Received: 8 March 2018 | Revised: 3 November 2018 | Accepted: 27 December 2018
DOI: 10.1111/lre.12255

ORIGINAL ARTICLE

WILEY Lakes & Reservoirs
Science, Policy and Management for Sustainable Use

A computationally efficient model for determining sound speed in shallow tropical freshwater systems, with field validation

Jyoti Sadalage[1] | **Arnab Das**[2] | **Yashwant Joshi**[1]

[1]Department of Electronics and Telecommunication Engineering, Shri Guru Gobind Singhaji Institute of Engineering and Technology, Nanded, India

[2]Maritime Research Centre, Pune, India

Correspondence
Jyoti Sadalage, Department of Electronics and Telecommunication Engineering, Shri Guru Gobind Singhaji Institute of Engineering and Technology, Nanded, India.
Email: jyoti.rangole@vpkbiet.org

Abstract

The optimum performance of any underwater system depends on the propagation characteristics of the acoustic signals in the local water medium. Shallow tropical freshwater systems suffer from sub-optimal performance of sonar systems deployed for any acoustic sensing because of random fluctuations of the water medium. The propagation characteristics depend largely on the sound speed variations defined by the site-specific physical parameters such as water temperature, salinity and depth. The present study focuses on analysing the sound speed profile of a typical shallow freshwater system (Khadakwasla Lake; 18.43°N, 73.76°E), using regression models with the goal of deriving a computationally efficient model. To this end, a linear and polynomial regression model was developed, and their performance compared with the results of the model of Chen and Millero (1977), based on root mean square error (RMSE). In situ measurements of electrical conductivity, temperature and density (CTD) were carried out using a Valeport 602 CTD meter. Approximately 125 CTD samples were obtained during a 2-day experimental study conducted at Khadakwasla Lake from 11 October 2017 to 12 October 2017. The data collection was undertaken throughout the day at multiple locations in the lake over a spatial distance of ~16 km. The Valeport 602 CTD meter uses the Chen and Millero (1977) formula, considered the most conventional sound speed equation. The computational complexity of the proposed models was measured in terms of the number of addition and multiplication operations required. The validation of both models was carried out by varying the model input parameters within defined limits. The model inputs have been derived from an in situ experimental data collection process in a typical shallow tropical freshwater system. The linear regression model exhibited an RMSE of 4.15 m/s, while the polynomial regression model exhibited a good agreement with an RMSE of 0.5 m/s.

KEYWORDS
regression model, sound speed profile, tropical shallow water

1 | INTRODUCTION

The acoustic waves play an important role in varied underwater applications, including underwater communication, object detection and many underwater investigations (Leif, 2017). The characteristic of sound propagation underwater is governed by the physical properties of the underwater channel. The tropical regions experience seasonal variation in the heat gain/loss that impacts the thermal stratification, correspondingly manifesting in acoustic propagation (Bertram & Martin, 2008). The acoustic propagation in the tropical region witnesses a twofold challenge, while ensuring optimal sonar performance for any underwater application. First is the multipath fading, and second is seasonal and diurnal variations in temperature along with spatial salinity changes, which manifest as the modification of the sound speed profile (Etter, 2003).

Freshwater systems such as lakes act as dynamic response systems that integrate environment, climate and tectonic forces into a continuous, high-resolution archive of local and regional changes (Gierlowski Kordesch & Kelts, 2000). Understanding the physical properties of the channel and the acoustic propagation at the experimental site plays a vital role in the deployment of SONAR systems for an acoustic application (Etter, 2003). Such scientific information about a lake is important for improved understanding of the lake's physical dynamics for guiding the stakeholders in policy design. There are multiple empirical models in the literature for the sound speed calculation of fresh water, including Wilson (1960), Medwin (1975), Chen and Millero (1977), Coppens (1981) and Leroy, Robinson, and Goldsmith (2008). The commercially available CTD meters calculate sound speed by using any one of said formulae.

