Article

Improved Direction-of-Arrival Estimation of an Acoustic Source Using Support Vector Regression and Signal Correlation

Faisal Alam [1], **Mohammed Usman** [2], **Hend I. Alkhammash** [3] **and Mohd Wajid** [4,*]

1   Department of Computer Engineering, Z.H.C.E.T., Aligarh Muslim University, Aligarh 202002, India;
     faisalalam@zhcet.ac.in
2   Department of Electrical Engineering, King Khalid University, Abha 61411, Saudi Arabia;
     omfarooq@kku.edu.sa
3   Department of Electrical Engineering, College of Engineering, Taif University, Taif 21944, Saudi Arabia;
     khamash.h@tu.edu.sa
4   Department of Electronics Engineering, Z.H.C.E.T., Aligarh Muslim University, Aligarh 202002, India
*   Correspondence: mohdwajid@zhcet.ac.in; Tel.: +91-571-270-0922 (ext. 3147)

Abstract: The direction-of-arrival (DoA) estimation of an acoustic source can be estimated with a uniform linear array using classical techniques such as generalized cross-correlation, beamforming, subspace techniques, etc. However, these methods require a search in the angular space and also have a higher angular error at the end-fire. In this paper, we propose the use of regression techniques to improve the results of DoA estimation at all angles including the end-fire. The proposed methodology employs curve-fitting on the received multi-channel microphone signals, which when applied in tandem with support vector regression (SVR) provides a better estimation of DoA as compared to the conventional techniques and other polynomial regression techniques. A multilevel regression technique is also proposed, which further improves the estimation accuracy at the end-fire. This multilevel regression technique employs the use of linear regression over the results obtained from SVR. The techniques employed here yielded an overall 63% improvement over the classical generalized cross-correlation technique.

1. Introduction

Applications such as hands-free mobile communication, hearing aids, target tracking, surveillance of aerial targets, etc., require a close estimation of the direction-of-arrival of a sound source [1–18]. Techniques based on microphone arrays and acoustic vector sensors (AVSs) can be employed for the accurate estimation of the DoA of the incoming acoustic wave [19–27]. However, the DoA estimation using the above techniques encounters practical challenges as the sound wave undergoes reflections and scattering from several objects and the surface enclosures in the surroundings. The irregular reflections of the sound waves create reverberations, and the presence of unwanted interfering sound sources causes a deterioration in the quality of the sound wave. Furthermore, ambient noise and sensor noise give rise to an additional disturbance in the acoustic wave. The estimation of the DoA can be carried out by acquiring signals impinging on a uniform linear array (ULA) of microphones/sensors. Different algorithms can be applied to the digital signals acquired on these microphones to compute the DoA. The classical known techniques/algorithms are beamforming, maximum-likelihood, the subspace method, time-difference of arrival (TDoA), etc. [28–32]. However, due to the presence of noise, interference and reverberations cause a higher deviation from the true value of the DoA, therefore being less suitable for many applications. In [33], the phase-mismatch error and gain-mismatch error among the sensors of the ULA were rectified using a compensated covariance matrix and phase

retrieval for DOA estimation. In [34], a high-resolution, low-complexity method was presented with the use of unfolded coprime linear arrays, where the uniform property of the sub-arrays and the polynomial root finding method were used. A strategy was proposed to overcome the effect of sensor failure in a co-prime array for DoA estimation by employing the singular-value thresholding algorithm [35]. To overcome the grid-mismatch limitation, a solution was proposed in [36] that addressed the DOA estimation problem in an off-grid mode under a sparse framework. The authors in [37] reported a method for near-field and far-field localization with higher accuracy for underdetermined cases by exploiting the co-array property.

The advent of machine learning (ML) in the present era has opened up avenues for the exploration of different ML algorithms for DoA estimation [38]. In this paper, the acoustic digital signals acquired by the ULA were used to compute the feature, Pearson product moment correlation coefficients (PPMCCs). The PPMCCs paired with the known DoA were used to train the ML model for the prediction of the DoA. We carried out a comparative study for the performance of DoA estimation using multiple ML algorithms viz. linear regression, multivariate polynomial regression, and support vector regression (SVR), and compared its result with a classical technique based on the TDoA using generalized correlation coefficients (GCCs). After the assessment of the best ML model for the DoA estimation, we further proposed a curve-fitting-based pre-processing technique for improving the DoA estimation. Furthermore, a multilevel regression scheme was proposed for reducing the error in the DoA estimate at the end-fire.

The rest of the paper is organized as follows. In Section 2, the signal model is explained. Section 3 gives a brief discussion of the techniques used. In Section 4, the methodologies used are explained, and the progressive improvement with comparative results is presented in each subsection. Section 5 concludes the paper.

2. Signal Model

The incoming acoustic waves moving with speed c were assumed to be from the source placed in the far-field; therefore, the incident wave-front was planar. The received signals were assumed to be a narrowband signal, $s(t)$, with the center frequency F (where $F = c/\lambda$ and λ is the wavelength). It was assumed that the sound source and the ULA were in the same plane. The DoA with respect to the normal of the ULA is denoted by θ. There were M number of microphones in a ULA, and each microphone was assumed to be of a point size and to have an omnidirectional pattern. The adjacent microphones in the ULA were separated by a distance d, as shown in Figure 1.
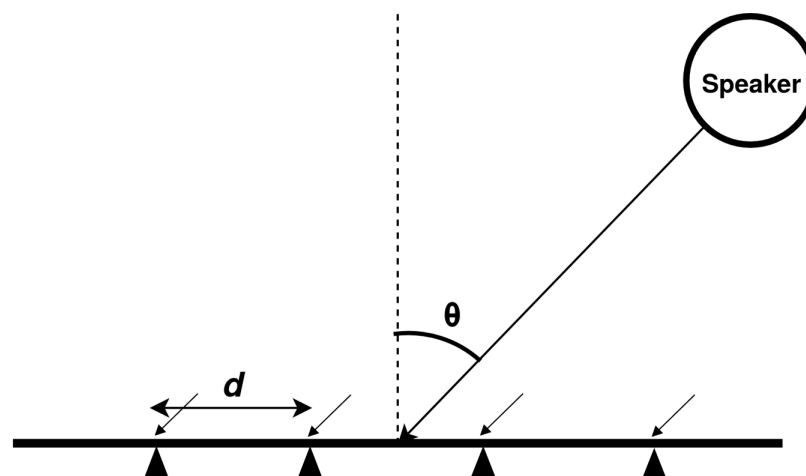

Figure 1. Signal impinging on a uniform linear array, where filled triangles indicate the position of omni-directional microphones.

The signal received by the mth (m = 1 to M) sensor of the array was a phase shifted signal in the frequency domain, which can be formulated as:

$$x_m(t) = s(t)e^{-j2\frac{\pi}{\lambda}D_m} + n_m(t) \tag{1}$$

where D_m is the wave path difference between different microphones in the ULA, which is expressed as $D_m = (m-1)d\sin\theta$, and $n_m(t)$ is the noise added to the mth sensor. We can rewrite Equation (1) as:

$$\mathbf{x}(t) = [x_1(t), x_2(t), ...x_m(t)]^T = \mathbf{a}(\theta)s(t) + \mathbf{n}(t). \tag{2}$$

where $\mathbf{a}(\theta)$ is the steering vector of the ULA, $\mathbf{n}(t)$ is the noise vector, and $[.]^T$ denotes the transpose. The $M \times M$ correlation matrix $\mathbf{R_{xx}}$ of received signal vector $\mathbf{x}(t)$ is expressed as:

$$\mathbf{R_{xx}} = E\Big[\mathbf{x}(t)\mathbf{x}^H(t)\Big] = \mathbf{a}(\theta)\mathbf{S}\mathbf{a}^H(\theta) + \mathbf{R_n}, \tag{3}$$

where $E[.]$ and $[.]^H$ denote the ensemble average and conjugate transpose, respectively. The signal and noise correlation matrices $\mathbf{S}$ and $\mathbf{R_n}$ can be expressed as:

$$\mathbf{S} = E\Big[\mathbf{s}(t)\mathbf{s}^H(t)\Big] \tag{4}$$

and

$$\mathbf{R_n} = E\Big[\mathbf{n}(t)\mathbf{n}^H(t)\Big], \tag{5}$$

respectively. We assumed that all the noise components were zero mean, mutually uncorrelated, and had the same power. Thus, we have:

$$\mathbf{R_n} = \sigma^2\mathbf{I}, \tag{6}$$

where $\mathbf{I}$ is the identity matrix and σ^2 is the noise power. Then, Equation (3) can be written as:

$$\mathbf{R_{xx}} = \mathbf{a}(\theta)\mathbf{S}\mathbf{a}^H(\theta) + \sigma^2\mathbf{I}. \tag{7}$$

3. Brief Discussion of the Techniques Used

In this paper, the process of DoA estimation employed several techniques, which when applied in tandem helped the near-accurate estimation of the DoA. A brief discussion of these techniques is given in the following subsections.

