بیست و هشتمین کنفرانس اپتیک و
فوتونیک ایران و چهاردهمین کنفرانس
مهندسی و فناوری فوتونیک ایران،

دانشگاه شهید چمران اهواز،

خوزستان، ایران.

۱۴-۱۲ بهمن ۱۴۰۰

مطالعه اثر ترکیب تحلیل مولفه اصلی با روش آماری مدل رگرسیون بردار پشتیبان در تکنیک اسپکتروسکوپی فروشکست القاییده لیزری

فاطمه رضائی[۱]، محسن رضائی[۲]، پروین کریمی[۳]

۱. دانشکده فیزیک، دانشگاه خواجه نصیرالدین طوسی، تهران- ایران، fatemehrezaei@kntu.ac.ir

۲. گروه مهندسی صنایع، دانشگاه علم و فناوری مازندران، بهشهر- ایران،
mohsen.rezaei@mazust.ac.ir

۳. گروه فیزیک، واحد تهران جنوب، دانشگاه آزاد اسلامی، تهران، ایران، p_karimi@azad.ac.ir

چکیده - در این مقاله، غلظت عناصر سازنده هفت آلیاژ آلومینیوم استاندارد با استفاده از روش‌های آماری مدل رگرسیون بردار پشتیبان (SVR و KSVR) و روش ترکیب تحلیل مولفه اصلی با مدل رگرسیون بردار پشتیبان (PCA_KSVR و PCA_SVR) در تکنیک طیف‌سنجی فروشکست القاییده لیزری پیش‌بینی شد. نتایج آنالیزها نشان داد که روش PCA_KSVR دقیق‌ترین غلظت را با حداقل خطا برای عنصر آهن گزارش کرده است.

کلید واژه- اسپکتروسکوپی فروشکست القاییده لیزری، پیش‌بینی غلظت، تحلیل مولفه اصلی، رگرسیون بردار پشتیبان.

Study of the effect of combination of principle component analysis with statistical method of support vector regression in laser induced breakdown spectroscopy

Fatemeh Rezaei [1], Mohsen Rezaei [2], and Parvin Karimi [3]

1) Department of Physics, K. N. Toosi University of Technology, Tehran, Iran, fatemehrezaei@kntu.ac.ir

2) Groups of Industrial Engineering, University of Science and Technology of Mazandaran, Behshahr, Iran, mohsen.rezaei@mazust.ac.ir

3) Department of Physics, South Tehran Branch, Islamic Azad University, Tehran, Iran, p_karimi@azad.ac.ir

Abstract- In this paper, concentration of constituent elements of seven standard aluminum'S alloys is predicted by using statistical methods of support vector regression model (SVR and KSVR) and the combinational method of principal component analysis with support vector regression model (PCA_SVR and PCA-KSVR) in laser induced breakdown laser spectroscopy technique. The results of the analyzes showed that the PCA_KSVR method reported the most accurate concentration with the least error for Fe element.

Keywords: Laser induced breakdown spectroscopy, Concentration prediction, Principle component analysis, support vector regression.

[Downloaded from opsi.ir on 2024-04-20]