The present study explores the experimental site for its acoustic behaviour that would help researchers in sonar design and deployment in tropical shallow freshwater systems. Limited resources are available in the literature about the thermal gradient, salinity change and sound speed profile of Khadakwasla Lake, a driving force behind the present study experiment plan. This study presents the temperature, salinity and sound speed profile at different locations in Khadakwasla Lake. A linear regression model and polynomial regression model were proposed for sound speed, based on in situ CTD measurements during 2 days of the research. The model was based on a sample size of 60 samples and tested over more than 50 samples. Because model validation is a crucial step in any model building, the model performance was tested by allowing the model input parameters to vary within certain limits. This approach provided essential insights on model behaviour and its response to changes in model inputs (Emanuele & Elmar, 2016). Validation of the proposed models was carried out for a water temperature in the range of 10–40°C, salinity in the range of 0.01–0.5 ppt and pressure in the range of 0.1–1.5 bar, which represent the limits of the lake. The RMSE was used as a measure to validate the results of the present study and determine how close they were to our results and to see how close they are to the model of Chen and Millero (1977).

1.1 | Experimental Site

Khadakwasla Lake is one of the main reservoirs for the city of Pune, in Maharashtra state, India. The lake spans a length of 17 km and has a width of 1 km. The lake's total catchment area is about 501.80 km^2, with its depth varying 12 and 36 m (National Defence Academy, 2017). Pune has a hot semi-arid climate, bordering between a tropical wet and dry region (Weather in Pune, 2017). The tropical climate at the experimental site is characterized by seasonal rainfall from June to October. Based on literature, the lake's siltation rate is estimated to be 23.920 ha 100 km^{-2} yr^{-1} (http://shodhganga.inflibnet.ac.in). The experimental location was chosen as being able to analyse the acoustic complexity of a tropical shallow freshwater system.

2 | METHODOLOGY

2.1 | Sound speed measurement

The most important underwater variable in the present study is the speed of sound. Most commonly used methods of underwater sound speed measurement include direct measurement and CTD-based measurement. The CTD values were recorded by manually deploying the CTD sensor to the maximum depth in 1-m intervals, with measurements taken at seven different sites at different points in time. This approach assisted in the analysis of diurnal variations in the water temperature and sound speed.

The Valeport CTD 602 derived sound speed was determined using the empirical relationship reported by Chen and Millero (1977) empirical relation and the salinity using the method of Fofonoff and Millard (1983), (Valeport, Model 602 CTD Operation Manual, 2018). The Chen et al. equation incorporates 42 coefficients with 41 addition and 139 multiplication operations, thereby being computationally very intense.

2.2 | Regression analysis

2.2.1 | Multiple linear regression

The general form of multiple linear regression model used for sound speed calculation (Multiple Linear Regression Model, 2018) is as follows,

$$y=\beta_0+\beta_1 x_1+\beta_2 x_2+\beta_3 x_3\cdots+\beta_k x_k+\varepsilon \tag{1}$$

where β_0, β_1, β_2, … β_k = regression coefficients associated with independent variables x_1, x_2, … x_k, respectively; β_0 = intercept; and ε = random error component reflecting the difference between the observed and fitted linear relationship.

2.2.2 | Multivariate polynomial regression

The second-order multivariate polynomial regression analysis used to evaluate the sound speed at Khadakwasla Lake (Polynomial Regression Models, 2018) is as follows,

$$y=\beta_0+\beta_1 x_1+\beta_2 x_2+\beta_{11} x_1{}^2+\beta_{22} x_2{}^2+\beta_{12} x_1 x_2+\varepsilon \tag{2}$$

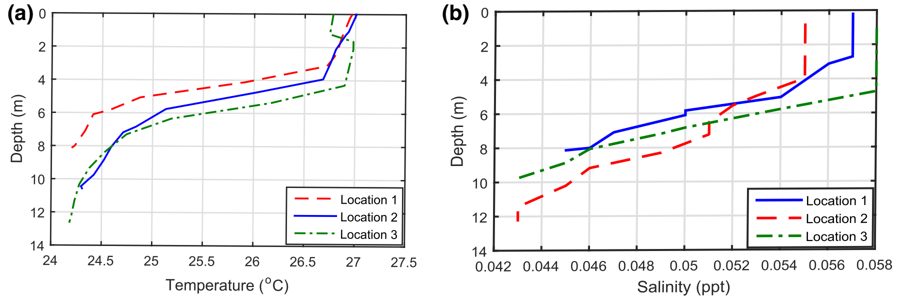

FIGURE 1 (a) Temperature profile at three different locations in Khadakwasla Lake. (b) Salinity profile at three different locations in Khadakwasla Lake