3.1. Polynomial Regression

Linear regression is a technique that helps to find a linear relationship between predictors x and the response y. On the other hand, polynomial regression is a technique that helps identify a non-linear relationship between predictors and the response. As explained in [39], the degree of polynomial regression has to be predetermined before the training. Based on the degree of the polynomial, n, a parameterized equation is of the form:

$$y = \mathbf{b}\mathbf{x} + \epsilon, \tag{8}$$

where $\mathbf{b} = [b_1, b_2...b_n]$ is the parameter vector to be estimated for the best fit and $\mathbf{x} = [x, x^2, x^3, ..., x^n]^T$. The non-linear terms like x^2, x^3, ... , are considered to be derived dimensions based on the base dimension x. The polynomial regression involves performing multivariate linear regression considering higher order terms to be a separate dimension. A univariate regression is a regression involving a single predictor variable, whereas a multivariate regression involves multiple predictor variables.

3.2. Support Vector Regression

Support vector regression (SVR) is a technique that finds a non-linear mathematical relationship between predictors and the response where the prior information of the polynomial degree is not required [40,41]. The technique involves projecting the predictor space into a multidimensional space using a kernel function. In this work, the radial basis function (RBF) was used as the kernel function. The RBF is given as:

$$K(\mathbf{x}, \mathbf{x}') = exp(-\gamma|\mathbf{x} - \mathbf{x}'|), \tag{9}$$

where $\mathbf{x}$ and $\mathbf{x}'$ are the two vectors in the feature space. This kernel function expands into multidimensional terms. The error estimation is measured using the following loss function,

$$L = \max \begin{cases} 0, & |\mathbf{y} - F(\mathbf{x}, \hat{\mathbf{w}})| < \epsilon \\ |\mathbf{y} - F(\mathbf{x}, \hat{\mathbf{w}})| - \epsilon, & \text{otherwise.} \end{cases} \tag{10}$$

As mentioned in [40], $F(\mathbf{x}, w)$ is a family of functions parameterized by $\mathbf{w}$, $\hat{\mathbf{w}}$ is that value of $\mathbf{w}$ that minimizes a measure of the error between $\mathbf{y}$ and F($\mathbf{x}$, $\mathbf{w}$). This loss function is termed as the ϵ-insensitive loss function. It identifies a high-dimensional tube of diameter ϵ. If the estimate is within the tube, then the loss is zero; otherwise, the loss is the distance from the point of estimation to the closest tube periphery. The objective is to flatten this tube to the maximum extent possible. In the training part, linear regression was performed on this high-dimensional space. As a result of the regression process, a linear hyperplane was identified that reduced the overall loss.

The roots of the SVR method are the same as the popular support vector machine (SVM) method, which is used in classification problems and utilizes the same underlying theory. SVM cannot be directly applied here as the DoA estimation was modeled in this work as a regression problem rather than a classification problem.

3.3. Pearson Product Moment Correlation Coefficient

As explained in [42], the Pearson product moment correlation coefficient (PPMCC) quantifies the degree of association between two statistical variables. It also ascertains whether the variables are directly or inversely associated with each other. The PPMCC ranges between +1 and −1. The value +1 indicates the highest degree of association between the variables, which indicates that an increase in one variable is commensurate with an increase in the other variables. The PPMCC value of −1 indicates the highest degree of dissociation between the two variables where an increase in one variable is commensurate with a decrease in another variable. The values in between intervals (−1, +1) are indicative of the corresponding association between the variables proportional to the measure of their values. If L is the sample size, $\{p_i\}_{i=1}^{L}$ and $\{q_i\}_{i=1}^{L}$ are two variables, $\bar{p}$ is the mean of $\{p_i\}_{i=1}^{L}$, and $\bar{q}$ is the mean of $\{q_i\}_{i=1}^{L}$, then the PPMCC, r_{pq}, is given by:

$$r_{pq} = \frac{\sum_{i=1}^{L}(p_i - \bar{p})(q_i - \bar{q})}{\sqrt{\sum_{i=1}^{L}(p_i - \bar{p})^2}\sqrt{\sum_{i=1}^{L}(q_i - \bar{q})^2}}. \tag{11}$$

3.4. Curve Fitting

Curve fitting is the process of identifying a curve that wraps around a series of data points in the best possible manner. The identification of the mathematical model or the curve starts with the proposition of a parameterized mathematical model. The objective of curve fitting identifies the value of these parameters such that the mathematical model minimizes the overall fitting error. For the best fit, the objective function has to be minimized with respect to the parameters, $\mathbf{p}$, using the Levenberg–Marquardt (LM) algorithm explained in [43,44]. The LM algorithm internally uses a combination of two different methods viz. the Gradient Descent Method (GDM) and the Gauss–Newton Method (GNM). In the LM algorithm, the GDM dominates when the parameters are far

apart from their actual values; however, the GNM dominates when the parameters are nearby. The goodness-of-fit is measured using the chi-squared error as given below:

$$\chi^2(\mathbf{p}) = \sum_{i=1}^{m} \left[\frac{g(t_i) - \hat{g}(t_i; \mathbf{p})}{\sigma_{g_i}} \right]^2 \tag{12}$$

where σ_{g_i} is the average error, $\hat{g}(t; \mathbf{p})$ is the fitted function of independent variable t and a vector of n parameters $\mathbf{p}$, and m is the number of data points in the data set.

The combination of the GDM and GNM is accomplished using a parameter λ that is tuned to fall into the appropriate method based on its magnitude,

$$\left[\frac{\partial \hat{g}}{\partial \mathbf{P}}^T \Theta \frac{\partial \hat{g}}{\partial \mathbf{P}} + \lambda \mathbf{I} \right] h_m = \frac{\partial \hat{g}}{\partial p}^T \Theta(g - \hat{g}), \tag{13}$$

where Θ is a diagonal matrix with elements $w_{ii} = 1/\sigma_{g_i}^2$ and h_m is the perturbation parameter that reduces the chi-squared error. To reach closer to the global minimum, the first few steps are taken to be small in the steepest direction. This is accomplished by keeping the value of λ small. The equation behaves closer to the Gauss–Newton update with a small λ. If an iteration gives a high error, then λ is increased and the equation behaves like gradient descent. The scipy package in Python provides the function *curve_fit* with its optimized class that provides the curve fitting function using multiple algorithms. In this work, we used this function for the LM algorithm. Since the acoustic waves received at the microphone array were sinusoidal with additive white Gaussian noise (AWGN), to fit the curve, sinusoidal function $a\ sin(bx + \phi)$ was been used where a, b, and ϕ are the parameters to be tuned to fit the curve.

4. Methodologies and Results

The ULA of omnidirectional microphones was used to record the spatial signals for developing the machine learning model for DOA estimation. The data set consisted of multiple recordings of microphones of a duration of 25 ms each, which were acquired for angles 0°, 2°, 4°, 6°, ..., 90° of a sound source with different signal-to-noise ratio (SNR) values. For each such angle with sensor noise of SNR = 26 dB, one-thousand four-hundred independent realizations of 25 ms in duration were used for training. For testing, a new data set was created with angles 0° to 90° with an increment of 1° for SNR = 22 dB, 18 dB, 14 dB, and 10 dB. For the recorded signals at each angle, we took the PPMCC between the discrete signals from the microphones. Discrete signals from the 1st, 2nd, 3rd, and 4th microphone produced six pairs of correlations. These signals were further processed to estimate the DoA. The following subsections explain the different methodologies that were applied for DoA estimation and exhibit the result obtained with these methods. Since various techniques and methodologies were assessed for improving the DoA estimation, each succeeding subsection employs a technique that improves over the best obtained in the preceding subsection. The result is obtained in terms of root mean squared angular error ($RMSAE$) and $\overline{RMSAE}$ and in each subsection, and it is compared with the results in the preceding subsections. The $RMSAE$ and $\overline{RMSAE}$ are defined in Equations (14) and (15), respectively,

$$RMSAE(\theta) = \sqrt{\frac{\sum_{i=1}^{N}(\theta - \theta_i)^2}{N}} \tag{14}$$

and

$$\overline{RMSAE} = \frac{1}{TOA} \sum_{\theta=0°}^{90°} RMSAE(\theta) \tag{15}$$

where θ_i is the ith prediction of the true angle θ, N is the total number of predictions realized, and the TOA is the total number of angles observed.

