

## تحلیل آماری داده‌های چند متغیره‌ی حاصل از بیناب‌نمایی فروشکست القایی لیزری و تقلیل ابعاد داده‌ها با استفاده از روش تحلیل مولفه اصلی به منظور شناسایی پلیمرها

محمد وحید دستجردی<sup>۱</sup>، سید جبار موسوی<sup>۲</sup>، سید محمد رضا دربانی<sup>۲</sup>، علی موسوی آذر<sup>۲</sup>، محمود سلطان‌الكتابی<sup>۱</sup>

<sup>۱</sup>گروه فیزیک، دانشگاه اصفهان، اصفهان

<sup>۲</sup>پژوهشکده علوم و فناوری اپتیک و لیزر، دانشگاه صنعتی مالک اشتر، اصفهان

چکیده- در این پژوهش روش تحلیل مولفه اصلی (PCA) به عنوان یکی از روش‌های تحلیل آماری داده‌های چند متغیره برای تقلیل بعد داده‌ها و همچنین برای یافتن یک الگو در داده‌های حاصل از بیناب LIBS به منظور تفکیک نمونه‌ها استفاده شد. ابتدا ۵ نمونه پلیمری تهییه و سپس شدت متناظر با ۱۷ طول موج مرتبط با خطوط گسیلی اتمی و یونی و باندهای مولکولی برای هر نمونه به عنوان ماتریس ورودی در روش PCA انتخاب شد. نتایج تحقیق نشان داد که تنها با ۳ متغیر خروجی از روش PCA پلیمرهای مورد آنالیز را می‌توان تشخیص و ابعاد داده‌ها را نیز کاهش داد.

کلید واژه- بیناب‌نمایی فروشکست القایی لیزری، تحلیل مولفه اصلی، شناسایی پلیمرها، تقلیل ابعاد داده‌ها.

### Statistical multivariate analysis in laser induced breakdown spectroscopy (LIBS) data and reduction dimensions of data by principal component analysis (PCA) with aim of polymers recognition

M. Vahid Dastjerdi<sup>1</sup>, S. J. Mousavi<sup>2</sup>, S. M. R. Darbani<sup>2</sup>, A. Mousavi Azar<sup>2</sup>, M. Soltanolkotabi<sup>1</sup>

<sup>1</sup> Department of Physics, University of Isfahan, Isfahan

<sup>2</sup> Optics & Laser Science and Technology Research Center, Malek Ashtar University of Technology, Isfahan

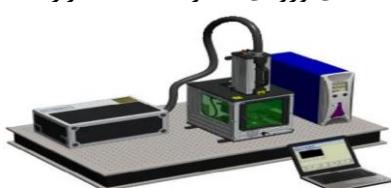
**Abstract-** In this research principal component analysis as one of statistical multivariate analysis methods with aim of reduction dimensions of data and also finding a pattern in laser induced breakdown spectroscopy data in order to classify of samples was applied. At first 5 polymer samples was prepared and then intensity corresponding to 17 wavelengths relating to atomic and ionic lines and molecular bands for each sample as input matrix in PCA method, was selected. The results show that with only 3 output variables from PCA method, analysis polymers can be recognized from each other and also dimensions of data, was reduced.

**Keywords:** Laser induced breakdown spectroscopy, Principal component analysis, Polymers recognition, Reduction dimensions of data.

پلی استایرن (PS)<sup>۶</sup> و پلی وینیل کلراید (PVC)<sup>۷</sup> که به طور معمول در ساخت انواع مواد پلاستیکی به کار می روند به صورت گرانول تهیه شد. سپس گرانول نمونه ها با استفاده از سیستم فشار هیدرولیکی به قرص های هم شکل و هم اندازه با سطحی کاملاً صاف با قطر ۱۳ میلی متر و ضخامت ۳-۴ میلی متر تبدیل گردید تا همه نمونه ها در شرایط کاملاً مشابه آنالیز شده و تنها عامل تفاوت در بیناب ثبت شده از پلیمرهای مختلف، تفاوت ساختاری آن ها باشد.