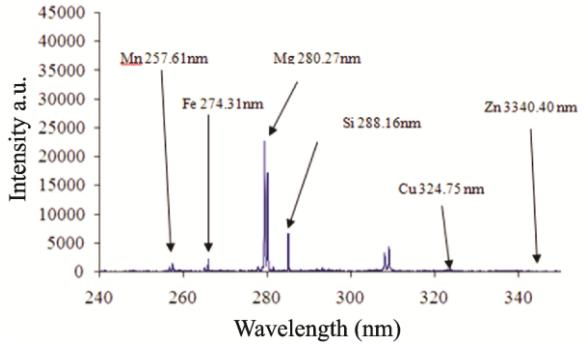
مقدمه

برآورد دقیق غلظت عناصر با استفاده از تکنیک طیف-سنجی فروشکست القاییده لیزری، فرایندی است که از دیرباز مورد توجه فیزیکدانان و شیمیدانان زیادی بوده است. به طور کلی آنالیز کمی را می‌توان در تکنیک طیف‌سنجی فروشکست القاییده لیزری با روش‌های متداولی از قبیل منحنی کالیبراسیون، نسبت شدت دو خط و معادلات ساها و ... انجام داد. شایان ذکر است که در سال‌های اخیر روش‌های آماری متعددی از قبیل شبکه عصبی مصنوعی ANN، رگرسیون بردار پشتیبان SVR، و مدل‌های رگرسیون چندگانه MLR جهت پیش‌بینی غلظت در تکنیک طیف‌سنجی فروشکست القاییده لیزری استفاده شده‌اند.

در این مقاله، اثر ترکیب تکنیک تحلیل مولفه اصلی با روش آماری مدل رگرسیون بردار پشتیبان بررسی می‌شود تا به بررسی نتایج کاهش ابعاد در آنالیز کمی تکنیک LIBS پرداخته شود.

چیدمان آزمایش

در این آزمایش از لیزر پالسی Nd:YAG با نرخ تکرار ۱۰Hz، طول‌موج ۱۰۶۴ nm و پهنای پالس ۶ نانوثانیه جهت تابش آلیاژهای آلومینیوم استفاده شده است. با تابش لیزر بر روی نمونه توسط لنزی با فاصله کانونی ۲۰ سانتی‌متر پلاسماهای آلومینیوم ایجاد می‌شود. تابش‌های پلاسما توسط لنز کوارتز دیگری با فاصله کانونی ۳۵ میلی-متر جمع شده و توسط یک لنز شی‌ای دیگری بر روی فیبر اپتیکی متمرکز می‌شوند. فیبر مذکور به اسپکترومتر اشل متصل می‌باشد که این طیف‌سنج، تابش‌های پلاسما را از لحاظ طول‌موجی تفکیک می‌نماید. سپس این تابش‌ها توسط دوربین ICCD نصب شده بر روی طیف‌سنج، در زمان‌های مختلف مورد آنالیز قرار می‌گیرند. در این تحقیق، از هفت آلیاژ استاندارد آلومینیوم به عنوان نمونه استفاده شده است. به عنوان مثال، نمونه‌ای از طیف دریافتی از آلیاژ استاندارد ۱۱۰۰ در شکل ۱ نشان داده شده است.



شکل ۱: طیف حاصل از آلیاژ استاندارد ۱۱۰۰ در زمان تاخیر یک میکروثانیه.

روش‌های آماری

مدل رگرسیون بردار پشتیبان (SVR) الگوریتم رگرسیون مبتنی بر کرنل است که بر اساس روش ماشین بردار پشتیبان (SVM) ساخته شده است [۱]. ماشین‌های بردار پشتیبان برای مسائل طبقه‌بندی به کار می‌روند و بعدها الگوریتم آنها برای کار با مسائل رگرسیون یا تخمین داده ها توسعه می‌یابد که به این الگوریتم جدید، رگرسیون بردار پشتیبان می‌گویند. در این روش، مجموعه‌ای از داده ها به شکل $(x_1,y_1), \ldots (x_n,y_n)$ به عنوان آموزش داده می-شوند که در این تحقیق، x_nها ورودی‌های طیف و y_nها شدت‌های مربوطه هستند و n نیز تعداد نمونه‌ها می‌باشد. در مدل SVR غیرخطی، داده‌ها به فضایی با ابعاد بزرگتر توسط توابع کرنل انتقال می‌یابند. تابع تخمین‌گر می‌تواند توسط تابع کرنل به صورت تابعی بین ورودی و خروجی با کمترین خطای ممکن ارتباط برقرار کند.

$$y = \sum_{i=1}^{l} (\alpha_n - \alpha_n^*).K(x_n, x) + b \qquad (1)$$

[Downloaded from opsi.ir on 2024-04-20]

در معادله بالا، b جمله بایاس، K تابع کرنل، α_n و α_n^* ضرایب غیرمنفی برای هر مشاهده x_n هستند. اگر تابع کرنل چندجمله‌ای باشد از معادله زیر تبعیت می‌کند:

$$K(x_n,x) = (x_n.x)^d. \qquad d=2, 3, \dots \qquad (۲)$$

همچنین، اگر تابع کرنل از تابع گوسی پیروی کند از معادله زیر محاسبه می‌شود:

$$K(x_n,x_j) = exp\left(-\frac{\|x_n - x_j\|^2}{2\sigma^2}\right). \qquad (۳)$$

در رابطه فوق، x_j ،x_n و به ترتیب بردارهای پشتیان nام و jام هستند و σ پهنای تابع کرنل است.