4.1. Regression Techniques with PPMCC

The PPMCCs were calculated between different signals $s_1(t)$, $s_2(t)$, $s_3(t)$, and $s_4(t)$ recorded by four microphones m_1, m_2, m_3, and m_4, respectively. The PPMCC between each microphone signal produced six coefficients, c_{12}, c_{13}, c_{14}, c_{23}, c_{24}, and c_{34}. These PPMCCs were used for training on the data set to produce the mathematical model of the DoA estimate. We trained the signals on the data set with SNR = 26 dB. The feature set was composed of the correlation coefficients mentioned above, and the response was the actual DoA. The regression techniques that were used and analyzed in this experiment were SVR, polynomial regression of order 1 (PR1), polynomial regression of order 2 (PR2), polynomial regression of order 3 (PR3), polynomial regression of order 4 (PR4), polynomial regression of order 5 (PR5), polynomial regression of order 6 (PR6), and polynomial regression of order 7 (PR7). In addition to these regression techniques, we also estimated the DoA for comparison with the conventional generalized cross correlation (GCC) technique.

Figure 2 shows the comparative assessment of SVR, PR1, PR2, and GCC. Polynomial regression of orders higher than two is not shown in this figure as their $RMSAE$ values were too high and difficult to show. In the following section, we propose a mechanism to reduce their $RMSAE$ values. Figure 2 also reveals that the SVR gave a low $RMSAE$ for higher SNR values, i.e., 22 dB and 18 dB. However, at lower SNR values of 14 dB and 10 dB, the GCC was robust. Moreover, the GCC provided an $RMSAE$ closer to SVR even on higher SNR values. In short, the GCC was a better estimator as it consistently gave good approximation at low, as well as high SNR values (the next section improves the performance of ML algorithms). Among the regression techniques, SVR performed better than PR1, which in turn performed better than PR2. The $\overline{RMSAE}$ of all the regression techniques and GCC are shown in Table 1. Another observation was that the regression techniques had a higher $RMSAE$ at the broadside (ranging between $0°$ to $5°$), as well as at the end-fire (ranging from $80°$ to $90°$). The cause for higher $\overline{RMSAE}$ near the end-fire was identified as follows. Consider the data set with SNR = 26 dB: the PPMCCs were computed. Then, we took the ensemble average of PPMCCs ($E[c_{12}]$, $E[c_{13}]$, $E[c_{14}]$, $E[c_{23}]$, $E[c_{24}]$, and $E[c_{34}]$) for each DoA, which is shown in Figure 3. This revealed that the correlations at the end-fire were steady with almost the same values, which was due to a smaller change in the relative time-delay with respect to the DoA. The relative time-delay, $\tau_{i,i+1}$, between the signals of two adjacent microphones ($m_i(t)$ and $m_{i+1}(t)$) for a planar wave-front are approximated by:

$$\tau_{i,i+1} \approx \frac{d}{c} sin(\theta), \tag{16}$$

where d is the microphone separation, c is the speed of sound, and θ is the direction of arrival of a planar wave with respect to the axis normal to the ULA. The rate of change of relative delay with respect to θ is given by:

$$\frac{d\tau_{i,i+1}}{d\theta} \approx \frac{d}{c} cos(\theta), \tag{17}$$

which shows that at $\theta \approx 90^o$, the $\frac{d\tau}{d\theta}$ is small, therefore the rate of change of the PPMCC is also small. This steady value of $\tau_{i,i+1}$ caused the same values of all ensemble PPMCC over a span of DoAs, thereby causing the regression techniques to have an error in the estimation of the DoA near end-fire. For the no noise case, the signal at all microphones would have similar waveforms, which would indicate that the PPMCC should be unity; however, random noise at the microphones caused the random variation of the PPMCC, and thereby, the learning/training of the machine was poor from the data at the broadside, hence the higher $\overline{RMSAE}$.
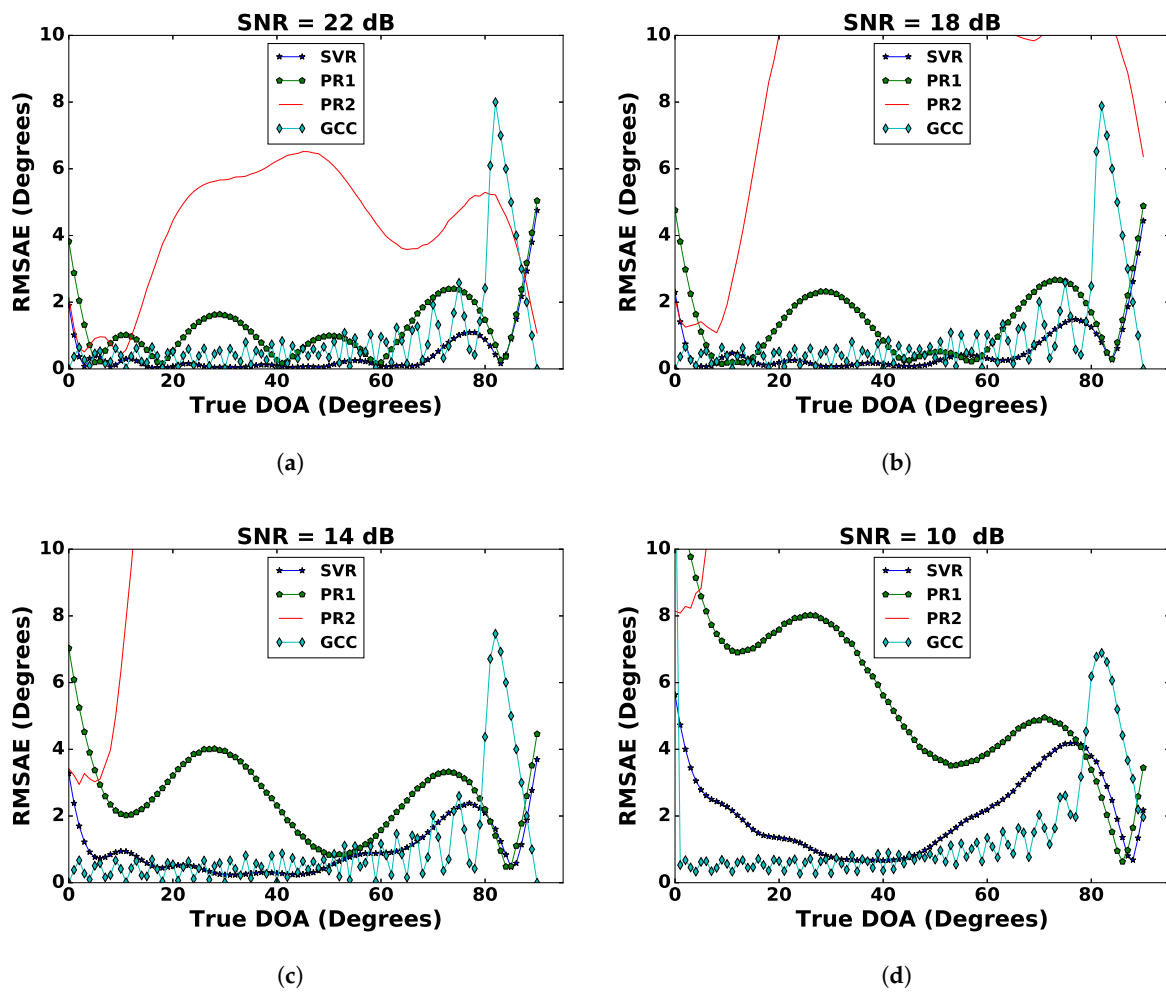
Figure 2. Comparison of different regression techniques (SVR, PR1, and PR2) and the GCC for (**a**) SNR = 22 dB, (**b**) SNR = 18 dB, (**c**) SNR = 14 dB, and (**d**) SNR = 10 dB.
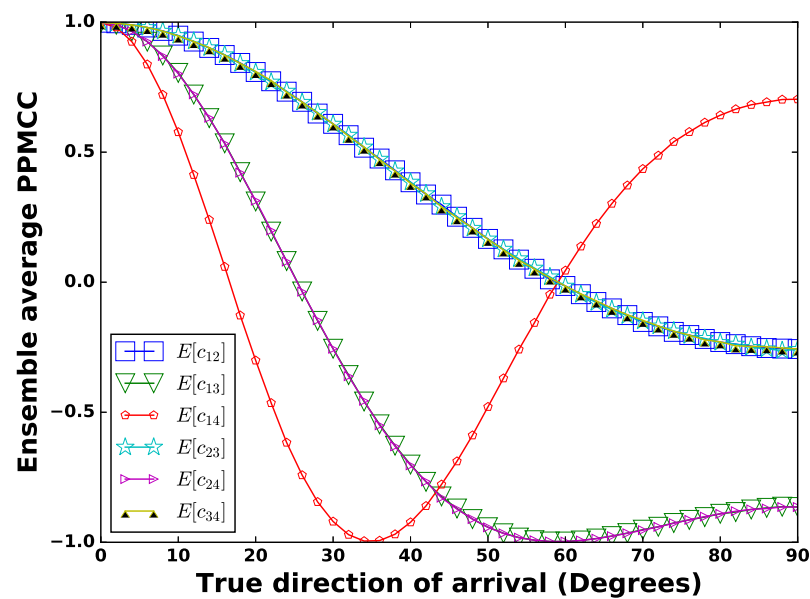

