Applied Acoustics 181 (2021) 108167

Contents lists available at ScienceDirect

Applied Acoustics

journal homepage: www.elsevier.com/locate/apacoust

Passive ocean acoustic thermometry with machine learning ☆

Fenghua Li [a,*], Kai Wang [a,b], Xishan Yang [a], Bo Zhang [a], Yanjun Zhang [a]

[a] State Key Laboratory of Acoustics, Institute of Acoustics, Chinese Academy of Sciences, No.21 North 4th Ring Road, Haidian District, Beijing 100190, PR China
[b] Physics Department, University of Chinese Academy of Sciences, No.19(A) Yuquan Road, Shijingshan District, Beijing 100049, PR China

ARTICLE INFO

Article history:
Received 26 May 2020
Received in revised form 25 April 2021
Accepted 4 May 2021

Keywords:
Passive ocean acoustic thermometry
Noise cross-correlation function
Random Forest

ABSTRACT

Passive ocean acoustic thermometry (POAT) needs long accumulation time to achieve high accuracy. This article provides a machine learning-based method, Random Forest, to obtain the averaged sound speed (AVSS). With supervised learning, the AVSS can be estimated from half an hour accumulated noise cross-correlation functions (NCFs). Based on the feature importance analysis, an empirical equation is proposed to briefly describe the relationships between the features. The results of estimations are compared among different methods to demonstrate the advantage of the machine learning-based approach.

1. Introduction

Ocean acoustic tomography (OAT) has been attracted much attention over decades [1,2]. Similar to the method of seismology, classical OAT inverts the ocean interior changes in sound speed by measuring sound travel-time perturbations between active acoustic sources and receivers [3,4]. But the complexity of the experiment greatly limits its application.

Passive ocean acoustic thermometry (POAT) provides an alternative remote sensing modality for detecting the ocean temperature based on the long time accumulated noise cross-correlation functions (NCFs) between two separated hydrophones [5–7]. With this method, Godin et al. [7] achieved the inversion of sound speed profile in the deep ocean and Woolfe et al. [8] successfully estimated the deep ocean temperature from one week accumulated NCFs.

The desired long accumulation time [9,10] in acquiring NCF is one of the main causes which limits the POAT's application to the fast-changing ocean. The acquisition of highly accurate ocean sound speed estimation with a short accumulation time is the goal for which always seeking [11,12]. Since only the noise sources in the end-fire direction between two receivers contribute to the NCFs [5,13], the beamforming technique is implemented on a vertical array [14,15] and on two separated arrays [16–19] to improve the convergence of the NCFs. By applying array signal processing, Li et al. [20] reduced the NCF accumulation time to 2 h.

The machine learning approaches may discover specific trends that would not be apparent and achieve more accurate predictions [21,22]. This motivates us to investigate the machine learning-based POAT with a short accumulation time.

The paper is organized as follows. The regular averaged sound speed (AVSS) estimation method is given in Section 2.1. The theory of the Random Forest (RF) and the detailed description of the machine learning-based method are discussed in Sections 2.2 and 2.3. The experimental results in Section 3 demonstrate the performance of the RF, the empirical equation, and the regular methods. In Section 4, the effect of features is analyzed and the physical explanation of the results is also presented. The conclusion is given in Section 5.

2. Method

2.1. Regular AVSS estimation from peak arrival time

Fig. 1(a) shows the sketch of POAT with two parallel horizontal arrays, a vertical chain consisting of 38 Temperature Depth sensors (TDs) is placed between the two arrays. The same beamforming process in the reference [20] is used here. The acquired data of each array are divided into K non-overlapping snapshots. Each snapshot data is converted into frequency bins by fast Fourier transform. The conventional beamforming output for the k^{th} snapshot in θ_0 direction can be written as

$$\mathbf{B}_i^k(\omega,\theta_0) = \mathbf{W}_i^{\dagger}(\omega,\theta_0)\mathbf{P}_i^k(\omega), \tag{1}$$

☆ This research was supported by the National Natural Science Foundation of China (Grant Nos. 11974017, 11874061) and the Youth Innovation Promotion Association CAS (Grant No. 2017028).
* Corresponding author.
E-mail address: lfh@mail.ioa.ac.cn (F. Li).

