euronoise 2021
25th to 27th of October
Madeira, Portugal - online

Learning the finite size effect for in-situ absorption measurement

Elias Zea[1], Eric Brandão[2], Mélanie Nolan[3], Joakim Andén[4], Jacques Cuenca[5], U. Peter Svensson[6]

[1]KTH Royal Institute of Technology, Department of Engineering Mechanics, The Marcus Wallenberg Laboratory for Sound and Vibration Research, Stockholm, Sweden
zea@kth.se
[2]Federal University of Santa Maria, Department of Structures and Construction, Acoustical Engineering, Santa Maria, Brazil
eric.brandao@eac.ufsm.br
[3]Technical University of Denmark, Department of Electrical Engineering, Acoustic Technology, Kgs. Lyngby, Denmark
melnola@elektro.dtu.dk
[4]KTH Royal Institute of Technology, Department of Mathematics, Stockholm, Sweden
janden@kth.se
[5]Siemens Industry Software, Leuven, Belgium
jacques.cuenca@siemens.com
[6]Norwegian University of Science and Technology, Department of Electronic Systems, Trondheim, Norway
peter.svensson@ntnu.no

Abstract
In this paper we propose the use of neural networks to predict the sound absorption coefficient spectra of finite porous samples with microphone arrays. The main goal is to train a model that can effectively mitigate the errors caused by the finite size effect. A convolutional neural network architecture is used to map the array data to the absorption coefficient at five frequencies. The training, validation and test data are numerically produced with a boundary element method; modelling a baffled, locally reacting porous absorber on a rigid backing with a Delany–Bazley–Miki model, for varying sample size, thickness, flow resistivity, incidence angle and frequency. The strength of using machine learning in this context is that no hypotheses are made about the sound field or the absorber, as the networks learn the necessary relationships from the data. We show that the network approximates well the absorption coefficient, as if the sample was infinite, in a wide range of cases.

Keywords: sound absorption, in-situ measurement, convolutional neural networks, finite size effect, Delany–Bazley–Miki model.

1 Introduction

Free-field or in-situ methods of measuring the absorption of acoustic materials aim at inferring the absorption properties (surface impedance, reflection and absorption coefficients) from measurements of the sound field in the vicinity of the measurement sample [1, 2, 3, 4]. Although not standardised, the attractiveness of these methods lies in the fact that they provide angle-dependent absorption data (which cannot be measured with standardised methods), and are applicable for materials mounted for their intended application.

These inverse methods rely on a mathematical model of the sound field above the material and generally assume a measurement sample of infinite extent (i.e., that the sample is large enough for the acoustic field to be sufficiently small at the edges). Yet, in practical data acquisition, measurement samples are limited in size, and the (total) measured sound field will differ from the case of an infinite sample due to diffraction phenomena evoked at the sample's edges. At low- and mid-frequencies, the so-called "edge-diffraction effect" (or "finite

size effect") leads to discrepancies between prediction and experimental data [5, 6]. This effect is negligible at high frequencies where the wavelengh is much smaller than the sample size.

Considerable effort has been spent on the problem of measuring the sound absorption of finite-size samples in-situ. In particular, a number of studies have compared experimental data with boundary element method (BEM) simulations in order to describe and account for the edge-diffraction effect (see e.g., [7, 8, 9, 10]). Yet, fewer studies have attempted to characterise the edge-diffraction effect experimentally [11, 12].

In recent years, data-driven deep learning approaches have yielded promising advances in acoustics [13]. In particular, convolutional neural networks (CNNs) [14] have successfully been applied for porous material parameter estimation [15], room acoustical parameter estimation [16], direction of arrival (DOA) estimation [17, 18], and near-field acoustic holography [19, 20]. In this paper, we propose to use CNNs to estimate the absorption coefficient of finite-size samples, by learning a mapping of pressure fields including edge diffraction effects to sound absorption. The main advantage of using a data-driven approach in this context is that no hypotheses on the nature of the sound field are necessary, as these are learned from the data. In this study, the data used for training, validation, and testing is generated numerically based on a boundary element model (BEM) of the sound field near a baffled, locally-reacting, homogeneous and isotropic porous layer on a rigid backing, for varying sample size, thickness, flow resistivity, incidence angle, and frequency. We assess the performance of the method to unseen data using two different test sets, and compare the predicted absorption coefficient against a benchmark solution based on the two-microphones method [1].

2 Methodology

2.1 Data generation via BEM

For the sound field simulation, a simplified BEM is considered, as in Ref. [9]. A 2D depiction of the system under consideration is given in Fig. 1. Here, a point source is located at coordinate $\mathbf{r}_\mathrm{q} = (x_\mathrm{q}, y_\mathrm{q}, z_\mathrm{q})$ and a receiver is located at coordinate $\mathbf{r} = (x, y, z)$. A finite rectangular absorber sample of dimensions $L_x \times L_y$ is flush mounted to an infinite hard baffle at the plane $z = 0$.