Figure 3. Ensemble average of the PPMCC versus the direction of arrival at SNR = 26 dB.

4.2. Improvement with Curve Fitting

To improve on the DoA estimation using regression techniques, curve fitting was applied on the noisy sinusoidal signal recorded by the microphones, which was described in Section 3.4. The process of curve fitting reduced the noise in the recorded signals, thereby providing a sanitized input for improved results. After pre-processing with curve fitting on the recorded noisy signals, the regression techniques mentioned in Section 4 were applied. The *RMSAE* versus DoA results with the SNR ranging from 22 to 10 dB with a decrement of 4 dB are shown in Figure 4. It also shows the results obtained from the GCC for comparison. A closer look at the results reveals that pre-processing with curve fitting reduced the *RMSAE* of the DoA estimates with regression techniques. Amongst all the regression techniques, SVR outperformed then in terms of the *RMSAE*. Polynomial regression of higher order yielded poor results of DoA estimation without curve fitting with a too high *RMSAE*. It can be seen that after pre-processing with curve fitting, the *RMSAE* of the DoA estimate was significantly reduced. The GCC continued to perform better with a consistently lower *RMSAE*, but it had a highly rugged curve with respect to the true DoA. Figure 5 shows the performance comparison of SVR and the GCC with and without curve fitting. It was observed that SVR with curve fitting had a lower *RMSAE*. A comparison in terms of $\overline{RMSAE}$ of different regression techniques and the GCC with and without curve fitting is shown in Table 1. It can be seen from this table that the $\overline{RMSAE}$ of SVR with curve fitting had the lowest values for each experiment with particular SNR values of 22 dB, 18 dB, 14 dB, and 10 dB. In contrast to the GCC, as the SNR value decreased, the curve fitting showed more improvement in the performance of SVR.
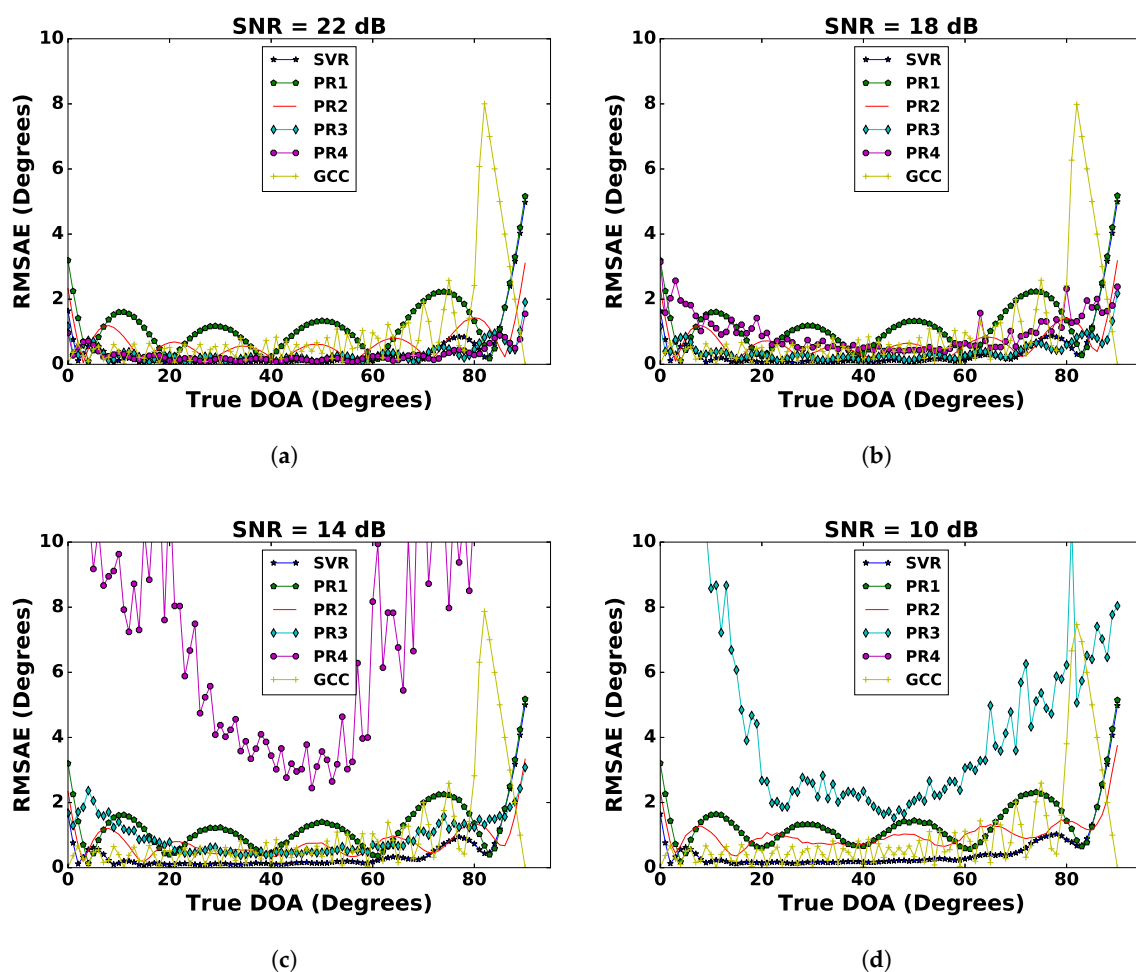