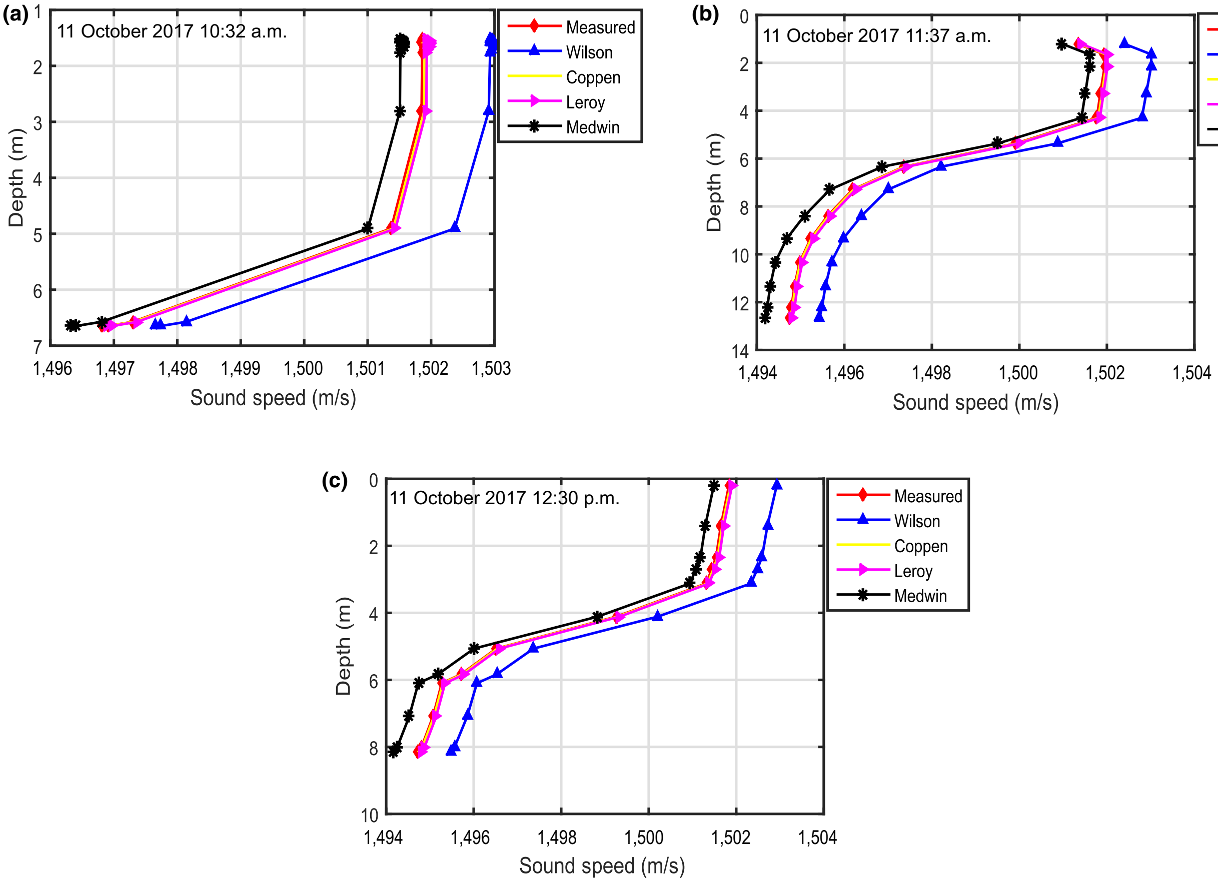

FIGURE 2 (a–c) Sound Speed Profiles for Khadakwasla Lake

where β_1, β_2, β_{11}, β_{22}, β_{12}, = regression coefficients associated with independent variables x_1 and x_2, respectively; β_0 = intercept; and ε = random error.

