

REMOTE IDENTIFICATION AND ESTIMATION OF THE CALCIUM CONCENTRATION IN BULK LIQUID UNDER HIGH-PRESSURE CONDITION

J. Sumathi*, V. Sathiesh Kumar, K. Veerappan

Department of Electronics Engineering, Madras Institute of Technology Campus – Anna University, Chennai-600044, Tamil Nadu, India; e-mail: sumathi.j05@gmail.com

Laser-induced breakdown spectroscopy (LIBS) based stand-off distance analysis of the calcium concentration in bulk liquid under high-pressure condition is carried out. The influence of salinity during the determination of the Ca concentration is studied. Machine learning classifiers are used in the estimation of the unknown Ca concentration in the liquid sample, for different experimental parameters or conditions. From the spectral information, the emission lines related to Ca II at 393 and 396 nm are visualized. Chlorine emission lines are not observed due to the requirement of high ionization energies. As the salinity (NaCl) of the sample solution is increased to 2500 ppm, the signal-to-noise ratio of the LIBS signal is improved by a factor of 0.85. The Ca II emission line peak intensity decreased with increase in ambient pressure conditions and stand-off collection distances. The opposite trend is observed with an increase in the laser fluence and the $\text{CaCl}_2 \cdot 2\text{H}_2\text{O}$ concentration in the sample solution. The Ca II 393 nm emission time period is estimated at $(1/e)I_0$ peak intensity. A typical Ca II 393 nm emission time period from 6.712 to 6.766 μs is observed. The specified value is obtained for a laser fluence of 9 J/cm², ambient pressure of 1 atm, stand-off collection distance of 0.6 m, 2500 ppm NaCl, and 1500 ppm $\text{CaCl}_2 \cdot 2\text{H}_2\text{O}$. It is observed that the emission time period increased with increase in the $\text{CaCl}_2 \cdot 2\text{H}_2\text{O}$ concentration in the sample solution and laser fluences. The opposite trend is observed for an increase in the ambient pressure and stand-off distances. The results obtained from the spectral and temporal measurements are in correlation with each other. The best system model's performance metrics of 100% (Accuracy = Precision = Recall = F1-Score) are obtained under the fixed experimental conditions using the k-nearest neighbor classifier.

Keywords: laser-induced breakdown spectroscopy, temporal measurement, remote mineral analysis, machine learning classifiers.

ДИСТАНЦИОННАЯ ИДЕНТИФИКАЦИЯ И ОЦЕНКА КОНЦЕНТРАЦИИ КАЛЬЦИЯ В ОБЪЕМЕ ЖИДКОСТИ В УСЛОВИЯХ ВЫСОКОГО ДАВЛЕНИЯ

J. Sumathi*, V. Sathiesh Kumar, K. Veerappan

УДК 543.42:546.41

Мадрасский технологический институт Университета Анны, Ченнай, 600044, Тамилнад, Индия; e-mail: sumathi.j05@gmail.com

(Поступила 6 июля 2020)

С помощью лазерно-искровой эмиссионной спектроскопии (LIBS) осуществлен удаленный анализ концентрации кальция в объеме жидкости в условиях высокого давления. Изучено влияние солености при определении концентрации Ca. Классификаторы машинного обучения использованы для оценки неизвестной концентрации Ca в жидкой пробе для различных экспериментальных параметров или условий. Наблюдаются эмиссионные линии Ca II при 393 и 396 нм. Линии излучения хлора не наблюдаются из-за требований высоких энергий ионизации. С увеличением солености (NaCl) раствора образца до 2500 ppm отношение сигнал/шум сигнала LIBS улучшается в 0.85 раза. С увеличением атмосферного давления и удалением от объекта исследования интенсивность эмиссионной линии Ca II уменьшается. Противоположная тенденция наблюдается при увеличении плотности энергии излучения лазера и концентрации $\text{CaCl}_2 \cdot 2\text{H}_2\text{O}$ в растворе пробы. Период времени излучения Ca II 393 нм

оценивается как $(1/e)I_0$, I_0 — максимальная интенсивность. Типичный период излучения Ca II 393 нм 6.712—6.766 мкс. Указанное значение получено для плотности энергии лазерного излучения 9 Дж/см², атмосферного давления 1 атм, при расстоянии от образца 0.6 м, 2500 ppm NaCl и 1500 ppm для CaCl₂ · 2H₂O. Обнаружено, что период излучения увеличивается с повышением концентрации CaCl₂ · 2H₂O в растворе образца и мощности лазерного излучения. Противоположная тенденция имеет место при увеличении давления окружающей среды и расстояния до образца. Результаты, полученные из спектральных и временных измерений, коррелируют друг с другом. Лучшие показатели производительности модели системы получены в фиксированных экспериментальных условиях с использованием классификатора k-ближайших соседей.

Ключевые слова: лазерно-искровая эмиссионная спектроскопия, временные измерения, дистанционный анализ минералов, машинное обучение классификаторов.

Introduction. Mineral identification in a harsh environment (ocean beds [1], space exploration [2]) is a challenging and time-consuming process. These environments require a technique with remote sensing capability to estimate its mineral content. Also, the technique needs to produce reliable data under strident operating conditions, such as high-pressure (ocean mining [3], Venus exploration [4]) and low pressure (space exploration - Moon [5], Mercury [6], Mars [4]). One such technique to estimate the elemental composition under extreme operating conditions is laser-induced breakdown spectroscopy (LIBS). It also possesses advantages such as remote sensing capability, sample analysis irrespective of its state (solid, liquid, and gas), multi-element identification, minimal sample preparation, and high sensitivity.

In this paper, a remote mineral analysis under high-pressure condition (mimicking the ocean environment to a certain extent) is carried out using the LIBS technique combined with the temporal measurement. The influence of salinity is studied during the process of remote mineral analysis. Also, the machine learning (ML) algorithms are used to predict the unknown mineral concentration in bulk liquid for various operating conditions.

Numerous research works has been reported on determining the mineral composition of samples under high pressure condition using the LIBS method. Jinjia Guo et al. [7] utilized the LIBS signal to identify the elements like sodium, potassium, and calcium in seawater (collected at different depths ranging from 100 to 1700 m). The experiments were carried out by varying the ambient pressure up to 168 atm. The authors reported that the emission line peak intensity (Na I 330.2, 589, and 589.6 nm; K I 766.5 and 769.9 nm; Ca I 422.7 nm) decreased with increase in the ambient pressure. Pressure broadening on the emission lines was also reported [7].

Anna P. M. Michel et al. [8] demonstrated the feasibility of the single-pulse LIBS method to detect the elements (Na, Mg, and Ca) under high pressure condition (269 atm). The emission lines related to the above-specified elements were visualized in the LIBS spectrum. The phenomenon of pressure broadening was observed in the LIBS signal at high pressure. A spectrometer gate delay of less than 200 ns resulted in a high signal-to-noise ratio. Also, the authors reported that the usage of laser energy of less than 60 mJ resulted in minimizing the plasma shielding effect and the phenomenon of continuous breakdown. The detection limit in the order of 50, 500, and 50 ppm was achieved for the elements Na, Mn, and Ca, respectively [8].

C. Goueguel et al. [9] studied the influence of the ionic compound (NaCl, Na₂SO₄, Na₂CO₃) solution in relation to the LIBS signal (Li, K) under high pressure condition. The authors reported that the LIBS signal intensity (Li I at 670.80 nm; K I at 766.49 nm) increased with the addition of the NaCl and Na₂SO₄ compounds. The opposite trend was reported for the Na₂CO₃ compound [9]. Christian Goueguel et al. studied the influence of NaCl in relation to the LIBS signal of the Ca and K elements. Two parameters (signal to background ratio (SBR) and signal to noise ratio (SNR)) were analyzed. Two emission lines (Ca I at 422.67 nm; K I at 769.49 nm) were considered. The authors reported that an increase in the NaCl concentration resulted in an increased SNR value, whereas the SBR decreased [10]. The authors studied the influence of the ambient pressure (CO₂ at 345 atm) in the determination of the Ca²⁺ and Ba²⁺ emission lines. The LIBS signal lines (Ca²⁺, Ba²⁺) were clearly visualized at different operating pressure (49 to 345 atm) conditions [11]. Also, the authors studied the influence of high ambient pressure (CO₂ at 10 to 395 atm) in the determination of other elements, such as Mg, Sr, Mn, and Ba. It was reported that the ambient pressure condition had a minimal effect on the detection limit of Mg, Sr, and Ba. The detection limit of Mn decreased with increase in the ambient pressure condition [12].