Figure 4. Comparison of different regression techniques (SVR, PR1, and PR2) and the GCC with curve fitting for (**a**) SNR = 22 dB, (**b**) SNR = 18 dB, (**c**) SNR = 14 dB, and (**d**) SNR = 10 dB.
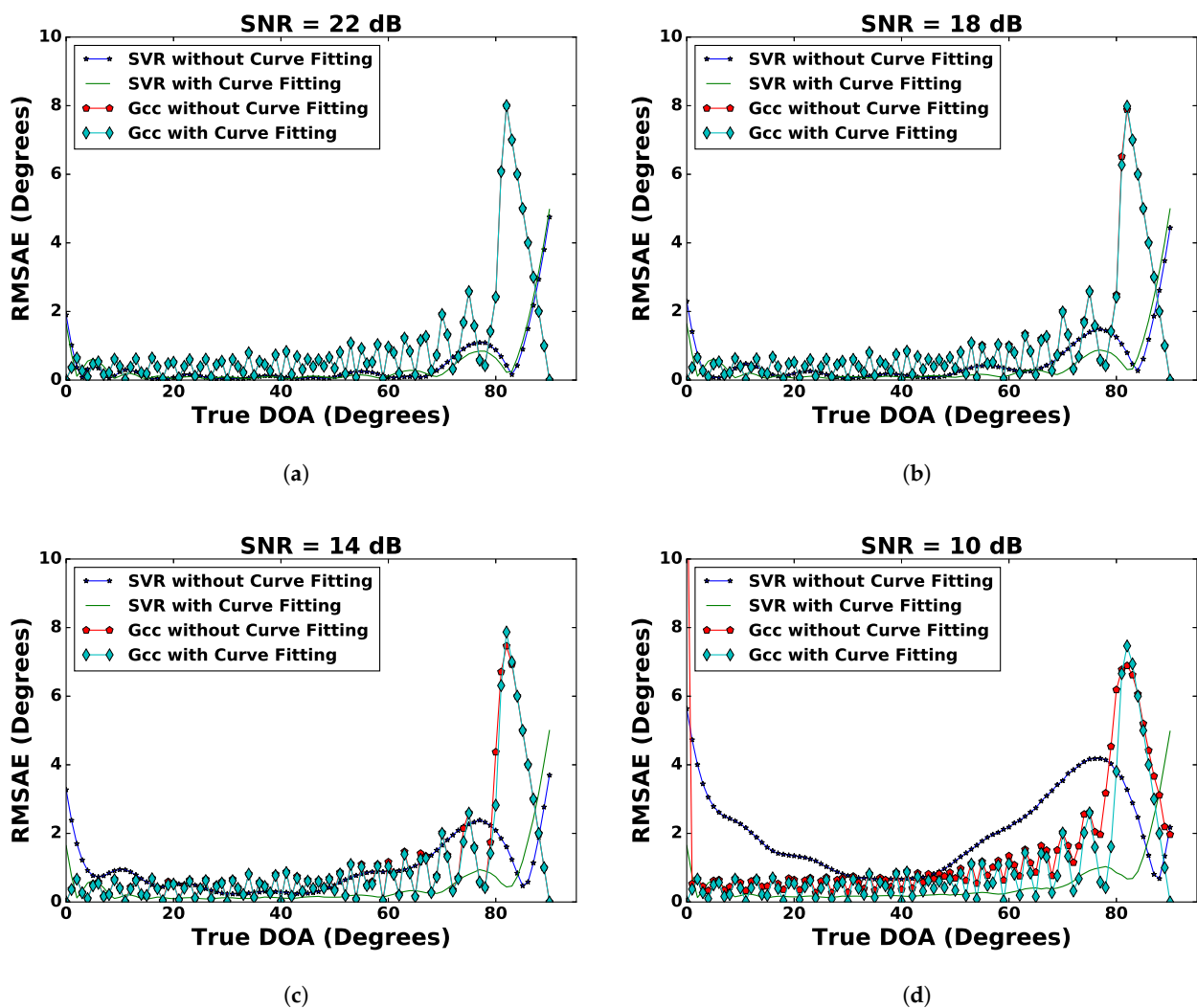
Figure 5. Comparison of SVR and the GCC with and without curve fitting for (**a**) SNR = 22 dB, (**b**) SNR = 18 dB, (**c**) SNR = 14 dB, and (**d**) SNR = 10 dB.

4.3. DoA Estimation Improvement at the End-Fire

As mentioned in Section 4.1, it can be observed from all the preceding results that the *RMSAE* error was consistently high at the end-fire ranging from 80° and above. To mitigate this high error, the bias and standard deviation analysis of the DoA estimates were performed. Based on this analysis, we proposed a multilevel regression to reduce the bias and further improve the accuracy at the end-fire.

4.3.1. Analysis of the DoA Estimation: Bias and Standard Deviation

The bias of a DoA estimator is the difference between the estimated DoA and the true DoA. If the bias of a DoA estimator is a certain value, then the estimation can be improved by subtracting the bias from the estimated value. The standard deviation of the DoA estimate indicates how much an estimated DoA differs from its expected value, which is a random error and cannot be compensated for a given estimate. Figures 6 and 7 show the bias and standard deviation of the DoA estimation for SVR and the GCC with curve fitting. It can be observed from Figures 6 and 7 that the bias for SVR was stable, whereas the bias for the GCC fluctuated with respect to true DoA. This stable bias in the SVR was beneficial as it aided in bias error compensation. The standard deviation of the DoA estimate for SVR

was relatively more stable than the GCC. In contrast to SVR, the standard deviation of the DoA estimate for the GCC showed spikes. Therefore, it can be concluded that SVR with curve fitting provided the best DoA estimate amongst all techniques considered.

Table 1. $\overline{RMSAE}$ (degrees) values with multiple regressions and conventional techniques.

| Base Technique | SNR (dB) | $\overline{RMSAE}$ (Degrees) | |
		Without Curve Fitting	With Curve Fitting
SVR	22	0.428	0.41
	18	0.551	0.429
	14	0.966	0.464
	10	2.039	0.53
PR1	22	1.212	1.194
	18	1.418	1.217
	14	2.587	1.264
	10	5.567	1.363
PR2	22	4.23	0.655
	18	10.351	0.687
	14	25.232	0.772
	10	59.379	1.034
PR3	22	22.086	0.326
	18	Too High	0.398
	14	Too High	0.948
	10	Too High	4.811
PR4	22	Too High	0.286
	18	Too High	0.983
	14	Too High	7.95
	10	Too High	73.604
PR5	22	Too High	0.27
	18	Too High	1.932
	14	Too High	19.671
	10	Too High	High
PR6	22	Too High	0.275
	18	Too High	2.183
	14	Too High	21.895
	10	Too High	High
PR7	22	Too High	0.326
	18	Too High	2.967
	14	Too High	31.947
	10	Too High	High
GCC	22	0.983	0.983
	18	0.988	0.988
	14	0.998	0.998
	10	1.027	1.027
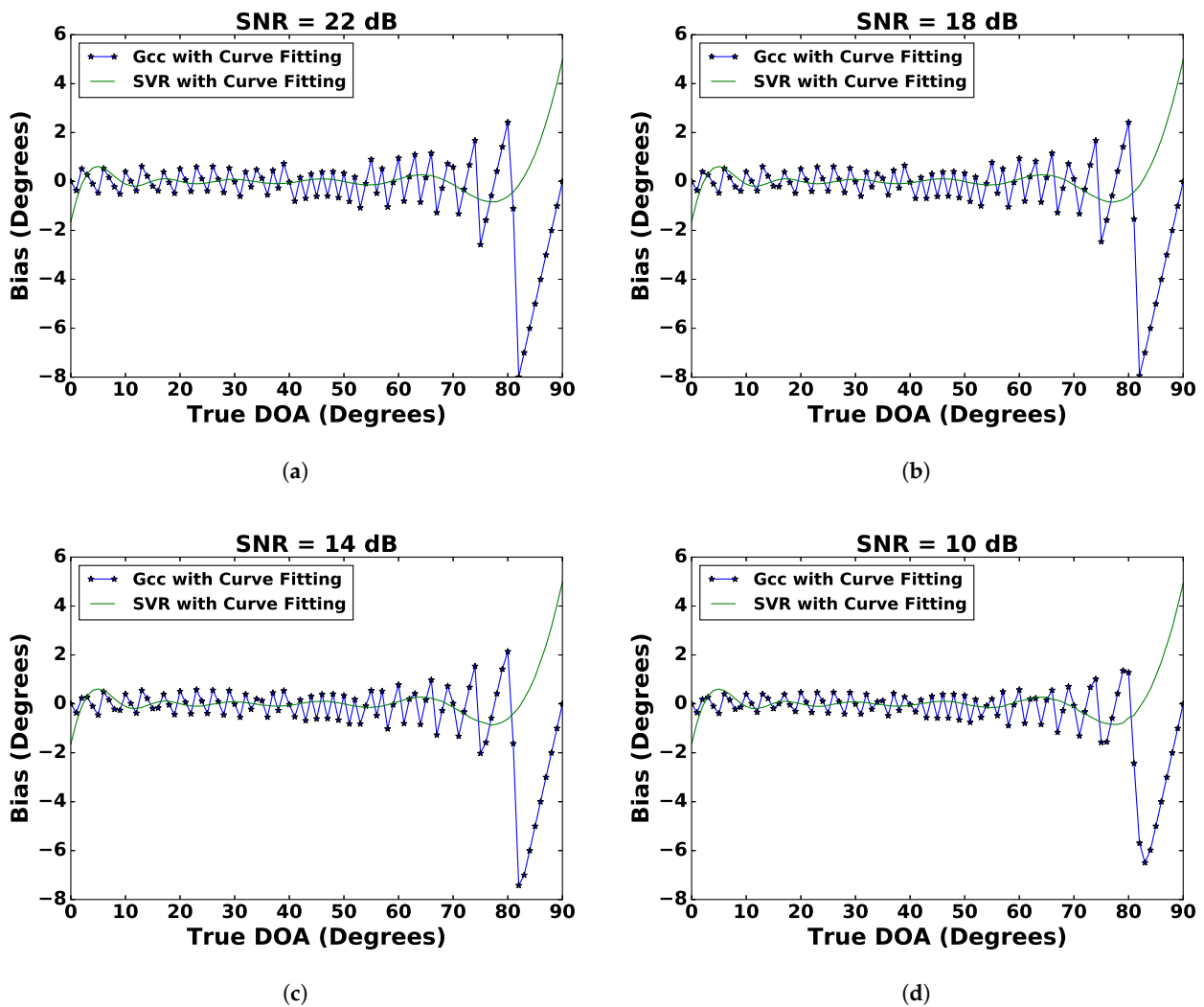
Figure 6. Bias comparison of SVR with curve fitting and the GCC with curve fitting for (**a**) SNR = 22 dB, (**b**) SNR = 18 dB, (**c**) SNR = 14 dB, and (**d**) SNR = 10 dB.