https://doi.org/10.1016/j.apacoust.2021.108167
0003-682X/

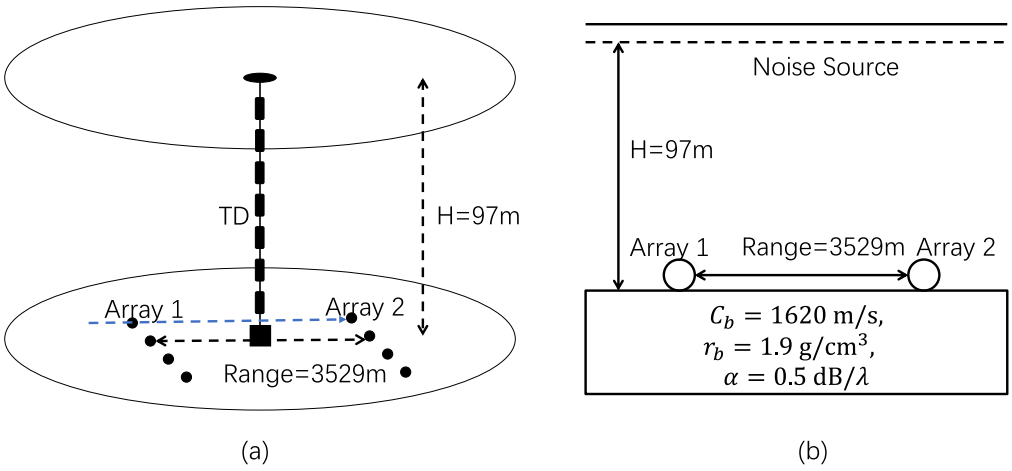

Fig. 1. (a) Sketch of the POAT with two parallel horizontal arrays. (b) Simulation fluid model with parameters.

where † denotes the Hermitian Conjugate of matrix, $\mathbf{P}_i^k(\omega) = [p_i^1(\omega), p_i^2(\omega), \ldots p_i^N(\omega)], i = 1,2$ denotes the frequency domain signal of the i^{th} array at the k^{th} snapshot, N denotes the number of array elements, and $\mathbf{W}_i(\omega,\theta_0) = [w_i^1(\omega,\theta_0), \ldots, w_i^N(\omega,\theta_0)], i = 1,2$ denotes the corresponding weighted vector, where $w_i^j(\omega,\theta_0) = \exp(-\mathrm{i}\omega\tau_j(\theta_0)), \theta_0$ denotes the end-fire direction of the first hydrophone of array 1 and the first hydrophone of array 2.

The cross-correlation function between two arrays in θ_0 direction is given by

$$\begin{aligned}\mathbf{C}^k(\omega) &= \mathbf{B}_1^k(\omega,\theta_0)\mathbf{B}_2^{k\dagger}(\omega,\theta_0) \\ &= \mathbf{W}_1^{\dagger}(\omega,\theta_0)\mathbf{P}_1^k(\omega)\mathbf{P}_2^{k\dagger}(\omega)\mathbf{W}_2(\omega,\theta_0).\end{aligned} \tag{2}$$

With K snapshots accumulation, we can obtain the ensemble average time-domain NCF as

$$\mathbf{C}(\tau) = \int_{-\infty}^{\infty}\left(\sum_{k=1}^{k=K}\mathbf{C}^k(\omega)\right)\exp(\mathrm{i}\omega\tau)d\omega. \tag{3}$$

The Empirical Green Function (EGF) can be extracted from the time derivative of the ensemble average time-domain NCF [23,24] as

$$\mathbf{G}(\tau) \approx -\frac{\mathrm{d}\mathbf{C}(\tau)}{\mathrm{d}\tau}. \tag{4}$$

In realistic multi-path channel, the sound arrives at different times because of the different path distances. The AVSS can be estimated as the arithmetic mean of the travel distance divided by the arrival time of the peak obtained from the EGF corresponding the i^{th} path [25]:

$$\hat{c} = \sum_i^N \frac{R_i}{\tau_i^p}. \tag{5}$$

The AVSS can also be estimated as the inverse-variance weighted average as

$$\hat{c} = \frac{1}{\sum_{i=1}^N \sigma_i^{-2}}\sum_{i=1}^N\left(\frac{R_i}{\tau_i^p} * \sigma_i^{-2}\right). \tag{6}$$

where R_i is the i^{th} path travels distance between the two arrays, τ_i^p is the corresponding peak arrival time, N is the number of the multi-paths and σ_i denotes the variance of the i^{th} peak arrival time. There is no big difference between sound speeds estimated from Eqs. (5) and (6), the inverse-variance weighted estimation's accuracy is slightly better than that of the arithmetic average. In the following text, the inverse-variance weighted estimation is used as the regular method in comparison to the machine learning based method.

The i^{th} path is estimated by the Bellhop model. And the parameters is shown in Fig. 1(b). The density of sediment material r_b and attenuation α are chosen based on the type of the adjacent seafloor. And the sediment compressional speed c_b is chosen from the reference [20].

2.2. Random forest

Formally, The RF regressor [26,27] trains several regression decision trees (RDTs) from several input variables to predict the value of a target.

The basic RDT is built sequentially by creating nodes that partition the data. The construction starts from the root node, each node can be split into two child nodes on the feature that results in the smaller cost function, then this splitting procedure is repeated at each child node until the leaves pure (all samples of leaves contain the same target).