Li et al. [13] studied the influence of salinity on the LIBS signal in bulk liquids. The salinity of the sample solution varied between 2 to 50%. The LIBS signal and imaging techniques were used to study the plas-

ma emission. It was reported that the atomic emission line intensity increased with increase in salinity, whereas the ionic emission line intensity was suppressed with increase in the salinity of the sample solution. It was also reported that the detection limit of elements did not change significantly with increase in the salinity of bulk liquid [13].

M. Dell' Aglio et al. [14] determined the mineral content in meteorite samples using the calibration-free LIBS technique. The measurements were carried out at a stand-off distance of 5 m. Elements such as Fe, Ni, and Co were identified in the Toluca meteorite. In the Sahara 98222 meteorite sample, the elements identified were Fe, Mg, Si, Na, Ti, Al, Cr, Mn, Ca, Ni, and Co. The calibration-free LIBS method was used to estimate the concentration of different elements in meteorite samples [14].

V. Sathiesh Kumar et al. [15] performed the remote identification and ranking of the salt (NaCl) deposit on the wind turbine blade material. The stand-off collection distance varied between 1 to 40 m. The quantification of the salt deposit was carried out by measuring the Na I (589 nm) emission intensity time period. The authors reported that the Na I emission intensity time period decreased with increase in the stand-off collection distance. This phenomenon is related to the amount of light coupled to the collection device (Telescopic type-Photometric device) [15].

Rajendhar Junjuri et al. [16] determined the elemental composition of explosives and nonexplosives using a compact Stand-off LIBS system combined with an artificial neural network (ANN). Five explosives and nineteen nonexplosive samples were analyzed from a distance of 6.5 m. Explosives and non-explosive materials were classified using the 2D scatter plot and principal component analysis (PCA). For the artificial neural network method, classification accuracies of 98% (explosives) and 94% (nonexplosives) were reported [16].

Tomoko Takahashi et al. [17] used the LIBS method to determine the elemental composition of deep-sea hydrothermal vents. The samples were analyzed at a pressure of 296 atm. The minerals identified were Cu, Pb, and Zn. The partial least square (PLS) regression method was applied to the LIBS data to determine the concentration of the above-specified elements. A relative concentration error of 25% was obtained from the PLS method when calibrated with the standard samples. The authors also reported that improvement in the concentration estimation is possible by collecting the plasma from multiple shots with a small positional offset [17].

Liwen Sheng et al. [18] performed the classification of iron ores from the LIBS signal using a random forest classifier. The classification accuracy was compared with the result obtained using the support vector machine (SVM) on a predetermined training dataset [18]. He Li'ao et al. [19] utilized the unsupervised learning algorithm such as hierarchical clustering analysis, *k*-means clustering, and the iterative self-organizing data analysis method to classify the plastics. The classification was related to the intensities of LIBS spectral emission lines (C 247.856 nm, CN 388.055 nm, C₂ 516.520 nm, C₂ 558.416 nm, H 656.285 nm, N 746.831 nm, O 777.194 nm) [19].

Based on the extensive literature survey, it has been identified that the number of reported works on the remote LIBS analysis of minerals in bulk liquid under high pressure condition is limited. It is also observed that most of the experimental analysis is based on spectral analysis. Temporal analysis (the plasma emission time period) is not considered as one of the influencing parameters. In quantitative analysis, the plasma emission time period plays a vital role. The emission time period is directly related to the concentration of the element present in the sample. The emission time period is influenced by the experimental condition and ambient environment. So, a detailed temporal analysis needs to be carried out. It helps in the estimation of the concentration of the elements present in the sample. The estimation of the sample in its natural environment requires a remote measurement technique. In the literature, the remote mineral analysis in bulk liquid is sparse. Since the experimental results (qualitative and quantitative measurements) are in close relation with the operating condition or parameters (stand-off distance, ambient pressure, laser fluence, LIBS signal), a mathematical model is required to predict the unknown mineral content in bulk liquid for different operating conditions.

The LIBS method combined with the temporal measurement is used to identify and estimate the calcium concentration in bulk liquid under high pressure conditions (1 to 8 atm). Also, the influence of salinity (NaCl) in the determination of the calcium concentration from the LIBS signal is studied. The stand-off collection distance varies between 0.6 to 2 m. Temporal analysis is carried out. A mathematical model (machine learning algorithms) is used to fit the LIBS data, thereby enabling prediction of the unknown calcium concentration in bulk liquid for different operating conditions. The machine learning algorithms considered in

the studies are the k -nearest neighbor (k -NN), Gaussian naive Bayes (GNB), classification and regression tree (CART), random forest classifier (RFC), logistic regression (LR), and linear discriminant analysis (LDA).

Experimental setup. The remote LIBS experimental setup to estimate the concentration of calcium in bulk liquid under high pressure condition is depicted in Fig. 1. A pulsed Nd:YAG laser with a 2ω module is used to create the plasma on the target surface. The optical emission from the plasma is collected using the combination of lens and optical fiber. The collected signal is fed to the spectrometer and photomultiplier tube. The spectral information is visualized using the spectrometer. The temporal characteristics of the plasma are obtained using the photomultiplier tube. A list of the equipment/components used in the experimental setup is detailed with their specifications in Table 1.

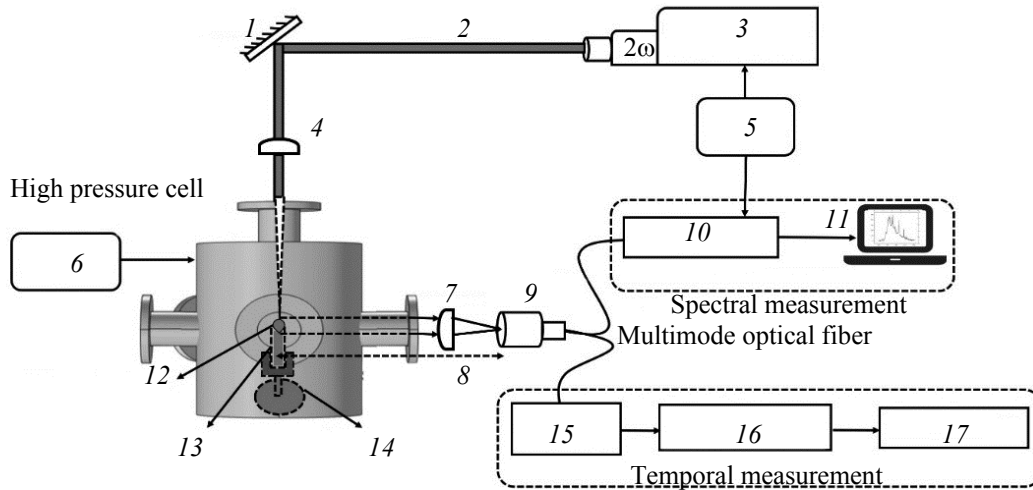


Fig. 1. Experimental setup: Remote LIBS technique combined with the temporal measurement: 1 – dichroic mirror, 2 – laser beam, 3 – Nd:YAG laser (1064 nm, 10 Hz, 6 ns), 4 – L_1 ($f = 300$ mm), 5 – delay generator, 6 – air compressor, 7 – L_2 ($f = 100/300$ mm), 8 – stand-off-distance, 9 – collimating lens, 10 – spectrometer, 11 – computer, 12 – plasma, 13 – quartz cuvette, 14 – cuvette holder, 15 – filter, 16 – photomultiplier tube, 17 – oscilloscope.