4.3.2. Improved DoA Estimation with Multilevel Regression

It can be observed from Figure 4 that despite curve fitting, the DoA estimation error in the range between 80° and 90° (end-fire) was relatively higher as compared to the DoA ranging between 0° and 80°. Even after applying curve fitting, there was no substantial improvement in DoA estimation at the end-fire. Figure 6 reveals that the SVR bias curve for DoA estimation was increasing near the end-fire. To remove this bias, a second-level regression model (SLRM) was applied at the end-fire. The input to the SLRM was the DoA estimated from the SVR with curve fitting (SVR-CF) model, and the output was expected to compensate for this bias. The SLRM was trained with linear regression (LR) for angles estimated to be between 80° and 90° at SNR = 26 dB. The response from this second-level regression for (a) SNR = 22 dB, (b) SNR = 18 dB, (c) SNR = 14 dB, and (d) SNR = 10 dB is shown in Figure 8. It was observed that the *RMSAE* was reduced significantly at the end-fire with the maximum *RMSAE* reducing from 6° to 2° by this improvement. Table 2 compares the results of the SVR-CF model and the SVR-CF tandem with LR (SVR-CF-LR) model in terms of $\overline{RMSAE}$. It can be seen that the $\overline{RMSAE}$ values significantly reduced, indicating the SLRM further improved the DoA estimation.
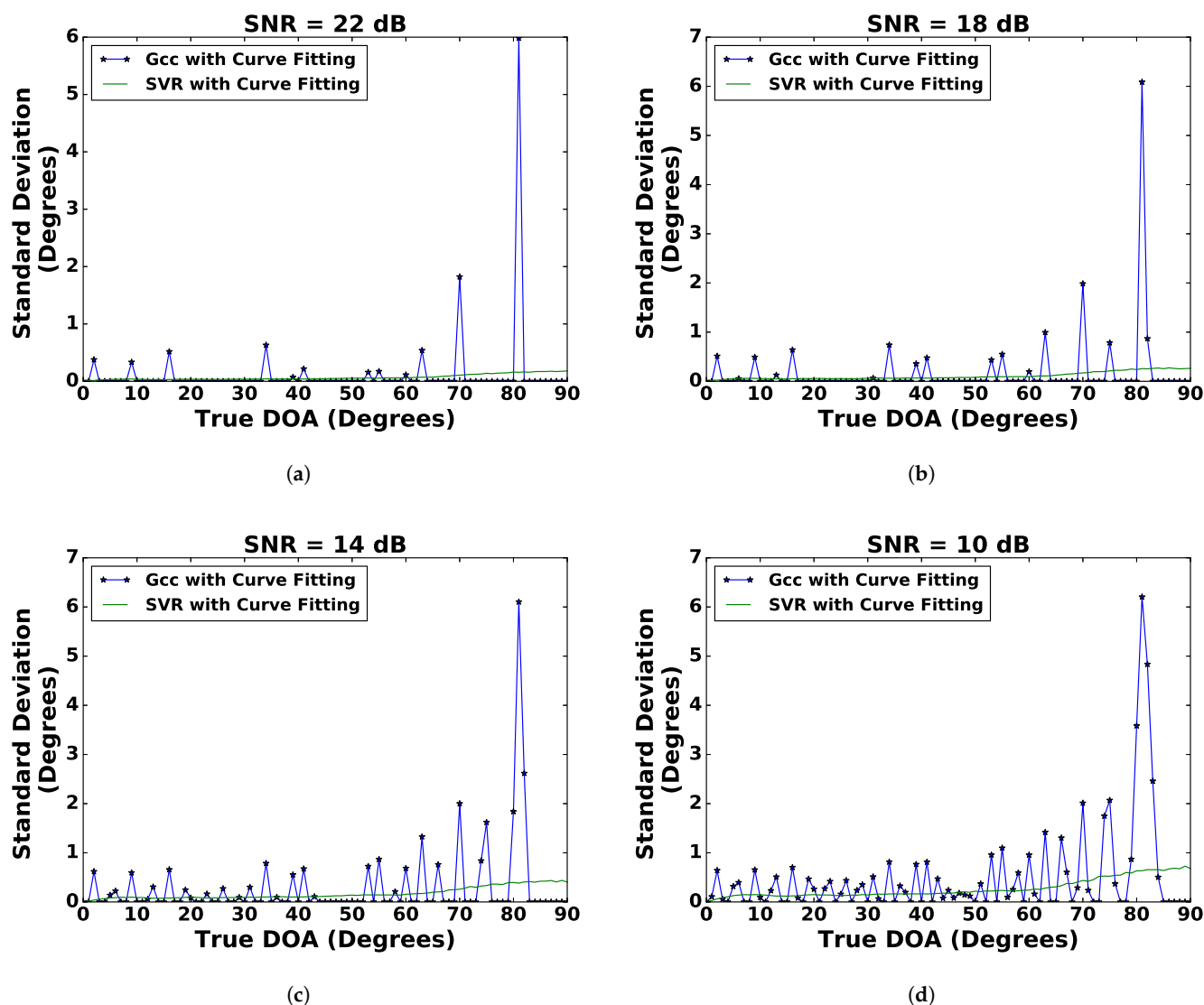
Figure 7. Comparison of the standard deviation of SVR with curve fitting (SVR-CF) and the GCC with curve fitting for (**a**) SNR = 22 dB, (**b**) SNR = 18 dB, (**c**) SNR = 14 dB, and (**d**) SNR = 10 dB.

Table 2. $\overline{RMSAE}$ of the DoA estimate for SVR-CF and SVR-CF-LR at the end-fire.