3 | RESULTS AND DISCUSSION

During the 2-day experiment, a Teledyne Odom Hydrographic Echotrac MK III Echo-sounder was deployed, operating at a frequency of 200 kHz for mapping the lake depth. The maximum depth over the planned survey area was ~15 m. The diurnal water temperature varied from ~24 to ~27°C (Figure 1a). Location 1, location 2 and location 3 are located at 18.43°N 73.76°E, 18.43°N 73.75°E and 18.41°N 73.73°E, respectively. The temperature variations were not significant to a depth of ~ 4 m, with this depth and higher being considered as the thermocline region (Figure 1a). The presence of the sonic layer at location 3 (Figure 1a) temperature profile could be observed from a depth of ~1 to ~1.7 m. The salinity at the three locations in Khadakwasla Lake is presented in Figure 1b. The experimental results indicated the average salinity of Khadakwasla Lake was 0.056 ppt. The sound speed values recorded at different time scale and location are presented (Figure 2a–c), along with the simulation results of the Wilson, Coppen, Leroy and Medwin models, in order to visualize the difference in the sound speed profile. The statistical analysis indicated a standard deviation of 0.66, 0.046, 0.078 and 0.28 m/s between the Chen Millero and Wilson, Coppen, Leroy, and Medwin models. The present study discusses the mathematical models developed for evaluating sound speed in shallow tropical freshwater systems. The proposed multiple linear regression model and multivariate polynomial regression model, respectively, are indicated as follows:

$$C(S,T,P) = 1431.46 + 2.609624 * T + 0.10489 * P + 1.307285 * S \quad (3)$$

$$C(T,P) = 1409 + 4.321 * T - 1.04 * P - 0.03285 * T^2 + 0.05144 * T * P \quad (4)$$

where C = sound speed (m/s); S = salinity (ppt); T = water temperature (°C$^{)}$; and P = pressure (bar). The multiple linear regression model is based on three independent variables, namely, salinity, water temperature, and pressure. The multivariate polynomial regression model expressed in Equation (4) is based on two independent variables, namely, water temperature and pressure. The regression coefficients β_0, β_1, β_2, … β_k for multiple regression and polynomial regression model could be easily obtained by comparing Equation (1) with (3), and Equation (2) with (4), respectively. The model performance is evaluated based on two measures, namely, RMSE and computational complexity. The present study also discusses the measures used for determining goodness of fit of the proposed models, using the coefficient of determination (R^2), Gaussianity of the residual plot and model validation.

The linear regression model and the polynomial regression model both provided an R^2 value of 0.98, based on Equation (5) (Regression Methods, 2018), as follows:

$$R^2 = 1 - \frac{SS_{res}}{SS_{tot}} \quad (5)$$

where SS_{res} = residual sum of squares and SS_{tot}= total sum of squares (variance of the data), as follows:

$$SS_{res} = \sum_i (y_i - f_i)^2 \quad (6)$$

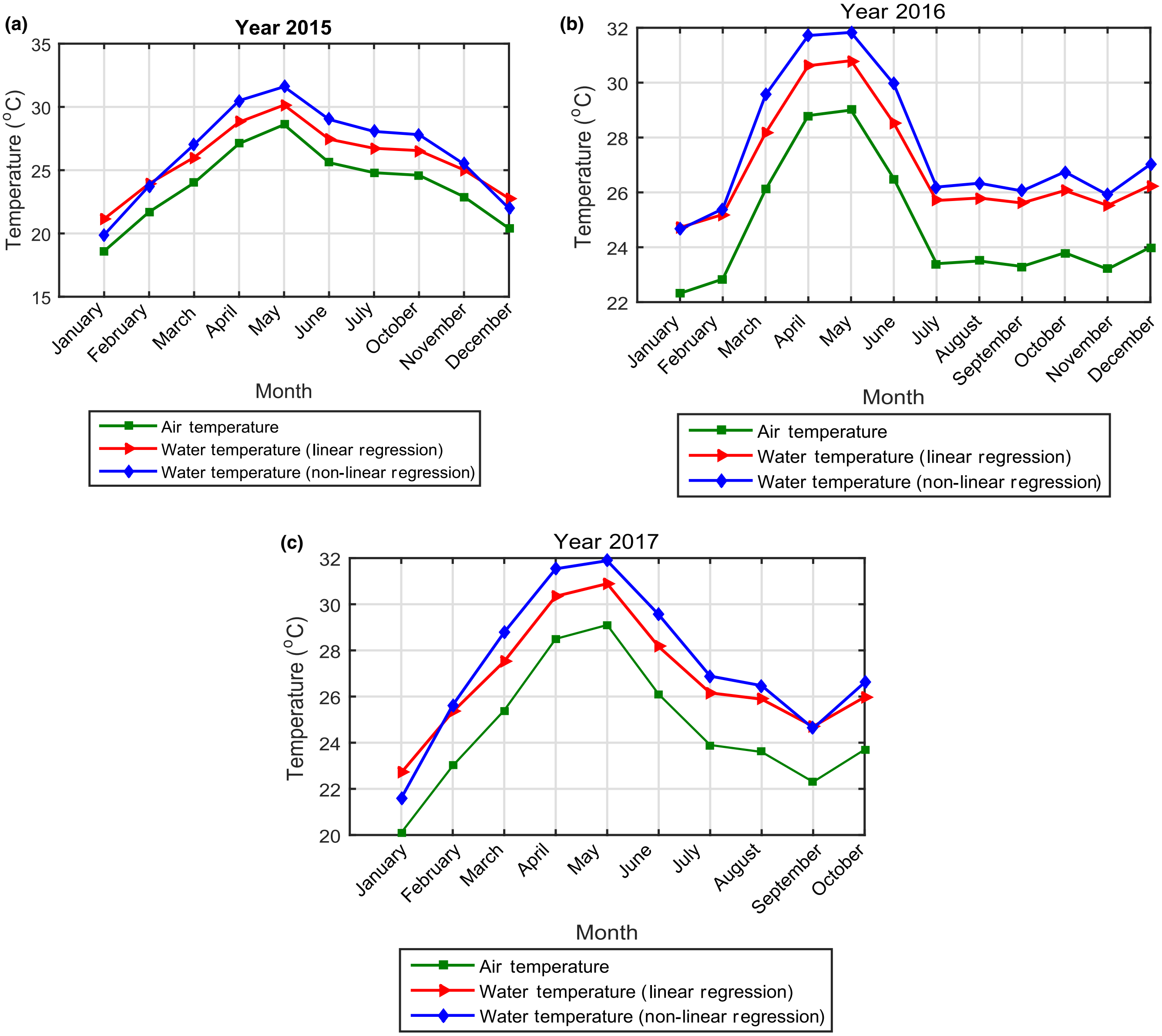

FIGURE 3 (a–c) Month-wise Average Air and Water Surface Temperatures for Khadakwasla Lake for the year 2015, 2016 and 2017