TABLE 1. List of Equipment and Components Used in Remote LIBS Experimental Setup

Equipment/Component	Model/Type	Specification
Laser	Quantel Q-smart 450	Fundamental wavelength: 1064 nm, Repetition rate: 10 Hz, Pulse duration: 6 ns
Focusing lens	Plano-convex lens	Focal length: 300 mm, Diameter: 50 mm
Collection lens	Plano-convex lens	Focal length: 300/100 mm, Diameter: 50 mm
Spectrometer	Ocean optics 2000+	Wavelength range: 200 to 1100 nm, Resolution: 1 nm
Fiber optic cable	BFH37-600	Core diameter: 600 μ m, Numerical aperture: 0.39, Spectral range: 300 to 2200 nm
Photomultiplier tube	R1925A	Response time: 1.5 ns (rise time), 17 ns (transit time), Spectral range: 300 to 850 nm, Gain: 5.0×10^5
Delay generator	DG645	Resolution: 5 ps, No. of channels: 4
High-pressure chamber	Own design & fabrication	Max. applied pressure: 8 atm

The sample solution is prepared by dissolving NaCl (2500 ppm) and $\text{CaCl}_2 \cdot 2\text{H}_2\text{O}$ (500–1500 ppm) in deionized water. The sample solution in a two-sided quartz cuvette is placed at the center of the high-pressure chamber. The high-pressure chamber is made up of stainless steel. The chamber encompasses six quartz optical windows. The optical windows are used to transmit/receive the optical signal to/from the chamber. The pressure inside the chamber is created using an air compressor (up to 10 atm).

A laser pulse of 532 nm wavelength is directed and focused on to the top surface of the sample solution in a quartz cuvette. The selection of the incident laser wavelength (532 nm) to create ablation is related to the sample absorption characteristics (high) and liquid splashing property (low). These properties play a vital role in the collection of plasma from different stand-off distances.

The plasma is generated on the top surface of the sample solution. It expands in the air. The expansion rate of the plasma is inversely proportional to the ambient pressure condition. The optical emission from the plasma is collected from one of the side quartz window in the high-pressure chamber. The stand-off collection distance is varied by changing the position of the collecting lens concerning the optical window. The distance between the plasma and the entrance point of the optical fiber is considered as the stand-off collection distance. It varies between 0.6 to 2 m. To minimize continuum radiation in the LIBS spectra, the spectrometer is triggered using a delay generator after 600 ns from the laser pulse.

The process of identifying the unknown elemental (Ca) concentration in the sample solution is shown in Fig. 2. The feature vector is formed by concatenating different measurements/parameters such as the incident laser wavelength (532 nm), laser fluence (9 to 11 J/cm²), ambient pressure (1 to 8 atm), relative elemental peak intensity (Ca II 396 nm/Ca II 393 nm), elemental line wavelengths (393, 396 nm), temporal elemental peak intensity (393 nm), elemental emission time period, and stand-off distances (0.6 to 2 m).

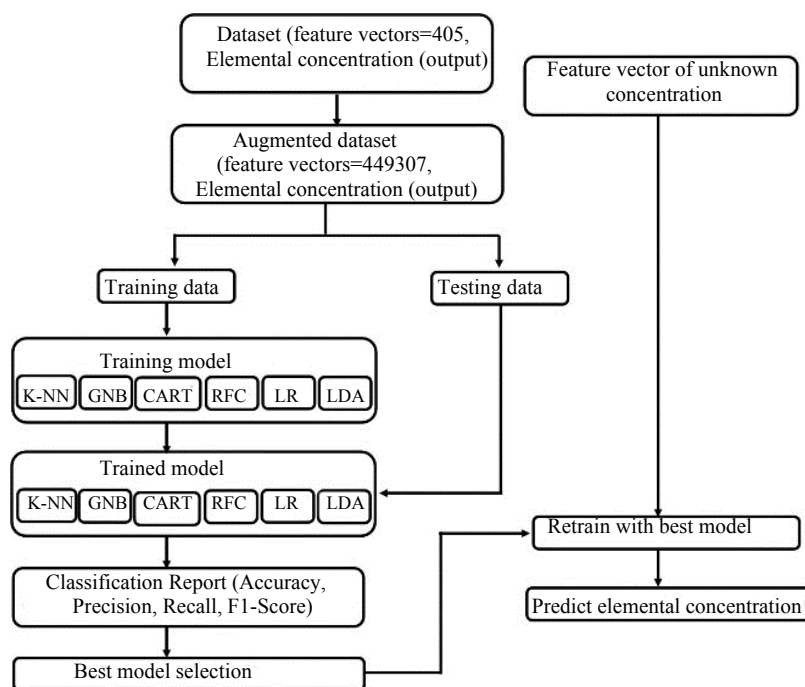


Fig. 2. Steps involved in identifying the unknown elemental concentration in the sample solution.

The number of experimental data collected consist of 405 feature vectors. The data augmentation is performed to increase the number of feature vectors to 449,307. The data augmentation process results in the improvement of accuracy in the prediction. The augmented dataset is split into training and testing datasets. The training/testing data split ratio (0.7/0.3 or 0.75/0.25 or 0.8/0.2 or 0.85/0.15 or 0.9/0.1) is varied.

A model is created by exposing the training data. The machine learning algorithms used in the creation of the model are *k*-NN, GNB, CART, RFC, LR, and LDA. Once the model is trained, it is exposed to a test dataset resulting in the generation of a classification report. The classification report provides details about the performance metrics of the model in the form of accuracy, precision, recall, and F1-score. The performance metrics are calculated using the equations given below [20]:

$$\text{Accuracy} = (\text{TP} + \text{TN}) / (\text{TP} + \text{TN} + \text{FP} + \text{FN}), \quad (1)$$

$$\text{Precision} = \text{TP} / (\text{TP} + \text{FP}), \quad (2)$$

$$\text{Recall} = \text{TP} / (\text{TP} + \text{FN}), \quad (3)$$

$$\text{F1-Score} = 2(\text{Precision} \times \text{Recall}) / (\text{Precision} + \text{Recall}), \quad (4)$$

where TP is the number of true positives, TN is the number of true negatives, FP is the number of false positives, and FN is the number of false negatives. The model with higher performance metrics is used in retraining the system before exposing it to a feature vector obtained from the unknown elemental (Ca) concentration in the sample solution. This results in the prediction of the elemental concentration.

k -NN classification is performed based on the distance metrics (Euclidean distance or chi-square distance or Manhattan distance). In the experimental analysis, the Euclidean distance metric is used. For $k \geq 2$, the majority voting scheme is utilized to determine the final output prediction [20].

GNB work on the principle of the Bayes theorem as represented in the following equation [21]:

$$p(x = v | C_k) = \frac{1}{\sqrt{2\pi}\sigma_k} \exp\left(-\frac{(v - \mu_k)^2}{2\sigma_k^2}\right), \quad (5)$$

where x is the continuous attribute in the training input data, μ_k is the mean, σ_k^2 is the variance, v is the observation value, and C_k is the number of classes. If the input data or features are continuous, then the values associated with each class are assumed to be distributed supporting Gaussian functionality. This algorithm is widely used in the binary classification or multi-class classification problem.

In the CART algorithm, the target variable is predicted based on other parameters (condition checking). It forms a binary tree. The root node of the tree represents a single input variable (x). Checking for a condition at the root node results in the formation of leaf nodes. Further, the leaf nodes are checked for a certain condition to induce further branching of the tree. This process is continued until the prediction (output variable (y)) of the input variable is obtained.

RFC is formed by the ensemble of CART classifiers. It is also a supervised machine learning algorithm. The final output prediction on the input data is dependent on the majority voting by the CART classifiers.

LR is a supervised learning algorithm. It is mainly used in a situation where the dependent target variables are categorical in nature. The decision boundary (non-linear) is set with a threshold to determine or estimate the classification group for the input data. The logistic regression of the independent input variable ($x_1, x_2, \dots, x_n$) is represented as [20]

$$P(Y = 1) = [1 + \exp(-(\beta_0 + \beta_1 x_1 + \dots + \beta_n x_n))]^{-1}, \quad (6)$$

where $P(Y = 1)$ is the probability of the true state, and $\beta_0, \beta_1, \dots, \beta_n$ are the regression coefficient.