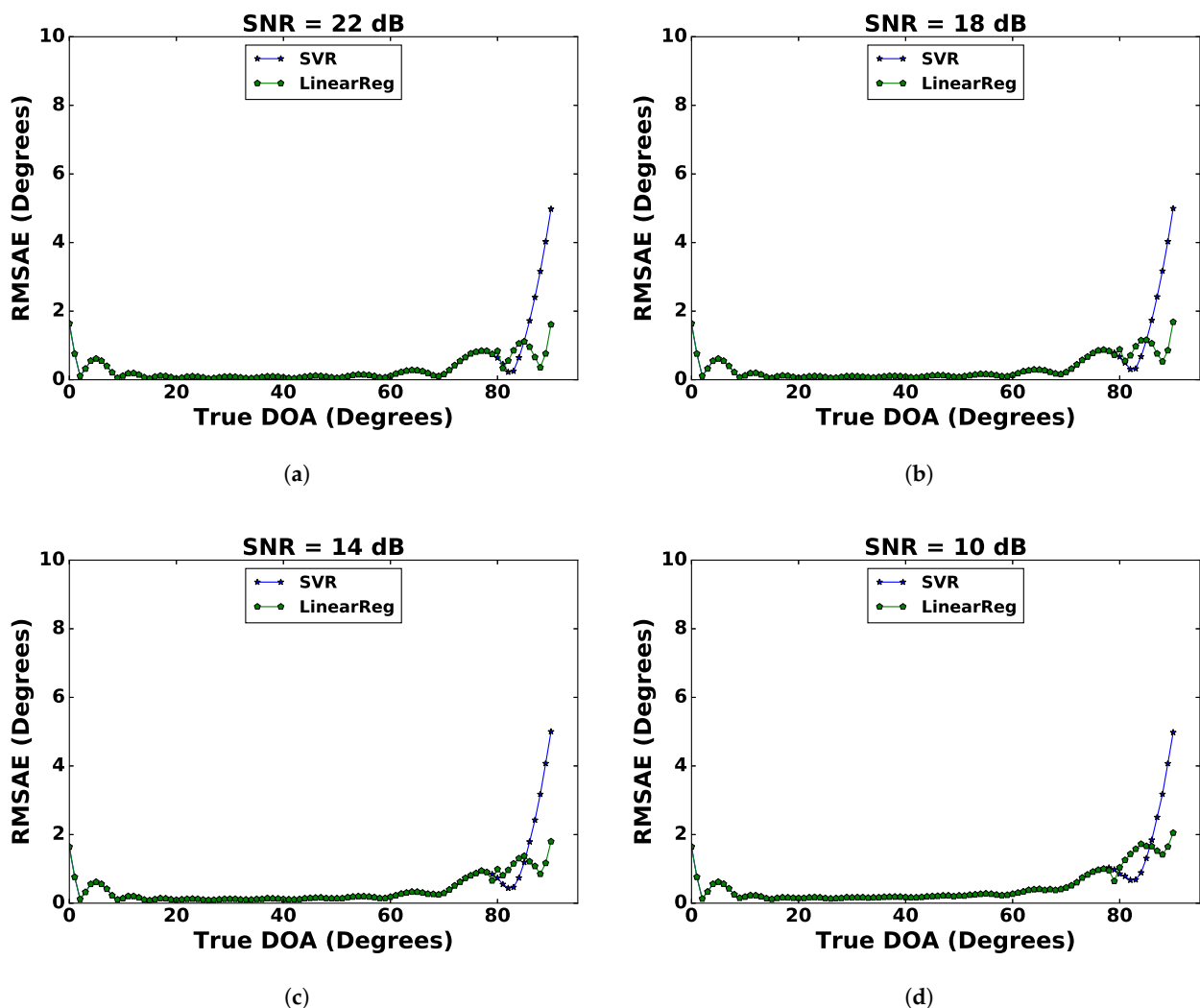| Technique	SNR (dB)	$\overline{\text{RMSAE}}$ (Degrees)
SVR-CF	22	0.410
	18	0.429
	14	0.464
	10	0.530
SVR-CF-LR	22	0.294
	18	0.322
	14	0.375
	10	0.473

Figure 8. Comparison between SVR-CF and SVR-CF-LR for the DoA estimate in terms of the *RMSAE* at the end-fire for (**a**) SNR = 22 dB, (**b**) SNR = 18 dB, (**c**) SNR = 14 dB, and (**d**) SNR = 10 dB.

5. Conclusions

DoA estimation of a sound source has many applications such as unmanned aerial vehicles, hearing aids, surveillance, etc. Many techniques such as beamforming and the GCC, employing a uniform linear array of microphones, have been studied in the past. However, these techniques have not been as effective and produced errors in the estimated angle owing to the impact of noise that blends with the received signals. This work aimed at increasing the efficacy of DoA estimation by exploring multiple machine learning techniques. The models for polynomial regression of order one to order seven and SVR were trained with the PPMCC as the feature selected, then compared with the classical GCC technique for comparative analysis. Since the results were not very impressive, we explored pre-processing techniques on the incoming signals for improved results. Curve-fitting as the pre-processing technique was applied, and the same regression models PR 1-7 and SVR were trained and tested after pre-processing and compared with each other. Among the regression techniques used, SVR fared the best with a minimum error when compared with the PR techniques and the GCC. However, all the techniques were shown to yield high errors near the end-fire. To reduce the error at the end-fire, multilevel regression was applied with linear regression as the second-level regression on the results of SVR. This technique proved to be nearly accurate, stable, and unbiased and produced approximately a 63% improvement in the estimated angle when compared with the classical GCC technique.

Abbreviations

The following abbreviations are used in this manuscript:

DoA	Direction of arrival
SVR	Support vector regression
AVS	Acoustic vector sensor
ULA	Uniform linear array
TDoA	Time-difference of arrival
ML	Machine learning
PPMCC	Pearson product moment correlation coefficient
GCC	Generalized correlation coefficient
SVM	Scalable vector machine
LM	Levenberg–Marquardt
RBF	Radial basis function
GDM	Gradient descent method
GNM	Gauss–Newton method
SNR	Signal-to-noise ratio
AWGN	Additive white Gaussian noise
PRn	Polynomial regression of order n
SLRM	Second-level regression model
SVR-CF	SVR with curve fitting
SVR-CF-LR	SVR-CF in tandem with linear regression

References

1. Zheng, X.; Ritz, C.; Xi, J. Encoding and communicating navigable speech soundfields. *Multimed. Tools Appl.* **2016**, *75*, 5183–5204. [CrossRef]
2. Asaei, A.; Taghizadeh, M.J.; Haghighatshoar, S.; Raj, B.; Bourlard, H.; Cevher, V. Binary sparse coding of convolutive mixtures for sound localization and separation via spatialization. *IEEE Trans. Signal Process.* **2016**, *64*, 567–579. [CrossRef]
3. Bekkerman, I.; Tabrikian, J. Target detection and localization using mimo radars and sonars. *IEEE Trans. Signal Process.* **2006**, *54*, 3873–3883. [CrossRef]
4. Wong, K.T.; Zoltowski, M.D. Closed-form underwater acoustic direction-finding with arbitrarily spaced vector hydrophones at unknown locations. *IEEE J. Ocean. Eng.* **1997**, *22*, 566–575. [CrossRef]
5. Sheng, X.; Hu, Y.-H. Maximum likelihood multiple-source localization using acoustic energy measurements with wireless sensor networks. *IEEE Trans. Signal Process.* **2005**, *53*, 44–53. [CrossRef]
6. Zhao, S.; Ahmed, S.; Liang, Y.; Rupnow, K.; Chen, D.; Jones, D.L. A real-time 3d sound localization system with miniature microphone array for virtual reality. In Proceedings of the 2012 7th IEEE Conference on Industrial Electronics and Applications (ICIEA), Singapore, 18–20 July 2012; pp. 1853–1857.
7. Clark, J.A.; Tarasek, G. Localization of radiating sources along the hull of a submarine using a vector sensor array. In Proceedings of the OCEANS 2006, Boston, MA, USA, 18–21 September 2006; pp. 1–3.
8. Carpenter, R.N.; Cray, B.A.; Levine, E.R. Broadband ocean acoustic (boa) laboratory in narragansett bay: preliminary in situ harbor security measurements. In *Defense and Security Symposium*; International Society for Optics and Photonics: Bellingham, USA , 2006; p. 620409.
9. DiBiase, J.H.; Silverman, H.F.; Brandstein, M.S. Robust localization in reverberant rooms. In *Microphone Arrays*; Springer: Berlin, Germany, 2001; pp. 157–180.

10. Bechler, D.; Schlosser, M.S.; Kroschel, K. System for robust 3d speaker tracking using microphone array measurements. In Proceedings of the 2004 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS) (IEEE Cat. No.04CH37566), Sendai, Japan, 28 September–2 October 2004; pp. 2117—2122.

11. Argentieri, S.; Danes, P. Broadband variations of the music high-resolution method for sound source localization in robotics. In Proceedings of the 2007 IEEE/RSJ International Conference on Intelligent Robots and Systems, San Diego, CA, USA, 29 October–2 November 2007; pp. 2009–2014.