تحلیل مولفه اصلی (PCA) نیز روشی چند متغیره است که به منظور کاهش ابعاد یا تعداد متغیرهای یک مجموعه چند متغیره استفاده می‌شود تا دقت در اندازه‌گیری‌ها را افزایش دهد. جزئیات این روش آماری در مرجع [۲] توضیح داده شده است. خطای میانگین مربعی (MSE) از معادله (۴) محاسبه می‌شود:

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(y_i - f_i)^2 \qquad (۴)$$

میانگین قدرمطلق خطا (MAE) نیز از معادله (۵) محاسبه می‌گردد [۳]:

$$MSE = \frac{1}{n}\sum_{i=1}^{n}|y_i - f_i| \qquad (۵)$$

که در رابطه‌های بالا، y_i و f_i به ترتیب، داده‌های پیش‌بینی شده در زمان t و داده واقعی در زمان t و نیز n تعداد داده‌ها می‌باشند.

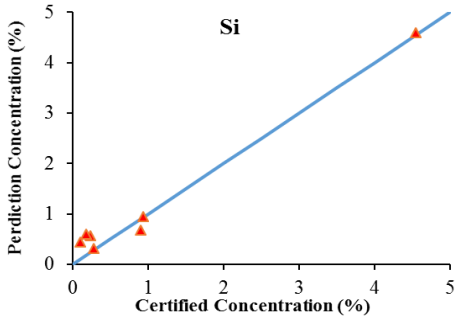
نتایج و بحث

غلظت عناصر مختلف با استفاده از توابع کرنل متفاوت ارزیابی شده‌اند و تابع کرنل مناسب برای هر عنصر انتخاب گردیده است. جدول ۱ نتایج این آنالیزها را نشان می‌دهد و رنگ قرمز بیانگر بهترین پیش‌بینی صورت گرفته با حداقل خطا توسط مطلوب‌ترین تابع کرنل است.

جدول ۱: مقایسه‌ای بین خطاهای محاسبه شده توسط انواع مختلفی از توابع کرنل در پیش‌بینی عناصر مختلف آلیاژ آلومینیوم.

Elements	Linear		Polynomial		Gaussian	
	MSE	MAE	MSE	MAE	MSE	MAE
Fe	0.078	0.280	0.0002	0.013	**0.0001**	**0.012**
Si	**0.012**	**0.109**	**0.012**	**0.109**	0.208	0.456
Mn	0.217	0.465	**0.001**	**0.035**	0.069	0.263
Cu	0.031	0.175	0.230	0.480	**0.002**	**0.045**
Mg	**0.053**	**0.230**	**0.053**	**0.230**	0.059	0.243
Zn	0.523	0.723	0.236	0.486	**0.014**	**0.120**

به عنوان مثال، نتایج پیش‌بینی غلظت عنصر Si در شکل ۱ نشان داده شده است.



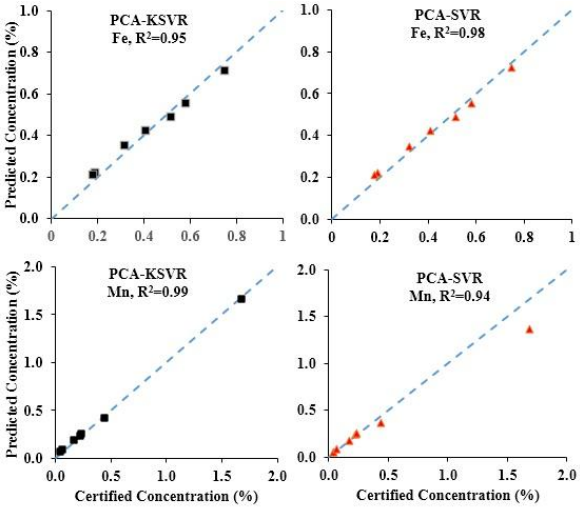
شکل ۱: نتایج پیش‌بینی غلظت Si با روش SVR.