Fig. 2 shows a simple RDT example. All input samples $X_n = [x_n^1, \ldots, x_n^i, \ldots, x_n^M], n \in [1, N]$, where N denotes the number of all training samples and M denotes the number of all features, at t^{th} node all belong to the same class as

$$\begin{aligned} X_n &\in c_1 \quad if \quad x_n^i > d^t, \\ X_n &\in c_2 \quad if \quad x_n^i \leqslant d^t. \end{aligned} \tag{7}$$

Here c_j denotes the j^{th} child node of the t^{th} node, d^t is the associated threshold of the t^{th} node that minimizes the cost function $\mathscr{L}$ as

$$\begin{aligned} d^t &= \mathrm{argmin} \quad \mathscr{L}\left(d^t\right), \\ \mathscr{L}(d) &= \sum_{j=1}^{N_{\mathrm{split}}}\frac{n_{c_j}}{n_p}H(c_j), \end{aligned} \tag{8}$$

where n_p is the total number of samples at the parent node, n_{c_j} is the number of samples at the j^{th} child node, N_{split} denotes the number of child nodes and $H(\cdot)$ is the impurity function as

$$H(c_j) = \frac{1}{n_{c_j}}\sum_{X_n \in c_j}\left(\hat{\mathbf{y}}_{c_j} - y^{(n)}\right)^2 \tag{9}$$

$$\hat{\mathbf{y}}_{c_j} = \frac{1}{n_{c_j}}\sum_{X_n \in c_j} y^{(n)}. \tag{10}$$

Here $y^{(n)}$ is the true target value of sample X_n, and $\hat{\mathbf{y}}_{c_j}$ is the predicted target value of all samples at node c_j.

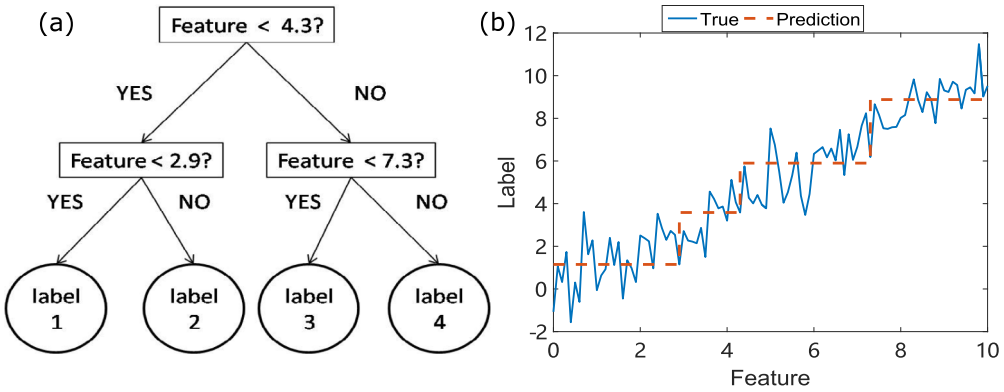

Fig. 2. Regression tree example. (a) is the visualization of a regression tree. (b) is the result of regression. Here, we have 100 samples ($N = 100$) divided into four labels with one feature ($N_f = 1$), d^t here equals to 2.9, 4.3, 7.3, respectively.

In practice, the RF algorithm firstly draws n_{tree} bootstrap samples from the original input, grows an unpruned RDT for each of the bootstrap samples [28], finally predicts new data by aggregating the predictions of the n_{tree} trees [29].

And the root mean squared error (RMSE) over N samples is used to quantify the prediction performance of the AVSS estimation as

$$\mathrm{RMSE} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(C_{p_i} - C_{t_i}\right)^2}, \tag{11}$$

where C_{p_i} and C_{t_i} denotes the predicted AVSS and the labelled truth AVSS respectively.

2.3. Step-by-step description of the method

We give a step-by-step description of the method here.

1. Data processing. The recorded signals were divided into segments with an accumulation time of T. For each segment data, we further divide the data into several snapshots with a length of 10s and apply a 20–400 Hz band-pass filter [30,20] to the data. Following Eqs. (1)–(4), we extract the EGFs from measured noises. The reference sound speed of beamforming used in this work is 1526 m/s. It is chosen as the water sound speed close to the bottom nearby.

2. Defining the features. One hundred and ninety-six features extracted from $\mathbf{G}(\tau)$'s both in the time and frequency domain are used. In the time domain, the extracted EGF is divided into 6 segments, including 4 envelopes, the part before 1st envelope, and the part after 4th envelope. For each segment, the peak arrival time, trough arrival time, maximum, minimum, mean, variance, skewness, and kurtosis are used. In the frequency domain, the extracted EGF is divided into 19 frequency bands (each frequency band corresponding to 20 Hz). For each frequency band, the maximum, minimum, mean, variance, skewness, kurtosis, and entropy are used. From the entire EGF, We extract the peak arrival time, the trough arrival time, maximum, minimum, mean, variance, skewness, and kurtosis in time domain again. And in the frequency domain, we extract the maximum, minimum, mean, variance, skewness, kurtosis, and entropy from the entire EGF.

3. Preparing the labels. To compensate for the lack of measured data, the AVSS inverted from the 2-h accumulated NCF with enough high SNR is interpolated as the ground truth, which is verified with the measured sound speed in reference [20].

4. Train-test split. In the experiment, we collected 1298 h of acoustic data. The first contiguous 780 h of the data are selected for training and the left contiguous 518 h are for testing.