12. Nakadai, K.; Matsuura, D.; Okuno, H.G.; Kitano, H. Applying scattering theory to robot audition system: Robust sound source localization and extraction. In Proceedings of the 2003 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2003) (Cat. No.03CH37453), Las Vegas, NV, USA, 27–31 October 2003; pp. 1147–1152.

13. Zhao, S.; Chng, E.S.; Hieu, N.T.; Li, H. A robust real-time sound source localization system for olivia robot. In Proceedings of the 2010 APSIPA Annual Summit and Conference, Biopolis, Singapore, 14–17 December, 2010.

14. Xiao, X.; Zhao, S.; Nguyen, D.H.H.; Zhong, X.; Jones, D.L.; Chng, E.-S.; Li, H. The ntu-adsc systems for reverberation challenge 2014. In *Proc. REVERB Challenge Workshop*; Spoken Language Systems MIT Computer Science and Artificial Intelligence Laboratory: Cambridge, USA, 2014 .

15. Delikaris-Manias, S.; Vilkamo, J.; Pulkki, V. Signal-dependent spatial filtering based on weighted-orthogonal beamformers in the spherical harmonic domain. *IEEE/ACM Trans. Audio Speech Lang. Process.* **2016**, *24*, 1507–1519. [CrossRef]

16. Delikaris-Manias, S.; Pulkki, V. Cross pattern coherence algorithm for spatial filtering applications utilizing microphone arrays. *IEEE Trans. Audio Speech Lang. Process.* **2013**, *21*, 2356–2367. [CrossRef]

17. Zhang, C.; Florêncio, D.; Ba, D.E.; Zhang, Z. Maximum likelihood sound source localization and beamforming for directional microphone arrays in distributed meetings. *IEEE Trans. Multimed.* **2008**, *10*, 538–548. [CrossRef]

18. Van den Bogaert, T.; Carette, E.; Wouters, J. Sound source localization using hearing aids with microphones placed behind-the-ear, in-the-canal, and in-the-pinna. *Int. J. Audiol.* **2011**, *50*, 164–176. [CrossRef]

19. Wajid, M.; Alam, F.; Yadav, S.; Khan, M.A.; Usman, M. Support vector regression based direction of arrival estimation of an acoustic source. In Proceedings of the 2020 International Conference on Innovation and Intelligence for Informatics, Computing and Technologies (3ICT), Sakheer, Bahrain, 20–21 December 2020; pp. 1–6.

20. Wajid, M.; Kumar, A.; Bahl, R. Design and analysis of air acoustic vector-sensor configurations for two-dimensional geometry. *J. Acoust. Soc. Am.* **2016**, *139*, 2815–2832. [CrossRef]

21. Wajid, M.; Kumar, A.; Bahl, R. Bearing estimation in a noisy and reverberant environment using an air acoustic vector sensor. *IUP J. Electr. Electron. Eng.* **2016**, *9*, 53.

22. Wajid, M.; Kumar, A.; Bahl, R. Direction-finding accuracy of an air acoustic vector sensor in correlated noise field. In Proceedings of the 2017 4th International Conference on Signal Processing, Computing and Control (ISPCC), Solan, India, 21–23 September 2017; pp. 21–25.

23. Wajid, M.; Kumar, A.; Bahl, R. Direction-of-arrival estimation algorithms using single acoustic vector-sensor. In Proceedings of the 2017 International Conference on Multimedia, Signal Processing and Communication Technologies (IMPACT), Aligarh, India, 24–26 November 2017; pp. 84–88.

24. Wajid, M.; Kumar, A. Direction estimation and tracking of coherent sources using a single acoustic vector sensor. *Arch. Acoust.* **2020**, *45*, 209–219.

25. Yadav, S.; Wajid, M.; Usman, M. Support vector machine-based direction of arrival estimation with uniform linear array. In *Advances in Computational Intelligence Techniques*; Springer: Singapore, 2020; pp. 253–264.

26. Wajid, M.; Kumar, B.; Goel, A.; Kumar, A.; Bahl, R. Direction of arrival estimation with uniform linear array based on recurrent neural network. In Proceedings of the 2019 5th International Conference on Signal Processing, Computing and Control (ISPCC), Solan, India, 10–12 October 2019; pp. 361–365.

27. Zhou, C.; Gu, Y.; Zhang, Y.D.; Shi, Z.; Jin, T.; Wu, X. Compressive sensing-based coprime array direction-of-arrival estimation. *IET Commun.* **2017**, *11*, 1719–1724. [CrossRef]

28. Liu, A.; Yang, D.; Shi, S.; Zhu, Z.; Li, Y. Augmented subspace music method for doa estimation using acoustic vector sensor array. *IET Radar Sonar Navig.* **2019**, *13*, 969–975. [CrossRef]

29. Shi, F. Two dimensional direction-of-arrival estimation using compressive measurements. *IEEE Access* **2019**, *7*, 20863–20868. [CrossRef]

30. Cui, X.; Yu, K.; Zhang, S.; Wang, H. Azimuth-only estimation for tdoa-based direction finding with 3-d acoustic array. *IEEE Trans. Instrum. Meas.* **2019**, *69*, 985–994. [CrossRef]

31. Meng, Z.; Zhou, W. Direction-of-arrival estimation in coprime array using the esprit-based method. *Sensors* **2019**, *19*, 707. [CrossRef] [PubMed]

32. Varma, K.M. Time Delay Estimate Based Direction of Arrival Estimation for Speech in Reverberant Environments. Ph.D. Thesis, Virginia Tech, Blacksburg, VG, USA, 2002.

33. Zhang, L.; Wu, S.; Guo, A.; Yang, W. A novel direction-of-arrival estimation via phase retrieval with unknown sensor gain-and-phase errors. *Sensors* **2019**, *19*, 2701. [CrossRef] [PubMed]

34. He, W.; Yang, X.; Wang, Y. A high-resolution and low-complexity doa estimation method with unfolded coprime linear arrays. *Sensors* **2020**, *20*, 218. [CrossRef]

35. Sun, B.; Wu, C.; Ruan, H. Array diagnosis and doa estimation for coprime array under sensor failures. *Sensors* **2020**, *20*, 2735. [CrossRef]

36. Wu, X.; Zhu, We.; Yan, J.; Zhang, Z. Two sparse-based methods for off-grid direction-of-arrival estimation. *Signal Process.* **2018**, *142*, 87–95. [CrossRef]

37. Wu, X. Localization of far-field and near-field signals with mixed sparse approach: A generalized symmetric arrays perspective. *Signal Process.* **2020**, *175*, 107665. [CrossRef]

38. Liu, Y.; Chen, H.; Wang, B. Doa estimation based on cnn for underwater acoustic array. *Appl. Acoust.* **2021**, *172*, 107594. [CrossRef]

39. Maia, A.; Ferreira, E.; Oliveira, M.C.; Menezes, L.F.; Andrade-Campos, A. 3-numerical optimization strategies for springback compensation in sheet metal forming. In *Computational Methods and Production Engineering*; Davim, J., Ed.; Woodhead Publishing Reviews: Mechanical Engineering Series; Woodhead Publishing: Cambridge, UK , 2017; pp. 51–82.

40. Drucker, H.; Burges, C.J.C.; Kaufman, L.; Smola, A.; Vapnik, V. Support vector regression machines. In *Advances in Neural Information Processing Systems*; Mozer, M.C., Jordan, M., Petsche, T., Eds.; MIT Press: Cambridge, MA, USA, 1997; Volume 9, pp. 155–161.

41. Smola, A.J.; Schölkopf, B. A tutorial on support vector regression. *Stat. Comput.* **2004**, *14*, 199–222. [CrossRef]

42. King, A.P.; Eckersley, R.J. Chapter 2–Descriptive Statistics II: Bivariate and Multivariate Statistics. In *Statistics for Biomedical Engineers and Scientists*; King, A.P., Eckersley, R.J., Eds.; Academic Press (Imprint of Elsevier): London, UK, 2019; pp. 23–56.

43. Gavin, H.P. The Levenberg-Marquardt Algorithm Fornonlinear Least Squares Curve-Fitting Problems. *J. Dept. Civ. Environ. Eng. Duke Univ.* **2011**, *28*, 1–5.

44. Moré, J.J. The levenberg-marquardt algorithm: Implementation and theory. In *Numerical Analysis*; Watson, G.A., Ed.; Springer: Berlin/Heidelberg, Germany, 1978; pp. 105–116.