همچنین، غلظت پیش‌بینی شده عناصر Fe و Mn با استفاده از روش‌های آماری PCA-SVR و PCA-KSVR در شکل ۲ نشان داده شده است. همانطور که در شکل ۲ مشاهده می‌شود افزودن تکنیک PCA بر تکنیک

[Downloaded from opsi.ir on 2024-04-20]

رگرسیون بردار پشتیبان سبب پیش‌بینی‌های بسیار دقیقی از غلظت گردیده است.

نتایج میزان خطای اندازه‌گیری شده کلیه عناصر در جدول ۲ ارائه گردیده است که نشان می‌دهد تکنیک PCA-KSVR می‌تواند جایگزین مناسبی بر روش‌های متداول LIBS در آنالیز کمی باشد.



شکل ۲: غلظت پیش‌بینی شده برای دو عنصر Fe و Mn با استفاده از روش‌های PCA-SVR و PCA-KSVR.

نتیجه‌گیری

در این تحقیق، به مطالعه اثر افزودن تحلیل مولفه اصلی بر روش‌های آماری SVR و KSVR در پیش‌بینی غلظت آلیاژهای آلومینیوم پرداخته شده است. نتایج نشان داد که تکنیک PCA-KSVR حداقل خطا را در آنالیز کمی پیش-بینی می‌کند.

جدول ۲: میزان خطای محاسبه شده ناشی از روش‌های آماری SVR، KSVR، PCA-SVR و PCA-KSVR.

Elements	Fe			Zn			Si		
Methods	MSE	RMSE	MAE	MSE	RMSE	MAE	MSE	RMSE	MAE
SVR	0.029	0.17	0.17	2.413	1.553	1.553	0	0	0.02
PCA-SVR	0	0	0.012	0.005	0.07	0.077	0.131	0.361	0.362
KSVR	0	0	0.011	0.014	0.118	0.119	0.011	0.104	0.108
PCA-KSVR	0	0	0.003	0.001	0.031	0.026	0.02	0.141	0.144

Mn			Cu			Mg		
MSE	RMSE	MAE	MSE	RMSE	MAE	MSE	RMSE	MAE
0.054	0.232	0.233	0.02	0.141	0.141	0.052	0.228	0.229
0.005	0.07	0.077	0.032	0.178	0.18	0.034	0.184	0.184
0.001	0.031	0.035	0.002	0.044	0.044	0.052	0.228	0.229
0.001	0.031	0.026	0.001	0.031	0.04	0.033	0.181	0.182

سپاسگزاری

از آقای دکتر سیدحسن توسلی بابت در اختیار قرار دادن تجهیزات آزمایشگاهی‌شان تشکر می‌نماییم.

مرجع‌ها

[1] U.N. Chowdhury, “Integration of principal-component analysis and support vector regression for financial time series forecasting”, Int. J. Comput. Sci, Vol. 15, No. 8, pp. 27-32, 2017.

[2] P. Pořízka, J. Klus, E. Képeš, D. Prochazka, D. W. Hahn, and J. Kaiser, “On the utilization of principal component analysis in laser-induced breakdown spectroscopy data analysis, a review”, Spectrochim. Acta Part B At Spectrosc. Vol. 148, No. 65, pp. , 2018.

[3] T. Lesieur, L. Miolane, M. Lelarge, F. Krzakala, L. Zdeborová, “Statistical and computational phase transitions in spiked tensor estimation”, J. Stat. Phys. arXiv:1701.08010v2, 2017.

[Downloaded from opsi.ir on 2024-04-20]