### چیدمان آزمایش

به منظور آنالیز نمونه ها در این پژوهش از سامانه LIBSCAN100 ساخت شرکت Applied Photonics استفاده شده است. این سامانه مجهز به لیزر Nd:YAG با انرژی خروجی متغیر تا  $100\text{ mJ}$ ، پهنه ای تپ  $7\pm2\text{ ns}$  فرکانس تکرار متغیر  $1\text{-}20\text{ Hz}$  بوده و بیناب پلاسمای تولیدی توسط آشکارساز مدل ۱۳-A ساخت شرکت Avantes ثبت می شود. این سامانه قابلیت بیناب نگاری در ناحیه طول موجی  $180\text{ nm}$  تا  $1057\text{ nm}$  را دارد. بوده که در بازه های مختلف طول موجی دارای توان تفکیک پذیری متغیر  $0.04\text{ nm}$  تا  $0.15\text{ nm}$  می باشد. نمایی از این سامانه در شکل ۱ نشان داده شده است. از هر نمونه پلیمری  $28\text{ }\mu\text{m}$  بیناب توسط این سامانه ثبت و بعد از میانگین گیری روی هر  $4\text{ }\mu\text{m}$  و تهیه  $7\text{ }\mu\text{m}$  بیناب نهایی برای هر پلیمر، شدت های متناظر با  $17\text{ }\mu\text{m}$  طول موج مرتبط با خطوط گسیلی اتمی و یونی و باندهای مولکولی در ماتریس داده های ورودی به فرایند PCA قرار داده شد.



شکل ۱: سامانه بیناب نمایی ۱۰۰ LIBSCAN

### ۳-۲- تحلیل مولفه اصلی

تحلیل مولفه اصلی یکی از روش های تحلیل آماری چند متغیره است که برای تقلیل بعد داده ها، شناسایی نقاط پرت احتمالی در داده های چند متغیره و کاهش پیچیدگی

### ۱- مقدمه

بیناب نمایی فروشکست القایی لیزری (LIBS)<sup>۱</sup>، یکی از روش های بیناب نمایی گسیل اتمی بوده که در آن از تپ لیزری به منظور تولید پلاسمای هدف استفاده می گردد<sup>[۱]</sup>. تشخیص انواع پلیمرها از یکدیگر یکی از نیازهای اساسی در صنعت بازیافت است. شناسایی پلیمرها با استفاده از بررسی وابستگی شدت های نسبی خطوط گسیلی بیناب LIBS به نسبت های استکیومتری متناظر با ساختار شیمیایی آنها همواره با چالش های متعددی روبرو است. به همین دلیل در سال های اخیر استفاده از تحلیل LIBS آماری داده های چند متغیره برای آنالیز بیناب نمونه ها افزایش پیدا کرده است<sup>[۲]</sup>. روش های آماری چند متغیره به روش های اطلاق می شود که به تحلیل توام (متغیرها) وجود دارد. تحلیل مولفه اصلی، تحلیل آیتمی و تحلیل عاملی برخی از روش های مدرن آنالیز چند متغیره هستند. هدف مشترک اغلب روش های چند متغیره کاهش بعد داده ها و در نتیجه قابل فهم کردن اطلاعات موجود در داده ها است<sup>[۳]</sup>. در این پژوهش به منظور تشخیص انواع پلیمرها از هم با استفاده از بیناب نمایی فروشکست القایی لیزری، در ابتدا رابطه بین نسبت  $\text{C}/\text{C}_2$  در حالت تئوری و تجربی بررسی شد. سپس با تعریف متغیرهای جدید با استفاده از روش PCA<sup>۲</sup> و تقلیل بعد داده ها به یافتن یک الگو در داده ها برای شناسایی و تفکیک نمونه ها پرداخته شد.

### ۲- مواد و روش ها

در این بخش نمونه های مورد آنالیز، چیدمان تجربی و همچنین روش آماری استفاده شده در این پژوهش، تشریح شده است.

### نمونه های مورد آنالیز

ابتدا پنج نمونه از پلیمرها شامل پلی اتیلن (PE)، پلی متیل متاکریلات (PMMA)<sup>۳</sup>، پلی پروپیلن (PP)<sup>۴</sup>

<sup>1</sup> Laser Induced Breakdown Spectroscopy

<sup>2</sup> Principal Component Analysis

<sup>3</sup> Polyethylene

<sup>4</sup> Polymethyl methacrylate

<sup>5</sup> Polypropylene

<sup>6</sup> Polystyrene

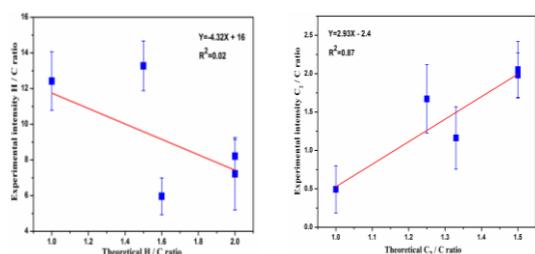
<sup>7</sup> Polyvinyl chloride

محاسبه نسبت تجربی نیز از شدت خطوط بینابی C<sub>2</sub>(473.6nm) و H<sub>a</sub>(656.24nm) استفاده شد.