The LDA technique is used to select the important combination of linear features related to the classification problem. It attempts to model the difference between the classes of input data. The model estimates the between-class and within-class variances from the mean and standard deviation of the input data in the dataset.

Results and discussion. Spectral measurement. Initially, the sample solution is prepared by dissolving $\text{CaCl}_2 \cdot 2\text{H}_2\text{O}$ (500 to 1500 ppm) in deionized water. The analysis is carried out by varying the incident laser fluence (9 to 11 J/cm²), ambient pressure (1 to 8 atm), and solution concentration (500 to 1500 ppm). Figure 3 represents the LIBS spectral information obtained using the sample solution ($\text{CaCl}_2 \cdot 2\text{H}_2\text{O}$) at a stand-off distance of 0.6 m. The complete spectrum is shown in Fig. 3a. From the spectral data, the emission lines related to Ca II at 393 and 396 nm, H I at 656 nm, and O I at 777 nm are visualized. The emission lines related to doubly ionized atoms and chlorine atoms are not observed due to the requirement of high ionization energies [22]. Figure 3b shows the influence of the laser fluence on the Ca II emission lines (393 and 396 nm). It is observed that the peak intensity of the Ca emission line increased with an increase in the laser fluence. A larger volume of the material is ablated with increase in the laser fluence. It increases the plasma components (excited state electrons, ions, and neutral atoms) and the plasma optical emission. Figure 3c shows the influence of the ambient pressure on the Ca II emission lines (393 and 396 nm). It is observed that the peak intensity of the Ca emission line decreased with increase in the ambient pressure. Once the plasma is created, it starts to expand at supersonic velocities. This expansion of the plasma is mitigated with an increase in the ambient pressure. This results in a smaller area of plasma emission at a higher pressure. Hence, the peak Ca emission line intensity is reduced with increase in the ambient pressure condition. A similar trend in the peak intensity of the elements under high pressure conditions is reported by Huaming Hou et al. [23]. Figure 3d shows the influence of the $\text{CaCl}_2 \cdot 2\text{H}_2\text{O}$ concentration on the emission lines Ca II 393 and 396 nm.

The spectrum is recorded for a laser fluence of 9 J/cm^2 , ambient pressure of 1 atm, and SOD of 0.6 m. It is observed that the Ca II emission peak intensity increased with increase in the $\text{CaCl}_2 \cdot 2\text{H}_2\text{O}$ concentration. As the concentration of the ionic compound in the sample solution increases, the number of interactions between the calcium atoms and incident photons also increases, which in turn increases the Ca II emission line peak intensity.

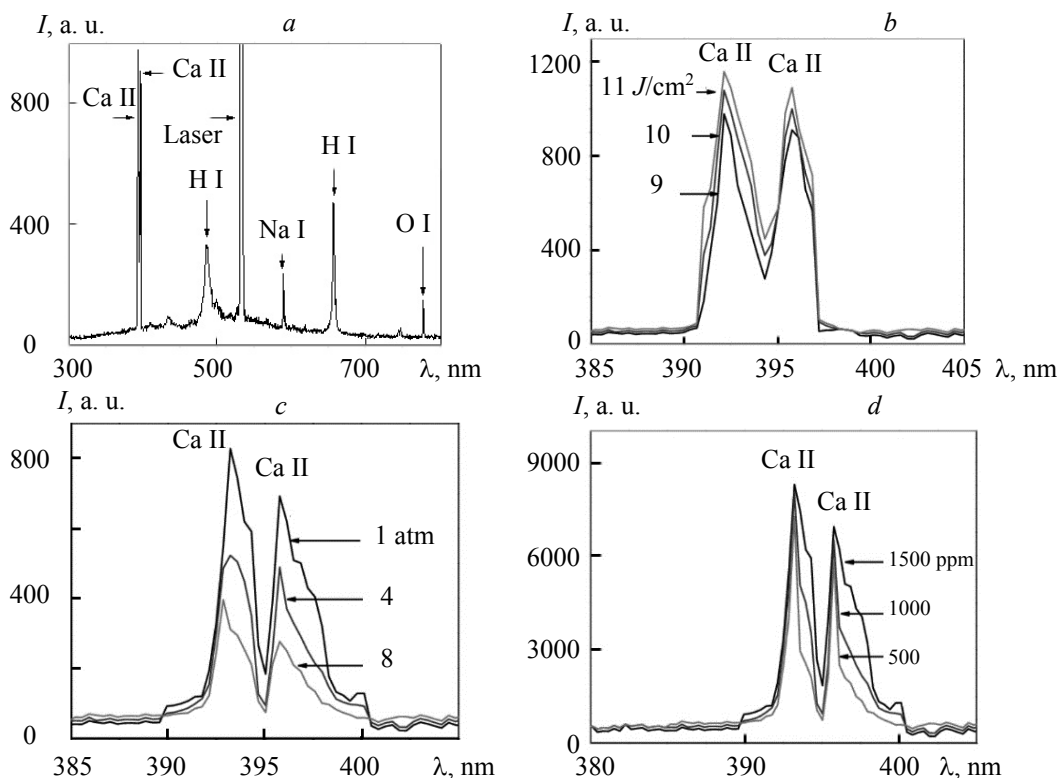


Fig. 3. LIBS spectra of the sample solution ($\text{CaCl}_2 \cdot 2\text{H}_2\text{O}$ + deionized water): a) spectrum; b) varying laser fluence; c) varying pressure condition; d) varying concentration.

Figure 4 shows the LIBS spectra of the sample solution (deionized water+2500 ppm NaCl+ $\text{CaCl}_2 \cdot 2\text{H}_2\text{O}$). The complete spectrum is shown in Fig. 4a. The spectrum is recorded for a laser fluence of 9 J/cm^2 , ambient pressure of 1 atm, stand-off distance of 0.6 m, and $\text{CaCl}_2 \cdot 2\text{H}_2\text{O}$ concentration of 1500 ppm. From the LIBS spectra, the singly ionized lines Ca II 393 and 396 nm are visualized. Also, the emission lines H I at 656 nm, Na I at 589 nm, and O I at 777 nm are observed. The emission lines related to doubly ionized atoms and chlorine atoms are not observed due to the requirement of high ionization energies [22]. Figure 4b represents the influence of the laser fluence in the determination of Ca II emission lines. The spectrum is recorded for an ambient pressure of 1 atm, stand-off collection distance of 0.6 m, and $\text{CaCl}_2 \cdot 2\text{H}_2\text{O}$ concentration of 1500 ppm in the sample solution. It is observed that the emission peak intensity at Ca II is increased with increase in the laser fluence. Figure 4c represents the influence of the ambient pressure condition in the determination of Ca II emission lines. The spectrum is obtained for a laser fluence of 9 J/cm^2 , stand-off collection distance of 0.6 m, and $\text{CaCl}_2 \cdot 2\text{H}_2\text{O}$ concentration of 1500 ppm in the sample solution. It is observed that the Ca II emission line peak intensity decreases with increase in the ambient pressure condition. The trend observed is in correlation with the LIBS measurement of the sample solution without NaCl. Figure 4d shows a comparison of the LIBS spectra obtained using the sample solution with and without 2500 ppm NaCl. The spectrum is recorded for a laser fluence of 9 J/cm^2 , ambient pressure of 1 atm, stand-off distance of 0.6 m, and 1500 ppm of $\text{CaCl}_2 \cdot 2\text{H}_2\text{O}$. It is observed that the Ca II emission line (393 and 396 nm) intensities increase with the addition of NaCl (2500 ppm) to the sample solution. The improvement of LIBS SNR is of the order of 0.85. Sodium atoms have lower ionization energy (5.14 eV) compared to that of Ca atoms (6.11 eV) [22]. This results in the generation of a large number of free electrons and Na ions. These free electrons may interact with the calcium neutral atoms in the plasma to release an electron and form singly ionized Ca atoms

(Ca II at 393 and 396 nm), which leads to increase in the Ca II emission peak intensity with the addition of NaCl to the sample solution. A similar trend is reported by Anna P.M. Michel et al. [8]: the addition of an easily ionizable element to the sample solution would enhance the SNR of other elements.