5. Training and tuning the hyper-parameters of the model. With the help of Scikit-learn [31], the grid search and k-fold cross-validation are used to search for the optimal parameters. Dividing the training set into k folds, a set of hyper-parameters is used to model on every $k-1$ folds and evaluated for its performance on the rest fold. Therefore the hyper-parameters that maximize the performance of models will finally be determined without overfitting.

6. Prediction and assessment. The model parameters trained in step 5 are used to predict the AVSS on the testing data and the prediction error is reported by the RMSE.

3. Experiment result

3.1. EGF extraction and AVSS prediction

To testify the validity, an experiment was conducted in the South China Sea. Two self-recorded horizontal arrays that consist of 15 elements are deployed on the bottom in 97-m depth water. The distance between these two horizontal arrays is 3.529 km. Both arrays are about 84 m long. In our experiment, these two arrays (array 1 and array 2) are physically connected by one cable. The noise signals are recorded at a sampling frequency of 5 kHz synchronously. The EGF with different accumulation time T can be obtained from Eq. (4). With the beamforming technique, only the signals from the pointing direction are retained, and the signals from other directions are filtered out [17,20]. Different reference sound speeds do not affect the extracted EGFs [20]. And the perturbation of RMSE is negligible (less than 0.02 m/s) by choosing different reference sound speeds.

The negative time delay part of NCF can be extracted by changing the direction to the opposite. Fig. 3 shows the extracted EGFs from two opposite directions with an accumulation time of 2 h.

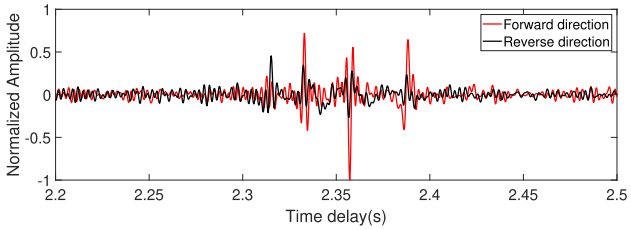

Fig. 3. EGFs with an accumulation time of 2 h. The red line is the extracted EGF from forward direction, the black line represents the extracted EGF from reverse direction. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

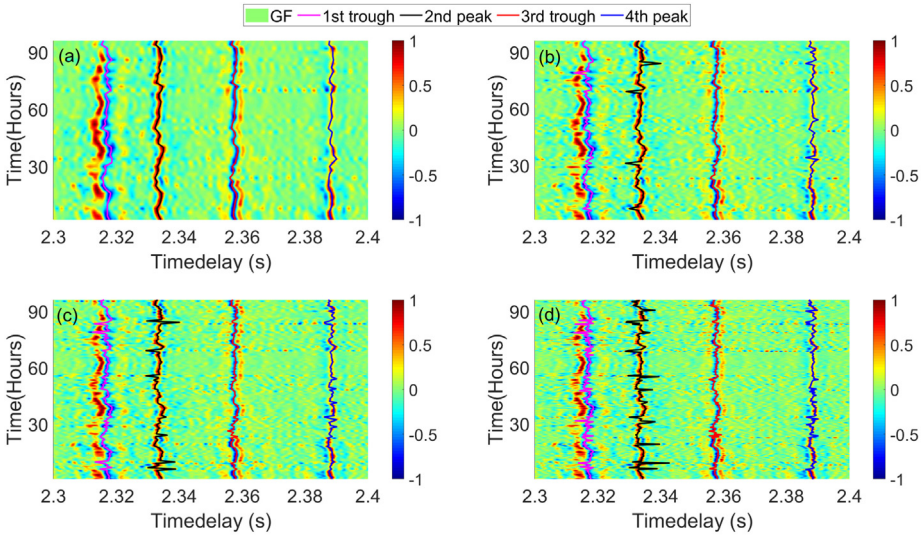

Fig. 4. EGFs for a total time of 96 h with an accumulation time of (a) 2 h, (b) 1 h, (c) 45 min and (d) 30 min. The pink line represents the first trough, black line represents the second peak, the red line represents the third trough and the blue line represents the fourth peak. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

The red line is the extracted EGF from the forward direction and the black line represents the extracted EGF from the reverse direction. It can been seen from Fig. 3 that EGFs from the reverse direction have inferior SNR because of the inhomogeneous ambient noise distribution in the real environment. Fig. 4 shows EGFs for 96 consecutive hours with different accumulation times T. Fig. 5 shows a typical EGF extracted from measured noise vs. different accumulation times T. It can be seen from the figure that SNR increases as the accumulation time increases.

From the extracted EGF, most abovementioned features are obtained in different time segments and frequency bands to ensure cleanliness and effectiveness. Higher-order moments of the acoustic data, as well as the variance, skewness, and kurtosis, are used to capture the evolution of the signal's energy.