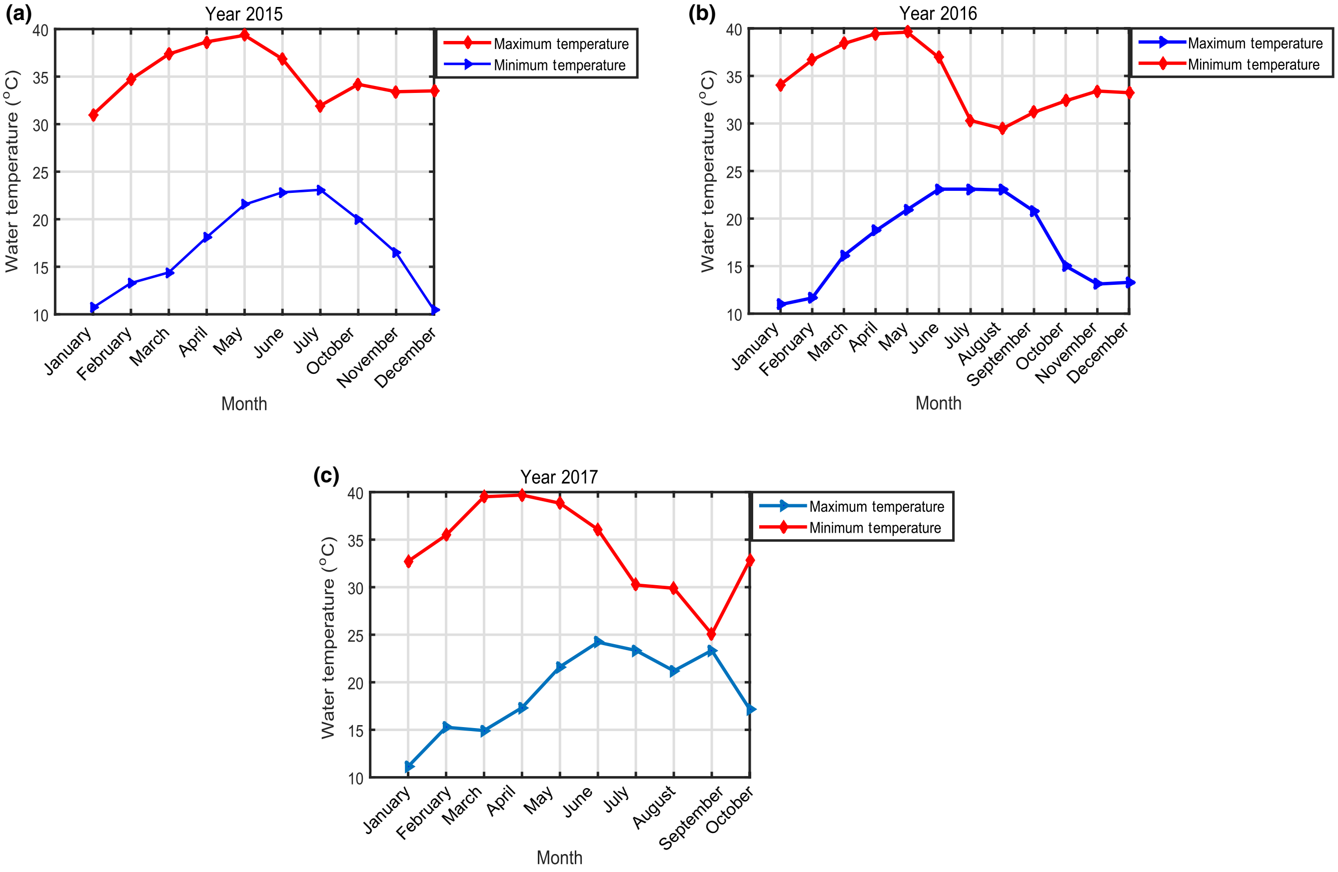

FIGURE 4 (a–c) Month-wise Minimum and Maximum Water Surface Temperatures for Khadakwasla Lake for the year 2015, 2016 and 2017

$$SS_{tot} = \sum_{i} (y_i - \bar{y})^2 \tag{7}$$

where y_i = measured sound speed; f_i = predicted or modelled sound speed; and y = mean sound speed. The confidence interval for the regression coefficients in Equations (3) and (4) is 95%. A hypothesis test was also undertaken to determine the significance of the regression coefficients. The null hypothesis (H_0) was that *the regression coefficients are zero or insignificant*, while the alternative hypothesis (H_A) was that the *regression coefficients are non-zero or significant*. The *p* value of each regression coefficient also was analysed, testing the null hypothesis that the coefficient is equal to zero, with a low *p* value (<0.05) indicating the null hypothesis can be rejected. To this end, the *p* value of each regression coefficient was <0.001, justifying rejection of the null hypothesis. The Gaussianity of the residual plot of the linear regression model also was tested with the Kolmogorov–Smirnov test, which indicated the residual plot had the normal distribution, allowing confidence in the model results.

To validate both the models, it is necessary to decide the range over which the independent variables are allowed to vary. Because the water temperature data for Khadakwasla Lake are not readily available, the relation presented by Ali (2013) was used to predict the surface water temperature from the available air temperature data. Air temperature data were taken at National Defence Academy station of the India Meteorological Department in Pune, which is situated in the vicinity of Khadakwasla Lake. The database contains minimum, maximum and average air temperature values, recorded at hourly resolution from the years 2015 to 2017. The month-wise average water surface temperature values were derived from the air temperature, using the linear and the nonlinear regression model proposed by Ali (2013). Analysis of month-wise average water surface temperature for the years 2015 to 2017 indicated that water surface temperature varied from 20 to 32°C (Figure 3a- c). The air temperature values for the months of August and September, 2015, were not available (Figure 3a). The daily water temperature variations were analysed to test the model robustness, considering the minimum and maximum water temperature values for 2015, 2016 and 2017, noting the minimum water temperature is ~10°C and the maximum water temperature is ~ 40°C (Figure 4a–c).