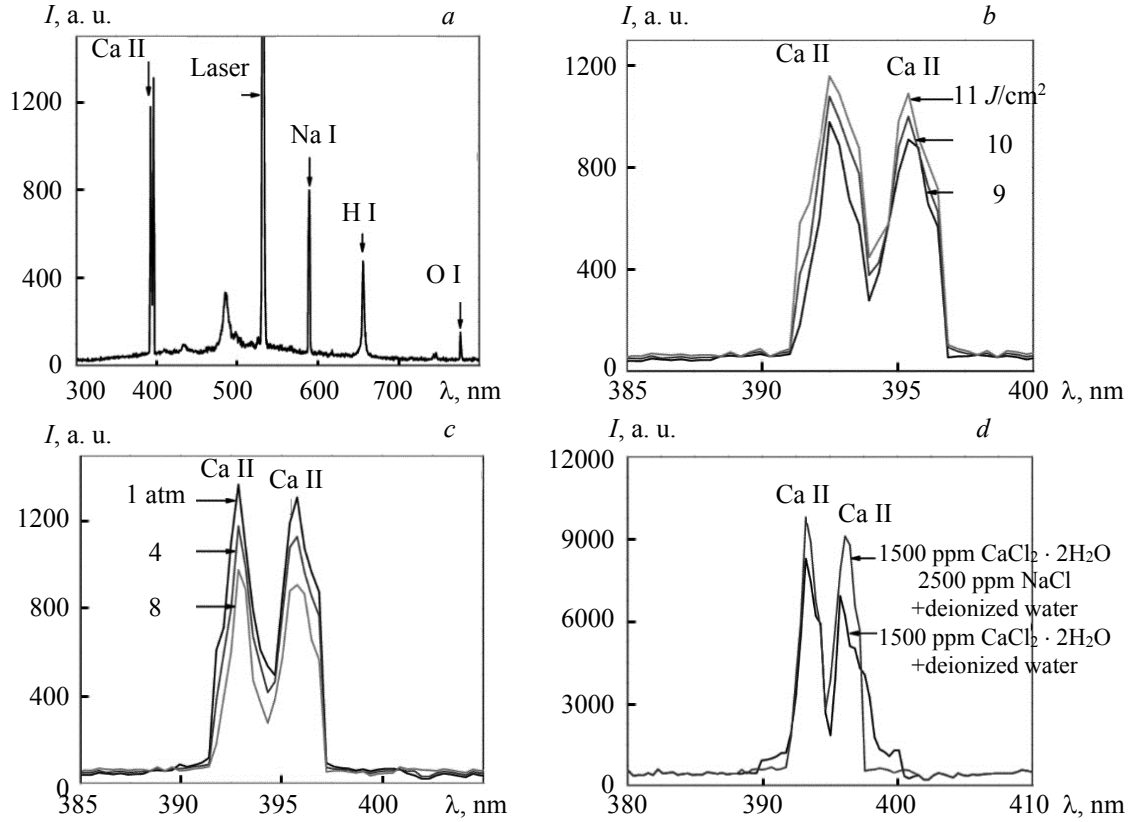


Fig. 4. LIBS spectra of the sample solution ($\text{CaCl}_2 \cdot 2\text{H}_2\text{O} + 2500 \text{ ppm NaCl} + \text{deionized water}$): a) spectrum; b) varying laser fluence; c) varying ambient pressure condition; d) with and without addition of NaCl.

Table 2 shows the influence of the stand-off collection distance in obtaining the LIBS information on the Ca II emission line. The table represents the variation of Ca II (393 nm) emission peak intensity in relation to the concentration of $\text{CaCl}_2 \cdot 2\text{H}_2\text{O}$ in the sample solution ($\text{CaCl}_2 \cdot 2\text{H}_2\text{O} + 2500 \text{ ppm NaCl} + \text{deionized water}$), ambient pressure (1 to 8 atm), and stand-off collection distance (0.6 to 2 m). The measurements are carried out by fixing an incident laser fluence of 9 J/cm^2 . From the spectral data, it is observed that the Ca II (393 nm) emission line peak intensity increased with increase in the $\text{CaCl}_2 \cdot 2\text{H}_2\text{O}$ concentration in the sample solution. This might be due to the increase in the number of interactions between the incident photons and Ca neutral atoms for the sample solution with a higher Ca concentration. It is also observed that the peak intensity of Ca II (393 nm) decreases with increase in the stand-off distance and the ambient pressure. As the stand-off collection distance increases, the amount of light (optical emission from the plasma) coupled to the optical fiber decreases due to the expansion of the plasma. Similarly, increase in the ambient pressure does not allow the plasma to expand freely. This results in a decrease in the emission line intensity.

Figure 5a shows the influence of the stand-off distance and the ambient pressure condition in the determination of the Ca II 393 nm emission peak intensity. It is observed that the peak intensity decreases with increase in the ambient pressure condition and the stand-off distance. The relationship between the stand-off distance and the Ca II 393 nm peak intensity for a varying pressure condition is estimated using the equation

$$I_{\text{Ca } 393} = S_{\text{Ca } 393}(f)/r^2, \quad (7)$$

where $I_{\text{Ca } 393}$ is the measured Ca II 393 nm peak intensity, r is the distance between the plasma creation point to the frontal apex of the optical fiber, and $S_{\text{Ca } 393}(f)$ is the Ca II 393 nm peak intensity at the plasma creation point. The peak intensity at the plasma creation point depends on the incident laser fluence, ambient pressure

condition, and concentration of the ionic compound in bulk liquid. For a fixed ambient pressure (1 atm) and $\text{CaCl}_2 \cdot 2\text{H}_2\text{O}$ concentration in bulk liquid (1500 ppm), the Ca II emission peak intensity is given by

$$S_{\text{Ca 393}}(f) = a_i \exp(f/t_i), \quad (8)$$

where a_i and t_i is the amplitude and decaying constant, respectively; f is the incident laser fluence. The estimated exponential fit parameters are listed in Table 3.

TABLE 2. Ca II 393 nm Emission Peak Intensity at Different Stand-off Collection Distances for Varying Ambient Pressure and $\text{CaCl}_2 \cdot 2\text{H}_2\text{O}$ Concentration in Bulk Liquid

Pressure, atm	Concentration, $\text{CaCl}_2 \cdot 2\text{H}_2\text{O}$, ppm	Emission line peak intensity Ca II 393 nm (a. u.)							
		SOD 0.6 m		SOD 1 m		SOD 1.5 m		SOD 2 m	
		Peak Intensity	Tolerance $\pm$	Peak Intensity	Tolerance $\pm$	Peak Intensity	Tolerance $\pm$	Peak Intensity	Tolerance $\pm$
1	500	7268.985	224.615	6444.985	243.435	5328.985	256.654	4068.985	267.615
	1000	8013.985	243.871	6517.985	265.543	5531.985	244.761	4861.985	286.453
	1500	8231.985	267.615	6731.985	243.564	5931.985	221.543	5131.985	269.324
4	500	6934.234	321.097	6034.543	301.653	5098.567	243.059	3787.432	299.504
	1000	7865.435	245.643	6213.856	256.987	5398.934	298.079	4512.654	319.048
	1500	7876.985	267.615	6731.985	244.543	5631.985	239.654	4993.267	309.234
8	500	6521.765	223.976	5823.948	287.654	4862.561	276.93	3426.749	245.894
	1000	7431.905	267.651	6354.076	239.065	5009.064	246.713	4065.073	245.954
	1500	7631.985	297.615	6531.985	255.345	5231.985	198.654	4623.789	276.543

TABLE 3. Estimated Parameters from the Exponential Fit in Fig 5a

Fluence, J/cm^2	1 atm		4 atm		8 atm	
	a_i	t_i	a_i	t_i	a_i	t_i
9	9784.21	3.03	9537.44	2.97	9262.02	2.72
10	10075.00	3.18	9787.11	3.11	9427.58	2.81
11	10529.90	3.43	10128.75	3.29	9672.11	2.94

Figure 5b shows the influence of the stand-off distance and the laser fluence in the determination of the Ca II 393 nm emission peak intensity. It is observed that the peak intensity decreases with increase in the stand-off distance, but the peak intensity increases with increase in laser fluence.