The estimations of the AVSS for three accumulation times (1 h, 45 min, 30 min) are shown in Fig. 6. Fig. 7 shows the comparison of corresponding RMSEs between different estimation methods. The blue line shows the RMSEs for the regular method and the purple line shows the RMSEs for the RF. It can be seen from the figure that the AVSS can be inverted accurately with the RF as the accumulation time decreases to 30 min, while the regular method has a relatively large RSME.

3.2. Feature importance and Empirical equation

With the feature importance technology, Fig. 8 gives the plot to show the feature importance of the input variables. It can be

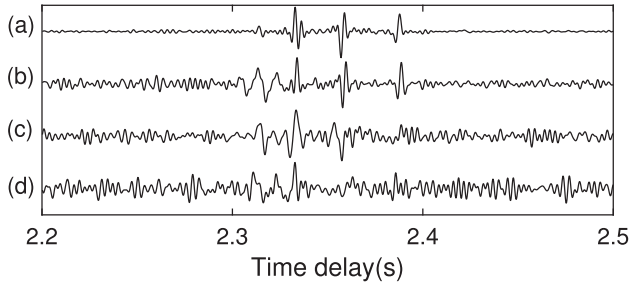

Fig. 5. an typical EGF with an accumulation time of (a) 2 h, (b) 1 h, (c) 45 min and (d) 30 min.

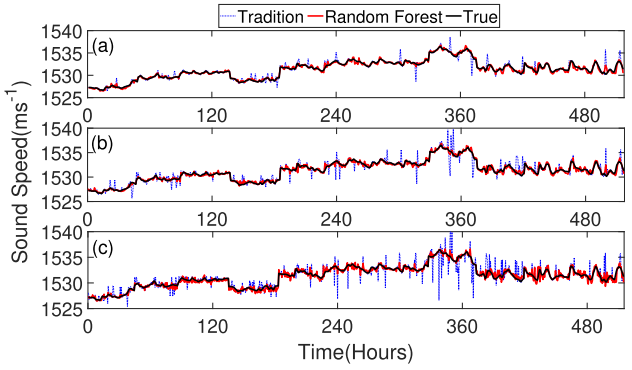

Fig. 6. AVSS predictions for a total time of 518 h with an accumulation time of (a) 1 h, (b) 45 min and (c) 30 min. The blue dot line represents the regular method results, the red line represents the inverted AVSS from the RF, and the black line represents the true AVSS. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

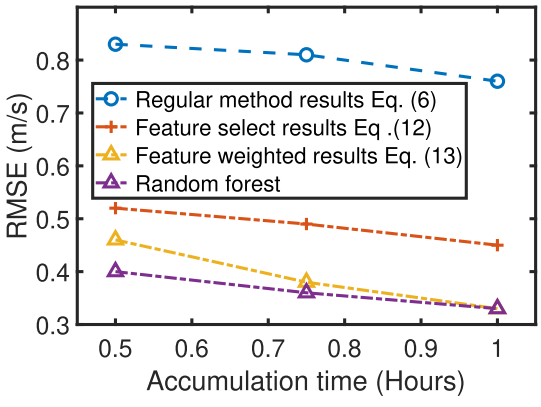

Fig. 7. The RMSEs for different methods with different accumulation times. The blue line represents the RMSEs for the regular method, the red line represents the RMSEs for the feature selected method, the yellow line represents the RMSEs for feature weighted method and the purple line represents the RMSEs for the RF. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

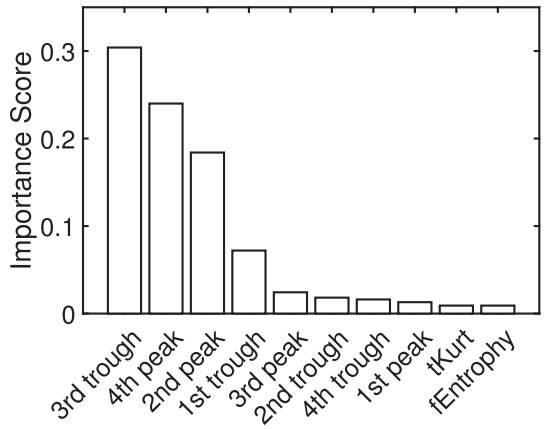

Fig. 8. shows the 10 most important features in the RF. Here tKurt denotes the kurtosis in the time domain.

observed that the most important feature is the arrival time of the 3rd trough, followed by the arrival time of the 4th peak, the arrival time of 2nd peak, the arrival time of 1st trough, and so on. None of the frequency-domain features is dominated. By averaging the feature importance between different accumulation times and normalizing proportions, the relative weights of evaluation for the four most important features are 38%, 30%, 23%, and 9%, respectively.