جدول ۱: نسبت‌های تئوری پلیمرهای مورد آنالیز

نمونه پلاستیک	نسبت H/C تئوری	نسبت C <sub>2</sub> /C تئوری	ساختار شیمیایی
PVC	1.5	1.5	
PMMA	1.6	1	
PS	1	1.25	
PE	2	1.5	
PP	2	1.33	

با برازش بهترین خط بر داده‌ها منحنی‌های کالیبراسیون ترسیم و در شکل ۲ نشان داده شده است. همان‌طور که شکل‌ها نشان می‌دهند یک رابطه خطی بین نسبت‌های تئوری و تجربی وجود نداشته که خود مانع شناسایی نمونه‌ها بر حسب این نسبت‌ها می‌شود.

شکل ۲: منحنی کالیبراسیون برای نسبت‌های شدت H/C و C<sub>2</sub>/C

اثرات ماتریسی شدید می‌تواند یکی از دلایل پراکندگی این نقاط نسبت به خط برازش شده باشد. دلیل دیگر می‌تواند مربوط به این واقعیت باشد که همه‌ی شدت گسیلی مربوط به خط کریم مربوط به مولکول‌های سازنده پلیمر نیستند و بخشی از آن مربوط افزودنی‌های موجود در پلیمر است. این محدودیتها شک ما در مورد عدم کارایی منحنی کالیبراسیون برای شناسایی و طبقه‌بندی

تحلیل متغیرهای اولیه مسئله به کار می‌رود. همچنین این روش یک تکنیک آماری بسیار کارآمد برای یافتن یک الگو در داده‌های با تعداد متغیرهای زیاد به منظور تفکیک نمونه‌ها است [۳].

### ۱-۳-۲ ساخت ماتریس داده‌ها و کوواریانس نمونه

تعداد اعضای نمونه (۳۵) با n و تعداد متغیرها (شدت‌های متناظر با ۱۷ طول موج مختلف) با p نشان داده می‌شود. ماتریس داده‌های نمونه که با X نشان داده می‌شود ماتریسی n×p است که نمونه‌ها روی سطر این ماتریس و متغیرها روی ستون آن قرار می‌گیرند.

$$X = \begin{pmatrix} x_{11} & \dots & x_{1p} \\ \vdots & \ddots & \vdots \\ x_{n1} & \dots & x_{np} \end{pmatrix} = \begin{pmatrix} x_1^T \\ \vdots \\ x_n^T \end{pmatrix}$$

ماتریس کوواریانس نمونه که آن را با S نمایش می‌دهیم ماتریسی p×p و متقارن است که درایه سطر i آم و ستون j آم آن در واقع کوواریانس نمونه بین متغیر i آم و j آم بوده و به طریق زیر تعریف می‌شود :

$$s_{ij} = \frac{1}{n-1} \sum_{r=1}^n (X_{ri} - \bar{X}_i) (X_{rj} - \bar{X}_j)$$

### ۲-۳-۲ تعریف مولفه‌های اصلی

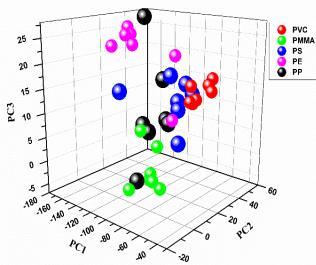
مولفه‌های اصلی، p متغیر جدیدند که آنها را با Y<sub>1</sub>, Y<sub>2</sub>, ..., Y<sub>p</sub> نشان می‌دهیم. هر یک از متغیرهای جدید یک ترکیب خطی از p متغیر اولیه هستند که بر اساس ویژه بردارهای ماتریس کوواریانس داده‌های اولیه تعریف می‌شوند. فرض کنید λ<sub>1</sub>, λ<sub>2</sub>, ..., λ<sub>p</sub> ویژه‌مقدادر مرتب شده ماتریس S هستند، به طوریکه λ<sub>1</sub> ≤ λ<sub>2</sub> ≤ ... ≤ λ<sub>p</sub>. فرض کنید q<sub>1</sub>, q<sub>2</sub>, ..., q<sub>p</sub> بردارهای ویژه یکه متعامد به ترتیب متناظر با λ<sub>1</sub>, λ<sub>2</sub>, ..., λ<sub>p</sub> هستند. در این صورت مولفه اصلی آم به صورت زیر تعریف می‌شود [۳] :

$$Y_i = q_i^T X = q_{i1} X_1 + q_{i2} X_2 + \dots + q_{ip} X_p \quad q_i = (q_{i1}, q_{i2}, \dots, q_{ip})^T$$

### ۳- نتایج و بحث

در ابتدا ارتباط بین نسبت‌های H/C و C<sub>2</sub>/C در حالت تئوری و تجربی بررسی شد. نسبت‌های تئوری H/C برای هر نمونه با استفاده از فرمول مولکولی آن محاسبه و نتایج در جدول ۱ نشان داده شده است. برای

همپوشانی دارند. در صفحه (PC1-PC2) نیز همه پلیمرها با انکی همپوشانی از هم تفکیک می‌شوند. همچنین تصویر داده‌های اولیه در فضای سه بعدی (PC1-PC2-PC3) برای هر مجموعه از پلیمرها در شکل ۷ رسم شده است.