The simulation results (Figure 5a–c) highlight the sound speed using the Chen and Millero (1977) model, along with the proposed linear and polynomial regression models with temperature in the range from 10 to 40°C, and pressure in the range of 0.1–1.5 bar, with an average salinity of 0.056 ppt. The RMSE value of the multiple linear regression model over the said range of the input variables was 4.5 m/s. Analysis of the second-order polynomial model

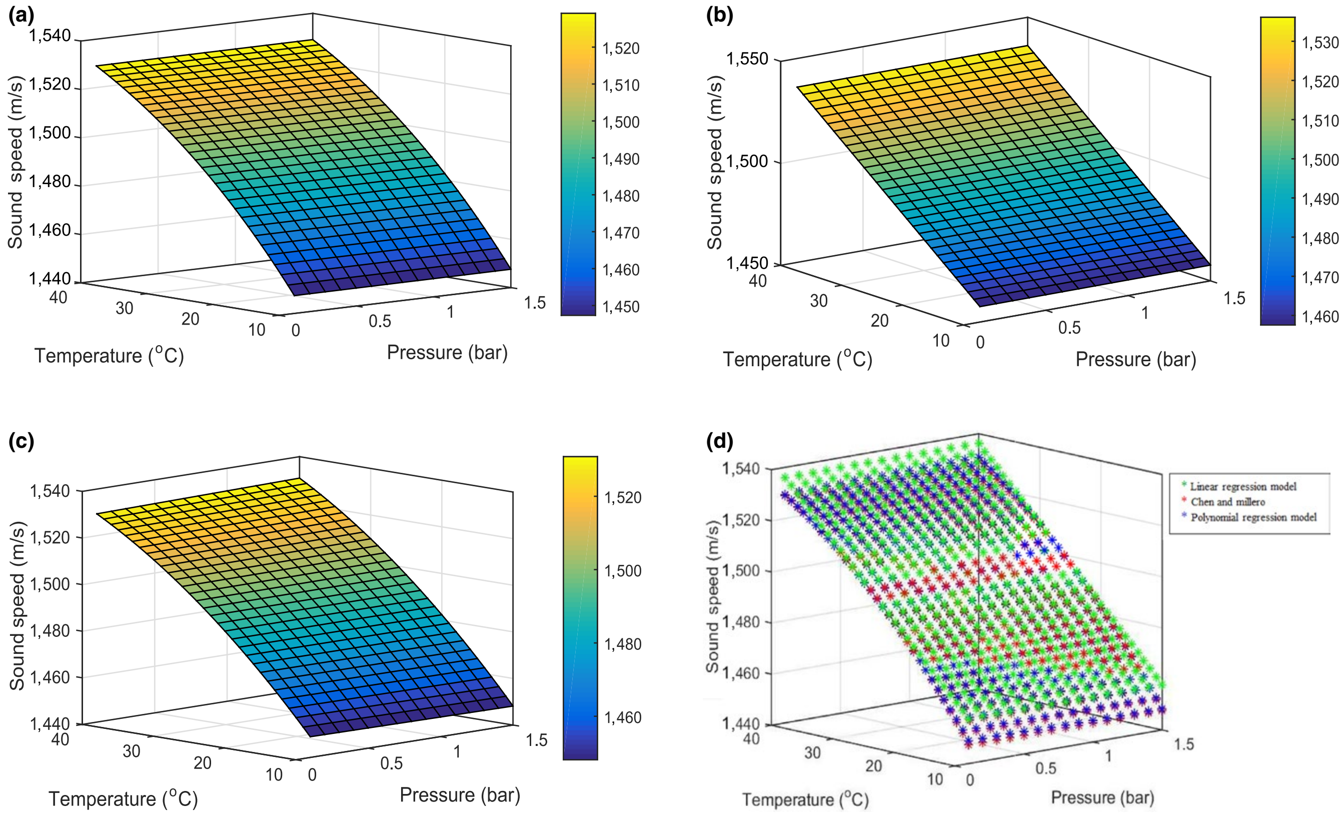

FIGURE 5 Sound Speed at Salinity 0.056 ppt using (a) Chen and Millero; (b) Proposed Linear Regression Model; (c) Proposed Polynomial Regression Model; and (d) Scatter Plot

presented a good agreement with an RMSE value of 0.5 m/s. The comparative analysis of the Chen and Millero (1977) model and the proposed models could be easily visualized from a 3-dimensional scatter plot (Figure 5d). The linear regression model was found to be comparable to the Chen and Millero (1977) model only in the temperature range 20 to 30°C. Outside this temperature range, however, there was a deviation of ~4 m/s. The polynomial model exhibited comparable results to the Chen and Millero (1977) model over the temperature range from 10 to 40°C, indicating it could be used over the extended pressure value up to 3.5 bar over the temperature range of 10–32°C.