Feature selection plays an important role in improving the inversion accuracy. If we select the 4 most important features, the corresponding sound speed can be estimated as

$$\widehat{C} = \frac{1}{\sum_{i=1}^{N}\sigma_i^{-2}}\left(\sigma_1^{-2}\frac{R_1}{\tau_1^t} + \sigma_2^{-2}\frac{R_2}{\tau_2^p} + \sigma_3^{-2}\frac{R_3}{\tau_3^t} + \sigma_4^{-2}\frac{R_4}{\tau_4^p}\right), \tag{12}$$

which is called feature selected inverse-variance weighted estimation. τ_i^t in Eq. (12) is the corresponding trough arrival time. The inversion results from each feature are calibrated against the labels to ensure R_i is corrected. If we further multiply $\tau_i^{t,p}$ by the weight value proportional to the feature importance, an empirical equation, which is called feature weighted estimation, can be obtained to estimate the AVSS as

$$\widehat{C} = 0.09\frac{R_1}{\tau_1^t} + 0.23\frac{R_2}{\tau_2^p} + 0.38\frac{R_3}{\tau_3^t} + 0.3\frac{R_4}{\tau_4^p} \tag{13}$$

For brevity, the results by Eqs. (12)–(13) are also shown in Fig. 7. It can be seen that the RMSEs are reduced significantly compared with those for the regular method.

4. Discussion

From Fig. 7, the RF has the smallest RMSE in comparison with other estimations when the accumulation time is reduced from 2 h to 30 min. The RF performs better than other methods because of the three reasons listed in Secs. 4.1–4.3. With the help of the RF, we develop the empirical equation Eq. (13) which can estimate the AVSS more accurately with a shorter accumulation time.

4.1. Feature selection

Machine learning recognizes the importance of features. The use of the RF allows us to report the most important features in our model and therefore enables us to gain a physical understanding of the problem - besides higher performance, which is one of the primary reasons why we choose to rely on these particular models.

Fig. 8 reports the importance scores of ten most important feature and Fig. 4 also shows the four most important features on the EGF. The four most features are different from traditional experience. We generally only consider the peak of the arrival signal, not the trough of the arrival signal. However, the arrival times of the first and third troughs are more important than the corresponding peak arrival time. Physically, this can be explained as the reversed-phase of the surface reflection. By selecting more suitable features, the AVSS can be inverted more accurately.

It can been seen from Fig. 8 that other features like the high-order moments or the frequency-domain features are not important. That implies the energy fluctuation is not sensitive to variations in the sound speed of water column.

4.2. Feature importance weights

Machine learning provides the feature importance weights. Feature importance score represents the number of times a given feature was selected at a tree node by the algorithm. On most nodes of the tree, the results are delineated by important features. Thus, the good or bad results of the model will depend largely on the weight of these good features in the model. Although it's not rigorous using this weighting method, the sound speed can be inverted more accurately by the feature weighted average than the regular average in Fig. 7.

In practice, the predictions of the arrival times tend to be biased relative to the true AVSS. Eq. (12) is based on the inverse-variance weighted average, the resulting weighted average will be biased if some of the arrival times are biased. As shown in Fig. 9, the predictions of the third trough and fourth peak have significantly fewer

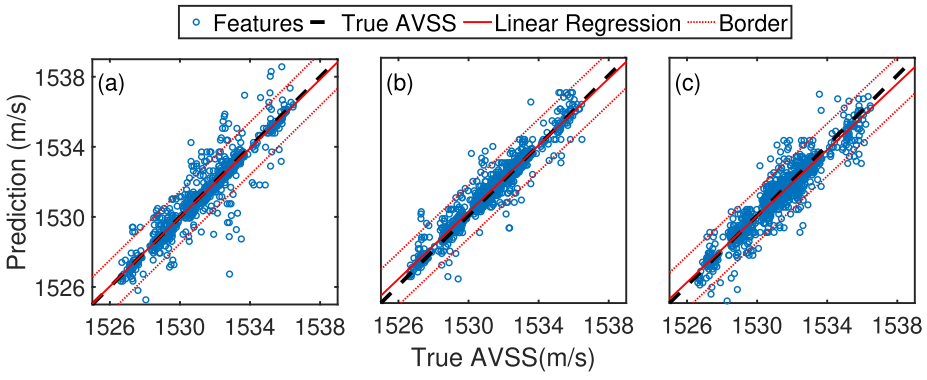

Fig. 9. With an accumulation time of 30 min, inversion results of (a) 2nd peak, (b) 3rd trough and (c) 4th peak. Black Line represents the true AVSS, red line represents the linear regression of the predictions, and red dot line the upper and lower boundaries (linear regression +/−1.5 m/s) that we use to determine outliers. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

outliers than the prediction of the second peak, but their magnitudes of the slope of linear regression are more deviant from the true AVSS.

It is difficult to assess the risk of bias without actual measurements, so the predictions of regular weighting methods tend to deviate from true AVSS overall. With enough data to train, we believe that the model can learn and reduce this bias. This is the reason why the inverse-variance weighted average performs worse than the feature weighted estimation Eq. (13).

Physically, the different feature importance may correspond to the effects of different ocean fluctuations on different paths. The machine learning approach can obtain relative importance among the features, and the feature weighted average estimation will have fewer outliers.