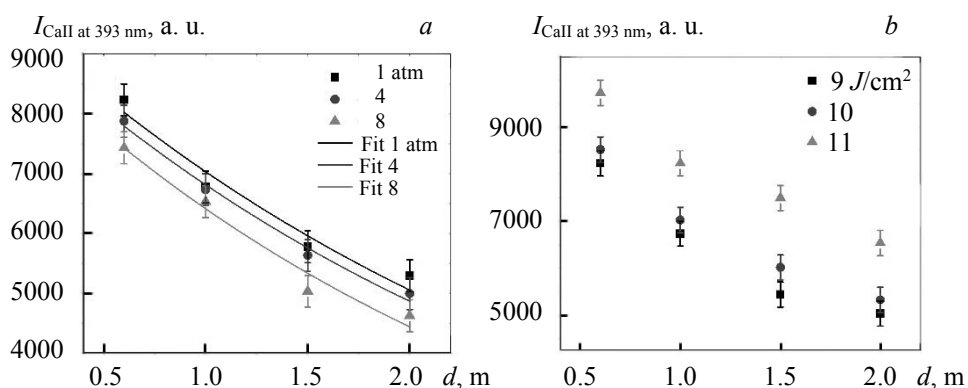


Fig. 5. LIBS spectral measurement at different stand-off collection distances: a) varying ambient pressure condition; b) varying laser fluence.

Temporal measurement. Figure 6a shows the temporal profile of the Ca II 393 nm emission line for varying ambient pressure conditions. The temporal profile is recorded for a 1500 ppm solution concentration of $\text{CaCl}_2 \cdot 2\text{H}_2\text{O}$ in deionized water. The profile is obtained by fixing the laser fluence at 9 J/cm^2 and the stand-off distance at 0.6 m. As the plasma is created, the Ca II 393 nm emission intensity increases rapidly to

a peak value and then starts to decay exponentially. The emission time period is measured at $(1/e)I_0$, where I_0 is the peak emission intensity. From Fig. 6a, it is observed that the emission time period of the Ca II 393 nm emission line decreases with increase in the ambient pressure. This might be due to the fact that at high pressure the plasma is not allowed to expand freely due to which the area of the plasma emission is greatly reduced. This results in the reduction of the light collected by the measurement system and leads to lowering of the Ca II 393 nm emission intensity and time period for the high ambient pressure. A similar trend in the temporal measurement is reported by H. Hou et al. [23] and Sathiesh Kumar et al. [15].

Figure 6b shows the variation of the Ca II 393 nm emission time period in relation to the solution concentration (500 to 1500 ppm) and the ambient pressure. It is observed that the emission time period increases with increase in the solution concentration. It decreases with increase in the ambient pressure. As the solution concentration increases, the number of ions in the solution increases too. This in turn increases the number of excited free electrons in the plasma state. It aids in the re-excitation of the Ca atoms in the solution. A Ca II 393 nm emission time period in the range from 6.712 to 6.766 μ s is obtained for a laser fluence of 9 J/cm², ambient pressure condition of 1 atm, stand-off distance of 0.6 m, and CaCl₂ · 2H₂O concentration of 1500 ppm. The variation in the Ca II 393 nm emission time period with the CaCl₂ · 2H₂O concentration and laser fluence is represented in Fig. 6c. From the figure, it is observed that the emission time period increases with increase in the laser fluence. As the laser fluence increases, the amount of the ablated material volume also increases. This results in an increase in the emission time period.

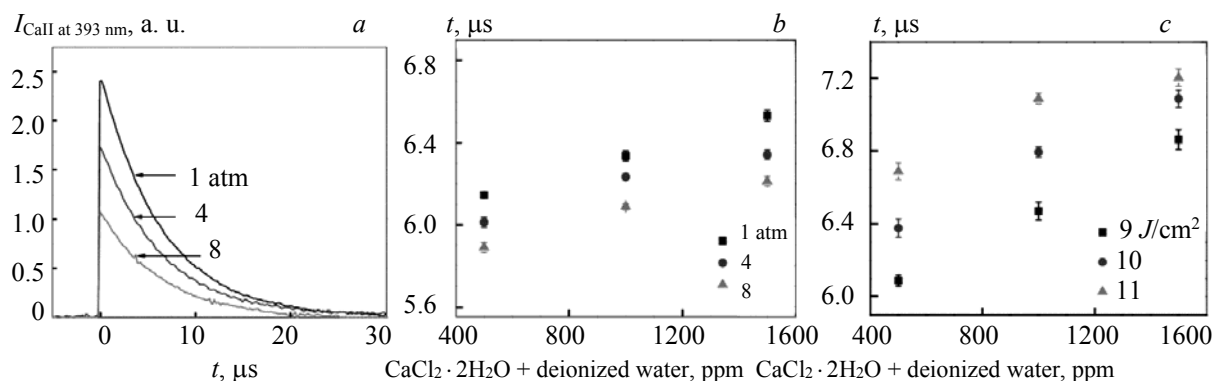


Fig. 6. Temporal measurement of the plasma emission in the sample solution (CaCl₂ · 2H₂O + deionized water): a) Ca II 393 nm emission profile for varying ambient pressure condition; b) Ca II 393 nm emission time period for varying concentrations and ambient pressure condition; c) Ca II 393 nm emission time period for varying concentrations and laser fluence.

Figure 7a shows the temporal profile of the Ca II 393 nm emission line obtained using the sample solutions (CaCl₂ · 2H₂O + deionized water, CaCl₂ · 2H₂O + 500 ppm NaCl + deionized water). The plot is obtained for a solution concentration of 1500 ppm CaCl₂ · 2H₂O. The experimental parameters such as the laser fluence and the ambient pressure are fixed at 9 J/cm² and 1 atm, respectively. It is observed that the emission time period is higher for the sample solution with NaCl. This might be due to the fact that the ionization energy of sodium is lower compared to that of the calcium ion. Due to the lower ionization energy, more sodium ions are generated. The electron-atom and ion-atom collision results in the generation of singly ionized Ca atoms and leads to the increase in Ca II 393 nm emission intensity and time period.

Figure 7b shows the relationship between the Ca II 393 nm emission time period and the CaCl₂ · 2H₂O concentration and the ambient pressure. The temporal measurements are carried out for a laser fluence of 9 J/cm² and, a stand-off distance of 0.6 m. The salinity (NaCl) of the sample solution is maintained at 2500 ppm. From the plot, it is observed that the Ca II 393 nm emission time period increases with increase in the CaCl₂ · 2H₂O concentration. This increase in the emission time period is related to the increase in the number of photon interactions with neutral Ca atoms for the sample solution with a higher CaCl₂ · 2H₂O concentration. It is also observed that the emission time period decreases with increase in the ambient pressure. This is related to the mitigation of plasma expansion at a high ambient pressure. A Ca II 393 nm emission time period in the range from 6.906 to 6.96 μ s is obtained for a laser fluence of 9 J/cm², ambient pressure condition of 1 atm, stand-off distance of 0.6 m, salinity of 2500 ppm NaCl, and CaCl₂ · 2H₂O concentration of 1500 ppm.

Figure 7c shows the relationship between the Ca II 393 nm emission time period and the $\text{CaCl}_2 \cdot 2\text{H}_2\text{O}$ concentration and the laser fluence. The temporal measurements are carried out for an ambient pressure condition of 1 atm and a stand-off distance of 0.6 m. The salinity (NaCl) of the sample solution is maintained at 2500 ppm. From the plot, it is observed that the emission time period increases with increase in the laser fluence. This is related to the increase in the amount of the material ablated for a larger incident laser fluence.

