp. 213-222, 2018.
[14] M. Mohammadi, M. K. Khorrami, A. Vatani et al., "Genetic algorithm based support vector machine regression for prediction of SARA analysis in crude oil samples using ATRFTIR spectroscopy," Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy, vol. 245, 2020.
[15] H. Tiernan, B. Byrne, S. G. Kazarian et al., "ATR-FTIR spectroscopy and spectroscopic imaging for the analysis of biopharmaceuticals," Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy, vol. 241, pp. 118-636, 2020.
[16] L. Zhihui, Z. Rui, Z. Yonghua, C. Qian, and Q. Weijun, "Discriminating wavenumbers selection of ATR-FTIR spectra for identifying graded asphalt," Construction and Building Materials, vol. 214, pp. 565-573, 2019.
[17] X. Sun, H. Yuan, C. Song et al., "Rapid and simultaneous determination of physical and chemical properties of asphalt by ATR-FTIR spectroscopy combined with a novel calibration-free method," Construction and Building Materials, vol. 230, 2020.
[18] C. Schramm, "High temperature ATR-FTIR characterization of the interaction of polycarboxylic acids and organotrialkoxysilanes with cellulosic material," Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy, vol. 243, 2020.
[19] American Standard Test Methods, Standard Specification for Chemical Admixtures for Concrete, ASTM C494/C494M-13, West Conshohocken, PA, USA, 2013.
[20] M. Mohammadi, M. K. Khorrami, A. Vatani et al., "Rapid determination and classification of crude oils by ATR-FTIR spectroscopy and chemometric methods," Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy, vol. 232, pp. 118-157, 2020.
[21] V. Pestov, "Is thek-NN classifier in high dimensions affected by the curse of dimensionality?" Computers \& Mathematics with Applications, vol. 65, no. 10, pp. 1427-1437, 2013.
[22] H. Guttmann-Beck, R. Rozen, and M. Stern, "Vertices removal for feasibility of clustered spanning trees," Discrete Applied Mathematics, vol. 2020, pp. 1-17, 2020.
[23] H. Ling, C. Qian, W. Kang, C. Liang, and H. Chen, "Combination of support vector machine and K-fold cross validation to predict compressive strength of concrete in marine environment," Construction and Building Materials, vol. 206, pp. 355-363, 2019.
[24] R. Ren, K. Han, P. Zhao et al., "Identification of asphalt fingerprints based on ATR-FTIR spectroscopy and principal component-linear discriminant analysis," Construction and Building Materials, vol. 198, pp. 662-668, 2019.
[25] M. A. Dritschel, D. Estévez, and D. Yakubovich, "Resolvent criteria for similarity to a normal operator with spectrum on a curve," Journal of Mathematical Analysis and Applications, vol. 463, no. 1, pp. 345-364, 2018.
[26] N. M. O'Boyle and R. A. Sayle, "Comparing structural fingerprints using a literature-based similarity benchmark," Journal of Cheminformatics, vol. 8, no. 1, p. 36, 2016.
[27] K. Wang, L. Cheng, and B. Yong, "Spectral-similarity-based kernel of SVM for hyperspectral image classification," Remote Sensing, vol. 12, no. 13, p. 2154, 2020.
[28] M. Xu, Y. Zhang, P. Zhao, and C. Liu, "Study on aging behavior and prediction of SBS modified asphalt with various contents based on PCA and PLS analysis," Construction and Building Materials, vol. 265, 2020.
[29] S. Chen, S. Sun, X. Chen et al., "Effects of core-shell polycarboxylate superplasticizer on the fluidity and hydration
behavior of cement paste," Colloids and Surfaces A: Physicochemical and Engineering Aspects, vol. 590, 2020.
[30] M. Lopes, V. Mouillet, L. Bernucci, and T. Gabet, "The potential of attenuated total reflection imaging in the mid-infrared for the study of recycled asphalt mixtures," Construction and Building Materials, vol. 124, pp. 1120-1131, 2016.
[31] X. Lin, B. Liao, J. Zhang, S. Li, J. Huang, and H. Pang, "Synthesis and characterization of high-performance crosslinked polycarboxylate superplasticizers," Construction and Building Materials, vol. 210, pp. 162-171, 2019.
[32] E. Källman, M. G. Delcey, M. Guo, R. Lindh, and M. Lundberg, "Quantifying similarity for spectra with a large number of overlapping transitions: examples from soft X-ray spectroscopy," Chemical Physics, vol. 535, 2020.