4.3. Decision tree mechanism

The decision tree can avoid outliers further. Fig. 10 shows an overfit decision tree built from the top four most important features. For example, there are some true AVSSs (like 1541.626 m/s and 1541.456 m/s in Fig. 10), for which the arrival time of the second peak should be after 2.32 s in the time window. However, this arrival time drifts and has a wrong value in the extracted EGF. We would get an outlier if we used this feature to invert AVSS.

In model training, it is impossible to reduce the RMSE by partitioning the data on this feature. So the judgment condition in the red box automatically determines that if the arrival time of the second peak has a wrong value ($\leqslant$2.32 s), the model will use the arrival time of the fourth peak to predict this sample in the next step. It can be seen from Fig. 11 that the judgment mechanism of RDTs greatly reduces the occurrence of outliers in the prediction.

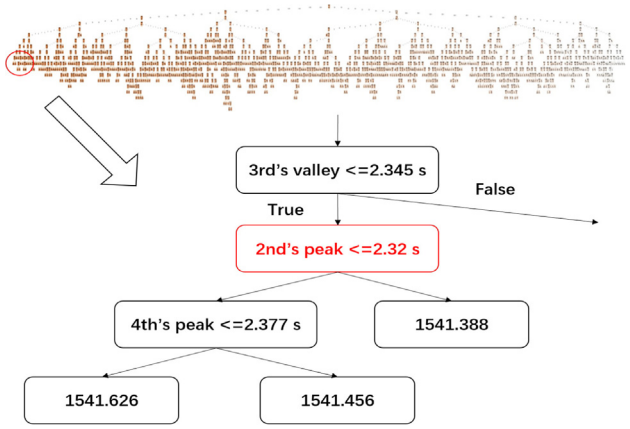

Fig. 10. a branch of the RDT in the RF built from the four most important features.

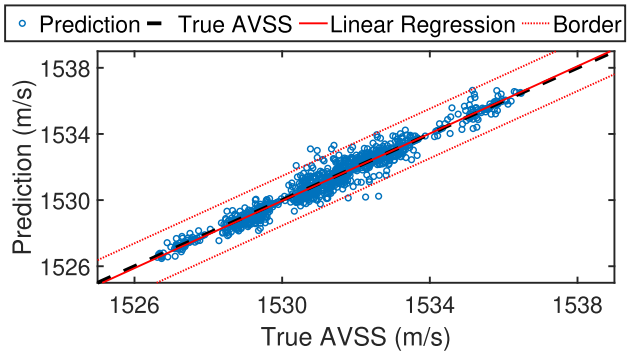

Fig. 11. Inversion results of the RF with an accumulation time of 30 min. Black Line represents the true AVSS, red line represents the linear regression of the predictions, and red dot line the upper and lower boundaries (linear regression +/−1.5 m/s) that we use to determine outliers. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

4.4. Sound speed profile inversion and deep learning

In our study, only AVSS is considered here. The sound speed profile (SSP) inversion problem is more challenging, especially for the SSP changes dramatically in a short time. And our model only takes manufacturing features based on physical knowledge as input. The accuracy of the prediction depends on the quality of the features we extract from the EGF. We still hope to reduce our reliance on the EGF and seek a better end-to-end solution in the AVSS inversion problem.

To solve this problem, additional information may need to be introduced. Although the Random Forest works well for data sets in this paper, more complicated machine learning algorithms, such as the Seq2Seq model with RNN or Attention, may be used to solve problems by extracting features directly from the noise signal or the EGF. Besides, more training data may be necessary for the SSP inversion problem in a more complicated ocean environment.

5. Conclusion

To summarize, this paper presents an approach for the POAT within a machine learning framework. By transforming the problem into a supervised learning problem, the AVSS can be estimated accurately within half an hour of accumulation time. Besides, by analyzing features and visualizing RDTs, an empirical equation is introduced to calculate the AVSS instead of choosing peaks and weights intuitively. The physical explanation of the results is also presented. The arrival times of the troughs are more stable than the corresponding peak arrival time, which can be explained as the reversed-phase of the surface reflection. Due to the effects of different ocean fluctuations on different paths, arrival times may be biased so that the inverse-variance weights between different features would be meaningless.

CRediT authorship contribution statement

Fenghua Li: Conceptualization, Methodology, Supervision, Writing - original draft, Writing - review & editing. **Kai Wang:** Methodology, Formal analysis, Software, Validation, Writing - original draft, Writing - review & editing. **Xishan Yang:** Software, Data curation, Writing - review & editing. **Bo Zhang:** Funding acquisition, Investigation. **Yanjun Zhang:** Investigation.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] Munk W, Wunsch C. Ocean acoustic tomography: A scheme for large scale monitoring. Deep Sea Res Part A, Oceanogr Res Pap 1979;26(2):123–61.
[2] Behringer D, Birdsall T, Brown M, Cornuelle B, Heinmiller R, Knox R, et al. A demonstration of ocean acoustic tomography. Nature 1982;299(5879):121–5.
[3] Munk W, Worcester P, Wunsch C. Ocean acoustic tomography. Cambridge University Press; 2009.
[4] Shang EC. Ocean acoustic tomography based on adiabatic mode theory. J Acoust Soc Am 1989;85(4):1531–7.
[5] Roux P, Kuperman WA. Group N. Extracting coherent wave fronts from acoustic ambient noise in the ocean. J Acoust Soc Am 2004;116(4):1995–2003.
[6] Sabra KG. Monitoring deep ocean temperatures using low-frequency ambient noise. J Acoust Soc Am 2015;137(4):2335.