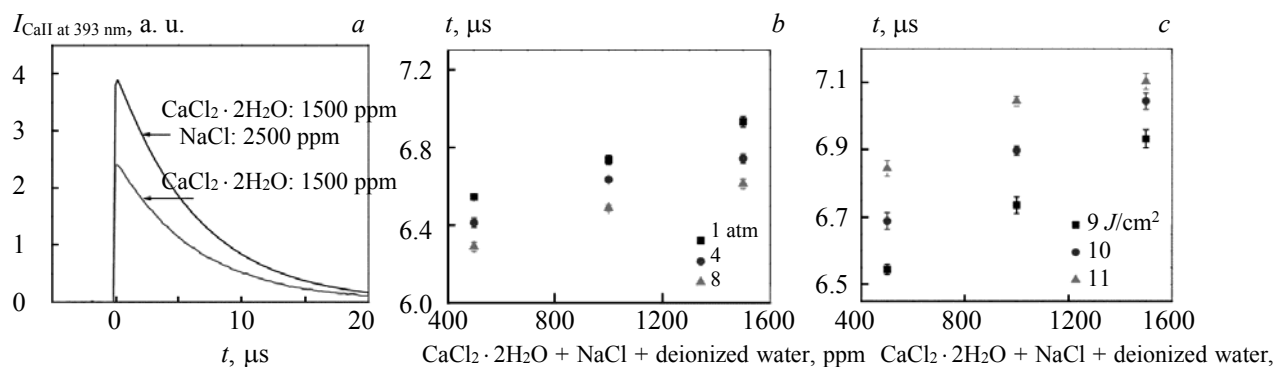


Fig. 7. Temporal measurement of the sample solution ($\text{CaCl}_2 \cdot 2\text{H}_2\text{O} + \text{NaCl} + \text{deionized water}$): a) Ca II 393 nm emission profile for the concentration of $\text{CaCl}_2 \cdot 2\text{H}_2\text{O} + \text{deionized water}$ and $\text{CaCl}_2 \cdot 2\text{H}_2\text{O} + \text{NaCl} + \text{deionized water}$; b) varying concentrations at different ambient pressures; c) varying concentrations at different laser fluence.

Table 4 presents the relationship between the Ca II 393 nm emission time period, $\text{CaCl}_2 \cdot 2\text{H}_2\text{O}$ concentration, ambient pressure, and stand-off collection distance. The data are obtained by fixing the laser fluence at 9 J/cm² and the NaCl concentration at 2500 ppm. From the table data, it is observed that the Ca II emission time period increases with increase in the concentration for a fixed stand-off distance. The emission time period decreases with increase in the stand-off collection distance. This is related to the reduction in the amount of light coupled to the measurement system for a larger stand-off distance.

TABLE 4. Ca II 393 nm Emission Time Period at Different Stand-off Collection Distances for Varying Ambient Pressure Condition and $\text{CaCl}_2 \cdot 2\text{H}_2\text{O}$ Concentration in Bulk Liquid

Pressure, atm	Concentration of $\text{CaCl}_2 \cdot 2\text{H}_2\text{O}$, ppm	Emission time period Ca II 393 nm (μs)							
		SOD 0.6 m		SOD 1 m		SOD 1.5 m		SOD 2 m	
		Time period	Tolerance $\pm$	Time period	Tolerance $\pm$	Time period	Tolerance $\pm$	Time period	Tolerance $\pm$
1	500	6.545	0.015	5.80219	0.0243	5.42019	0.0232	4.75019	0.03887
	1000	6.736	0.0241	5.96923	0.0139	5.50723	0.0145	4.91723	0.02787
	1500	6.933	0.027	6.50428	0.0238	5.94228	0.0239	5.45228	0.03887
4	500	5.9321	0.3887	4.7856	0.4312	3.9678	0.3241	3.0321	0.3452
	1000	6.5296	0.2331	5.2296	0.2331	4.3706	0.2331	3.5706	0.2331
	1500	6.8576	0.2341	5.8673	0.3887	5.0231	0.2562	4.1431	0.3887
8	500	5.4655	0.3887	4.2112	0.3211	3.4534	0.3523	2.7867	0.3732
	1000	6.3129	0.1555	4.8629	0.1555	3.6649	0.1555	3.0649	0.1555
	1500	6.4321	0.432	5.2341	0.2543	4.7654	0.2343	3.7121	0.1543

Figure 8a shows the relationship between the Ca II 393 nm emission time period, stand-off distance, and ambient pressure. The measurement is recorded for a $\text{CaCl}_2 \cdot 2\text{H}_2\text{O}$ concentration of 1500 ppm. The laser fluence is fixed at 9 J/cm². The salinity of the sample solution is maintained at 2500 ppm. As the ambient pressure increases, the plasma emission time period decreases. The increase in the ambient pressure quenches the plasma. This results in the reduction of the emission time period for the high-pressure condition. The relationship between the stand-off distance and the Ca II 393 nm emission time period for varying pressure condition is estimated using the equation

$$T_{\text{Ca } 393} = Z_{\text{Ca } 393}(f)/r^2, \quad (9)$$

where $T_{\text{Ca } 393}$ is the measured Ca II 393 nm emission time period, r is the distance between the plasma creation point to the frontal apex of the optical fiber, and $Z_{\text{Ca } 393}(f)$ is the Ca II 393 nm emission time period at the plasma creation point.

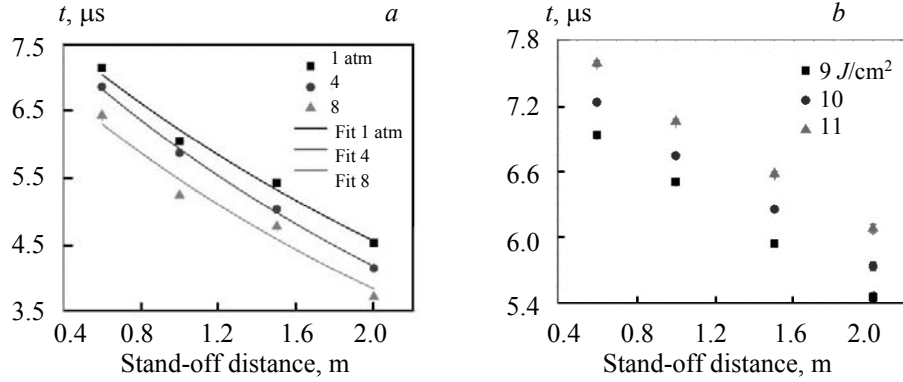


Fig. 8. Ca II 393 nm emission time period at different stand-off collection distances: a) varying pressure condition; b) varying laser fluence.

The emission time period at the plasma creation point depends on the incident laser energy density, ambient pressure condition, and concentration of the ionic compound in bulk liquid. For a fixed ambient pressure condition (1 atm) and $\text{CaCl}_2 \cdot 2\text{H}_2\text{O}$ concentration in bulk liquid (1500 ppm), the Ca II emission time period is given by the equation

$$Z_{\text{Ca } 393}(f) = a_i \exp(f/t_i), \quad (10)$$

where a_i and t_i are the amplitude and decaying constant, respectively, and f is the incident laser energy density. The estimated exponential fit parameters are listed in Table 5.

Figure 8b shows the relationship between the Ca II 393 nm emission time period, stand-off distance, and laser fluence. The measurements are carried out for a $\text{CaCl}_2 \cdot 2\text{H}_2\text{O}$ concentration of 1500 ppm and ambient pressure of 1 atm. The salinity of the sample solution is fixed at 2500 ppm. It is observed that the emission time period increases with increase in the laser fluence.

TABLE 5. Estimated Parameters from the Exponential Fit in Fig. 8a

Fluence, J/cm ²	1 atm		4 atm		8 atm	
	a_i	t_i	a_i	t_i	a_i	t_i
9	7.70	5.80	7.44	5.81	7.29	5.47
10	8.05	6.13	7.81	5.90	7.58	5.69
11	8.44	6.49	8.18	6.26	7.76	5.47

Machine learning models. The information from the spectral measurement, temporal measurement, and experimental parameters is used to form the feature vectors. The augmented data set of feature vectors is used in the training-cum-testing phase of the system model. The performance metrics are obtained at the end of the testing phase.