شکل ۷: تصویر داده‌ها در فضای سه بعدی

همانطور که شکل ۷ نشان می‌دهد، در فضای سه بعدی (PC1-PC2-PC3) این پلیمرها با همپوشانی بسیار کمی از هم جدا می‌شوند. البته همواره نقاط پرتی وجود دارد که از هم گروهی‌های خود در فواصل دوری قرار دارند که بایستی این نقاط را باید به عنوان داده پرت حذف کیم. بنابراین به روش PCA با استفاده از ۳ متغیر جدید (مولفه‌های اصلی اول تا سوم) به جای استفاده از ۱۷ متغیر اولیه ابعاد داده‌ها تقلیل داده شد و هم پلیمرهای مورد آزمایش از هم شناسایی و تفکیک شدند.

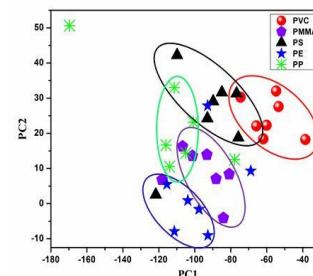
#### ۴- نتیجه‌گیری

هدف اصلی این تحقیق بررسی توانایی LIBS در شناسایی و تفکیک پلیمرهای با ساختار مشابه است. به همین منظور دو روش کالیبراسیون و تحلیل مولفه اصلی مورد بررسی قرار گرفت. نتایج تحقیق نشان داد که روش کالیبراسیون در تعیین ساختار پلیمرهای انتخاب شده در این تحقیق کارآمد نبوده، در صورتی که روش PCA علاوه بر تقلیل ابعاد داده‌ها به شناسایی و تفکیک پلیمرهای با ساختار مشابه کمک می‌کند.

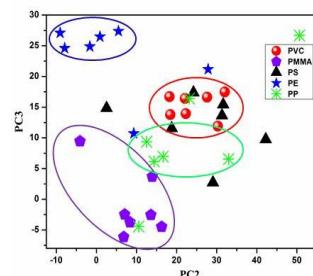
#### مراجع

- [1] Noll, R., *Laser-induced breakdown spectroscopy*, Springer, 2012.
- [2] Godoi, Q., et al. "Preliminary studies of laser-induced breakdown spectrometry for the determination of Ba, Cd, Cr and Pb in toys." *Spectrochimica Acta Part B: Atomic Spectroscopy*. 64(6): 573-581, 2009.
- [3] Johnson, R. A. and D. W. Wichern., *Applied multivariate statistical analysis*, Prentice hall Englewood Cliffs, NJ, 1992.

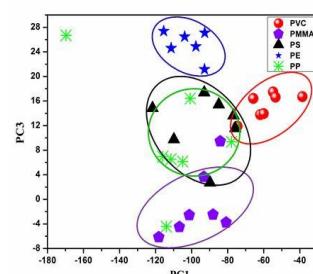
پلیمرها مورد استفاده در این آزمایش را تأیید می‌کند و ما را به استفاده از روش‌های تحلیل آماری داده‌های چند متغیره برای طبقه‌بندی پلیمرها ترغیب کرد. به همین منظور پس از انجام فرایند PCA در نرم افزار Matlab و تعیین ویژه‌مقادیر و ویژه‌بردارهای ماتریس کوواریانس نمونه، تصویر داده‌های اولیه در صفحه مولفه‌های اصلی اول و دوم (PC1-PC2)، دوم و سوم (PC2-PC3) و اول و سوم (PC1-PC3) ترسیم و به ترتیب در شکل‌های ۴، ۵ و ۶ نشان داده شده است.



شکل ۴: تصویر داده‌ها در صفحه (PC1-PC2)



شکل ۵: تصویر داده‌ها در صفحه (PC2-PC3)



شکل ۶: تصویر داده‌ها در صفحه (PC1-PC3)

همانطور که شکل‌ها نشان می‌دهند در صفحه (PC1-PC3) پلیمرهای PE و PVC با PMMA و PP به خوبی قابل تفکیک اما پلیمرهای PS و PP با همپوشانی دارند. در صفحه (PC2-PC3) پلیمرهای PMMA و PE به خوبی PMMA با PE تفکیک می‌شوند. در صفحه (PC1-PC3) پلیمرهای PVC و PMMA به خوبی قابل تفکیک ولیکن سایر پلیمرها در این صفحه با هم