[7] Godin OA, Zabotin NA, Goncharov VV. Ocean tomography with acoustic daylight. Geophys Res Lett 2010;37(13).
[8] Woolfe KF, Lani S, Sabra KG, Kuperman WA. Monitoring deep-ocean temperatures using acoustic ambient noise. Geophys Res Lett 2015;42(8):2878–84.
[9] Sabra KG, Roux P, Kuperman WA. Emergence rate of the time-domain Green's function from the ambient noise cross-correlation function. J Acoust Soc Am 2005;118(6):3524–31.
[10] Sabra KG, Roux P, Thode AM, D'Spain GL, Hodgkiss WS, Kuperman WA. Using ocean ambient noise for array self-localization and self-synchronization. IEEE J Ocean Eng 2005;30(2):338–47.
[11] Goncharov VV, Shurup AS, Godin OA, Zabotin NA, Vedenev AI, Sergeev SN, et al. Tomographic inversion of measured cross-correlation functions of ocean noise in shallow water using ray theory. Acoust Phys 2016;62(4):436–46.
[12] Sergeev SN, Shurup AS, Godin OA, Vedenev AI, Goncharov VV, Mukhanov PY, et al. Separation of acoustic modes in the Florida Straits using noise interferometry. Acoust Phys 2017;63(1):76–85.
[13] Sabra KG, Roux P, Kuperman WA. Arrival-time structure of the time-averaged ambient noise cross-correlation function in an oceanic waveguide. J Acoust Soc Am 2005;117(1):164–74.
[14] Siderius M, Harrison CH, Porter MB. A passive fathometer technique for imaging seabed layering using ambient noise. J Acoust Soc Am 2006;120(3):1315–23.
[15] Siderius M, Song H, Gerstoft P, Hodgkiss WS, Hursky P, Harrison C. Adaptive passive fathometer processing. J Acoust Soc Am 2010;127(4):2193–200.
[16] Leroy C, Lani S, Sabra KG, Hodgkiss WS, Kuperman WA, Roux P. Enhancing the emergence rate of coherent wavefronts from ocean ambient noise correlations using spatio-temporal filters. J Acoust Soc Am 2012;132(2):883–93.
[17] Fried SE, Walker SC, Hodgkiss WS, Kuperman WA. Measuring the effect of ambient noise directionality and split-beam processing on the convergence of the cross-correlation function. J Acoust Soc Am 2013;134(3):1824–32.
[18] Lani SW, Sabra KG, Hodgkiss WS, Kuperman W, Roux P. Coherent processing of shipping noise for ocean monitoring. J Acoust Soc Am 2013. 133(2):EL108–EL113..
[19] Evers LG, Wapenaar K, Heaney KD, Snellen M. Deep ocean sound speed characteristics passively derived from the ambient acoustic noise field. Geophys J Int 2017;210(1):27–33.
[20] Li F, Yang X, Zhang Y, Luo W, Gan W. Passive ocean acoustic tomography in shallow water. J Acoust Soc Am 2019;145(5):2823–30.
[21] Bianco MJ, Gerstoft P, Traer J, Ozanich E, Roch MA, Gannot S, et al. Machine learning in acoustics: Theory and applications. J Acoust Soc Am 2019;146(5):3590–628.
[22] Niu H, Reeves E, Gerstoft P. Source localization in an ocean waveguide using supervised machine learning. J Acoust Soc Am 2017;142(3):1176–88.
[23] Godin OA. Recovering the acoustic Green's function from ambient noise cross correlation in an inhomogeneous moving medium. Phys Rev Lett 2006;97(5):54301.
[24] Zang X, Brown MG, Godin OA. Waveform modeling and inversion of ambient noise cross-correlation functions in a coastal ocean environment. J Acoust Soc Am 2015;138(3):1325–33.
[25] Jensen FB, Kuperman WA, Porter MB, Schmidt H. Computational ocean acoustics. Springer Science & Business Media; 2011.
[26] Biau G. Analysis of a random forests model. J Mach Learn Res 2012. 13(Apr):1063–1095..
[27] Breiman L. Bagging predictors. Mach Learn 1996;24(2):123–40.
[28] Breiman L. Random forests. Mach Learn 2001;45(1):5–32.
[29] Liaw A, Wiener M. Classification and regression by randomforest. R news 2002;2(3):18–22.
[30] Wenz GM. Acoustic ambient noise in the ocean: Spectra and sources. J Acoust Soc Am 1962;34(12):1936–56.
[31] Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, et al. Scikit-learn: Machine learning in python. J Mach Learn Res 2011;12:2825–30.