4 | CONCLUSION

The site-specific seasonal and diurnal variability in the acoustic behaviour in shallow tropical freshwater systems is a challenge for any acoustic survey effort. The results of the present study represent a rare on-site experimental validation of the medium parameters and their manifestation as the sound speed profile. Insight regarding the acoustic behaviour at Khadakwasla Lake is a critical input to understanding shallow tropical freshwater system behaviour. The key contribution of the present study was to present a complete framework of environmental changes at the experimental site and their effects on the acoustic field, along with the computationally efficient mathematical formulation of sound speed for such tropical systems. The proposed multiple linear regression model requires three addition and three multiplication operations. The second-order polynomial regression model provided a computationally efficient solution for sound speed with only four addition and six multiplication operations over the specified bounds of the model input parameters. Validation of the polynomial regression model also highlighted comparable results with an RMSE of 0.5 m/s with the Chen and Millero (1977) model.

ORCID

Jyoti Sadalage https://orcid.org/0000-0002-3164-2642

REFERENCES

Ali, S. (2013). Estimation of surface water temperature in small recharge pond from air temperature. *Indian Journal of Soil Conservation*, *41*(1), 1–7.

Bertram, B., & Martin, S. (2008). Stratification of lakes. *Reviews of Geophysics*, *46*, RG2005.

Chen, C., & Millero, F. (1977). Speed of sound in seawater at high pressures. *Journal of Acoustic Society of America*, *62*(5), 1129–1135. https://doi.org/10.1121/1.381646

Coppens, A. (1981). Simple equations for the speed of sound in Neptunian waters. *Journal of Acoustic Society of America*, *69*(3), 862–863. https://doi.org/10.1121/1.385486

Emanuele, B., & Elmar, P. (2016). Sensitivity analysis: A review of recent advances. *European Journal of Operational Research*, *248*(3), 869–887.

Etter, P. (2003). *Underwater Acoustic Modeling and Simulation*, 3rd ed.. Taylor and Francis Group: CRC Press. https://doi.org/10.4324/9780203417652

Fofonoff, P., & Millard, R. (1983). Algorithms for computation of fundamental properties of seawater. *UNESCO Technical Papers in Marine Science*, *44*, 53.

Gierlowski Kordesch, E. H., & Kelts, K. (2000). *Lake basins through space and time AAPG Studies in Geology 46* (p. 648). Tulsa, OK: American Association of Petroleum Geologists. http://shodhganga.inflibnet.ac.in/bitstream/10603/99856/12/12chapter1.pdf.

Leif, B. (2017). *Applied Underwater Acoustics*, *Thomas Neighbors*, *David Bradley*. Amsterdam, The Netherlands: Elsevier.

Leroy, C., Robinson, S., & Goldsmith, M. (2008). A new equation for the accurate calculation of sound speed in all oceans. *Journal of Acoustic Society of America*, *124*(5), 2774–2782. https://doi.org/10.1121/1.2988296

Medwin, H. (1975). Speed of sound in water: A simple equation for realistic parameters. *Journal of Acoustical Society of America*, *58*(6), 1318–1319. https://doi.org/10.1121/1.380790

Multiple Linear Regression Model, 2018. Retrieved from http://home.iitk.ac.in/~shalab/regression/Chapter3-Regression-multipleLinearRegressionModel.pdf

National Defence Academy, 2017. Retrieved from http://www.nda.nic.in

Polynomial Regression Models, 2018. Retrieved from http://home.iitk.ac.in/~shalab/regression/Chapter12-Regression-PolynomialRegression.pdf.

Regression Methods, 2018. Retrieved from https://onlinecourses.science.psu.edu/stat501/node/255

Valeport, Model 602 CTD, operational manual, 2018. Retrieved from https://www.Valeport.co.uk/Portals/0/docs/Manuals/CTD%20&%20Multiparameter/Model%20602/0602804A.pdf

Weather in Pune, 2017. Retrieved from http://mutha.ncra.tifr.res.in/ncra/ncra1/asi-2015/weather-in-pune

Wilson, W. (1960). Equation for the speed of sound in sea water. *Journal of the Acoustical Society of America*, *32*(10), 1357. https://doi.org/10.1121/1.1907913

How to cite this article: Sadalage J, Das A, Joshi Y. A computationally efficient model for determining sound speed in shallow tropical freshwater systems, with field validation. *Lakes & Reserv*. 2019;00:1–7. https://doi.org/10.1111/lre.12255