The classification report (accuracy, precision, recall, F1-score) obtained by implementing the machine learning classifiers is listed in Table 6. The performance metrics are also reported for varying the split ratio of the training and test data. From the table, it is observed that the k -NN classifier (accuracy = precision = recall = F1-score = 100%) performs better compared to other classifiers. The 100% accuracy in the prediction of the concentration by the k -NN classifier is achieved at certain experimental conditions such as laser fluence (9, 10, and 11 J/cm²), incident laser wavelength = 532 nm, ambient pressure condition (1, 4, and 8 atm), stand-off distance (0.6, 1.0, 1.5, and 2.0 m), and $\text{CaCl}_2 \cdot 2\text{H}_2\text{O}$ concentration (500, 1000, and 1500 ppm) in bulk liquid. Still, in the real-time scenario the operating condition may vary significantly.

In such cases, the accuracy in the prediction of the $\text{CaCl}_2 \cdot 2\text{H}_2\text{O}$ concentration in bulk liquid will drop. Hence, calibrating the instrument for different operating conditions is vital to utilize this methodology in real-time scenarios.

TABLE 6. Performance Metrics Obtained by Using Different Machine Learning Classifiers

Classifiers	Test Size, %	Accuracy, %	Precision	Recall	F1-Score
KNN	10	100.00	1	1	1
	15	100.00	1	1	1
	20	100.00	1	1	1
	25	100.00	1	1	1
	30	100.00	1	1	1
GNB	10	52.54	0.52	0.53	0.52
	15	52.93	0.52	0.53	0.52
	20	52.84	0.52	0.53	0.52
	25	52.73	0.52	0.53	0.52
	30	52.72	0.52	0.53	0.52
CART	10	99.93	1	1	1
	15	99.93	1	1	1
	20	99.94	1	1	1
	25	99.95	1	1	1
	30	99.94	1	1	1
RFC	10	99.98	1	1	1
	15	99.98	1	1	1
	20	99.98	1	1	1
	25	99.98	1	1	1
	30	99.98	1	1	1
LR	10	86.62	0.87	0.87	0.87
	15	86.71	0.87	0.87	0.87
	20	86.6	0.87	0.87	0.87
	25	86.57	0.87	0.87	0.87
	30	86.6	0.87	0.87	0.87
LDA	10	84.51	0.85	0.85	0.85
	15	84.58	0.85	0.85	0.85
	20	84.47	0.85	0.84	0.85
	25	84.43	0.85	0.84	0.85
	30	84.53	0.85	0.85	0.85

In summary, the remote LIBS technique combined with the temporal measurement and machine learning classifiers can be used to estimate the unknown calcium concentration in bulk liquid under high pressure condition. The spectral information and temporal data are highly dependent on the experimental parameters or conditions (laser wavelength, laser fluence, ambient pressure condition, salinity of the sample solution, elemental concentration in bulk liquid, stand-off collection distance). Hence, calibration of data (spectral and temporal measurements) for different operating condition becomes necessary to utilize this method to identify the unknown elemental concentration in bulk liquid under high pressure condition.

Conclusions. An investigation of the possibility of effective identification and estimation of the Ca concentration in bulk liquid under high pressure condition is carried out. Experimental analysis is performed by combining the LIBS method with the temporal measurement and machine learning classifiers. Also, the influence of salinity in relation to the LIBS signal-to-noise ratio is studied. From the experimental investigation, it is observed that the spectral and temporal information is highly dependent on the operating parameters or conditions such as ambient pressure, laser fluence, elemental concentration in the sample solution, salinity of the sample, and stand-off collection distance. Hence, calibration of the spectral and temporal information for different operating conditions becomes necessary to determine the unknown elemental concentration in the sample solution.

Acknowledgments. This work is supported by a grant-in-aid from the Science and Engineering Research Board-File no: ECR/2015/000028 dated 10.08.2016.

REFERENCES

1. B. Thornton, T. Takahashi, T. Sato, T. Sakka, A. Tamura, A. Matsumoto, et al., *Deep Sea Res. I: Oceanogr. Res. Papers*, **95**, 20–36 (2015).
2. A. K. Knight, N. L. Scherbarth, D. A. Cremers, M. J. Ferris, *Appl. Spectrosc.*, **54**, No. 3, 331–340 (2000).
3. K. A. Miller, K. F. Thompson, P. Johnston, D. Santillo, *Front. Mar. Sci.*, **4**, 418 (2018).
4. S. M. Clegg, R. Wiens, A. K. Misra, S. K. Sharma, J. Lambert, S. Bender, et al., *Appl. Spectrosc.*, **68**, No. 9, 925–936 (2014).
5. J. Lasue, R. C. Wiens, S. M. Clegg, D. T. Vaniman, K. H. Joy, et al., *J. Geophys. Res.: Planets*, **117** (2012).
6. P. Zheng, H. Liu, J. Wang, M. Shi, X. Wang, B. Zhang, et al., *J. Anal. At. Spectrom.*, **30**, No. 4, 867–874 (2015).
7. G. U. O. Jinjia, A. S. Mahmoud, L. I. Nan, S. O. N. G. Jiaojian, R. Zheng, *Plasma Sci. Technol.*, **21**, No. 3, 034022 (2019).
8. A. P. Michel, M. Lawrence-Snyder, S. M. Angel, A. D. Chave, *Appl. Opt.*, **46**, No. 13, 2507–2515 (2007).
9. C. Goueguel, D. L. McIntyre, J. Jain, A. K. Karamalidis, C. Carson, *Appl. Opt.*, **54**, No. 19, 6071–6079 (2015).
10. C. Goueguel, J. P. Singh, D. L. McIntyre, J. Jain, A. K. Karamalidis, *Appl. Spectrosc.*, **68**, No. 2, 213–221 (2014).
11. C. L. Goueguel, J. C. Jain, D. L. McIntyre, C. G. Carson, H. M. Edenborn, *J. Anal. At. Spectrom.*, **31**, 1374–1380 (2016).
12. C. L. Goueguel, C. R. Bhatt, J. C. Jain, C. L. Lopano, D. L. McIntyre, *Opt. Laser Technol.*, **108**, 53–58 (2018).
13. N. Li, J. Guo, C. Zhang, Y. Zhang, Q. Li, Y. Tian, R. Zheng, *Appl. Opt.*, **58**, No. 14, 3886–3891 (2019).
14. M. Dell’Aglia, M. Lopez-Claros, J. J. Laserna, S. Longo, A. De Giacomo, *Spectrochim. Acta B: At. Spectroscopy*, **147**, 87 (2018).
15. V. Sathiesh Kumar, Nilesh J. Vasa, R. Sarathi, *J. Phys. D: Appl. Phys.*, **48**, 435504 (2015).
16. Rajendhar Junjuri, Arun Prakash Gummadi, Manoj Kumar Gundawar, *Optic*, **204**, 163946 (2020).
17. Tomoko Takahashi, Soichi Yoshino, Yutaro Takaya, Tatsuo Nozaki, Koichi Ohki, Toshihiko Ohki, Tet-suo Sakka, Blair Thornton, *Deep Sea Res. I: Oceanogr. Res. Papers*, **158**, 103232 (2020).
18. Liwen Sheng, Tianlong Zhang, Guanghui Niu, Kang Wang, Hongsheng Tang, Yixiang Duan, Hua Li, *J. Anal. At. Spectrom.*, **30**, 453–458 (2015).
19. H. Li’ao, W. A. N. G. Qianqian, Z. H. A. O. Yu, L. Li, P. Zhong, *Plasma Sci. Technol.*, **18**, No. 6, 647 (2016).
20. S Anubha Pearline, V. Sathiesh Kumar, *IET Image Process.*, **13**, No. 12, 2176–2182 (2019).
21. S. Marsland, *Machine Learning a Algorithm Perspective*, 2nd ed., CRC Press (2015).
22. NIST Handbook of basic atomic Spectroscopic data. <http://physics.nist.gov/PhysRefData/Handbook/index.html>.
23. H. Hou, Y. Tian, Y. Li, R. Zheng, *J. Anal. At. Spectrom.*, **29**, 169–175 (2014).