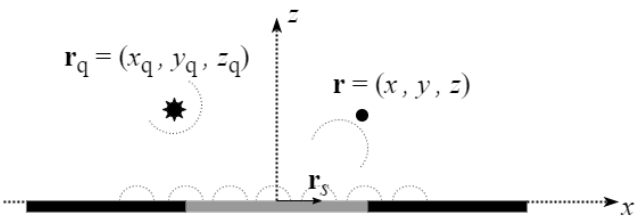

Figure 1 – Schematic of the BEM simulation. The point source at $\mathbf{r}_\mathrm{q}$ excites the sound field. A receiver is located at $\mathbf{r}$ and $\mathbf{r}_s$ is a point at the surface of the sample. The incident, reflected and scattered waves are also depicted schematically as wave fronts.

The sound pressure, p($\mathbf{r}$), can be written as the Helmholtz/Huygens integral

$$c(\mathbf{r})p(\mathbf{r}) = \frac{\mathrm{e}^{-\mathrm{j}k|\mathbf{r}-\mathbf{r}_\mathrm{q}|}}{|\mathbf{r}-\mathbf{r}_\mathrm{q}|} + \frac{\mathrm{e}^{-\mathrm{j}k|\mathbf{r}-\mathbf{r}_\mathrm{q}^{'}|}}{|\mathbf{r}-\mathbf{r}_\mathrm{q}^{'}|} - \frac{\mathrm{j}k}{Z_s}\int_S p(\mathbf{r}_s)\frac{\mathrm{e}^{-\mathrm{j}k|\mathbf{r}-\mathbf{r}_s|}}{4\pi|\mathbf{r}-\mathbf{r}_s|}\,\mathrm{d}S, \tag{1}$$

where $k = 2\pi f/c_0$ is the wavenumber in air, f is the frequency, c_0 the speed of sound and Z_s is the surface impedance of the finite sample. The vector $\mathbf{r}$ can be above or on the surface of the sample; $\mathbf{r}_s$ is any point at the surface of the finite sample; $c(\mathbf{r})$ is 0.5 if $\mathbf{r}$ is located on the absorptive surface and 1.0 if $\mathbf{r}$ is located at any point above the absorptive surface. The first term on the right-hand side of Eq. (1) is the Green's function between the sound source and the receiver. The second term is the Green's function between the image sound source, at $\mathbf{r}_\mathrm{q}^{'} = (x_\mathrm{q}, y_\mathrm{q}, -z_\mathrm{q})$, and the receiver. Together they form the unperturbed sound field, as if the sample itself is not present. The last term carries the information of the absorption and diffraction on the finite absorber, formulated as an integral over the finite absorber sample area S. The absorption term is

modelled by a prescribed surface impedance, Z_s, constant across the finite sample's surface. The characteristic impedance and the wavenumber for the material are computed with the Delany–Bazley–Miki model [21]. In the following, we assume a locally-reacting, hard-backed porous layer of thickness d. Thus, for a given angle of incidence, the surface impedance and absorption spectra can be calculated from a well-known transfer matrix method (TMM) [22]. These constitute the reference absorption spectra, which are used as labels to perform the supervised training of the CNN.

The surface of the sample can be discretised into N square elements, with $p(\mathbf{r}_s)$ considered constant over each element. Therefore, Eq. (1) can be rewritten as

$$c(\mathbf{r})p(\mathbf{r}) = \frac{\mathrm{e}^{-\mathrm{j}k|\mathbf{r}-\mathbf{r}_\mathrm{q}|}}{|\mathbf{r}-\mathbf{r}_\mathrm{q}|} + \frac{\mathrm{e}^{-\mathrm{j}k|\mathbf{r}-\mathbf{r}_\mathrm{q}^{'}|}}{|\mathbf{r}-\mathbf{r}_\mathrm{q}^{'}|} - \frac{\mathrm{j}k}{Z_s}\sum_{n=1}^{N} p(\mathbf{r}_{sn}) \int_{S_n} \frac{\mathrm{e}^{-\mathrm{j}k|\mathbf{r}-\mathbf{r}_s|}}{4\pi|\mathbf{r}-\mathbf{r}_s|}\,\mathrm{d}S_n, \tag{2}$$

where the collocation method is used by placing $\mathbf{r} = \mathbf{r}_i$ at $i = 1, 2, \ldots, N$ on the surface of the sample. Thus a system of equations is formed and the surface pressure, $p(\mathbf{r}_{sn})$, at each element can be found. Once the surface pressure is known, it can be re-inserted into Eq. (2) to calculate the pressure at any receiver point for $z > 0$. An array of receivers is considered in this study. The array aperture is 0.6×0.6 m and the receivers are arranged in a regular grid of 12×12 at a distance of 2 cm from the surface of the sample. The highest simulated frequency is 2 kHz, with six elements per wavelength. The integrals in Eq. (2) are calculated with linear interpolation and Gauss–Legendre quadrature with 36 points on each element [23, 24]. With the implemented configuration, the Gauss points do not coincide with the element center, which avoids singularities. Experimental validation for such BEM simulations using single point estimates can be found in Ref. [9].

2.2 Two-microphone method

As a reference for the validation of the proposed method, the classical two-microphone method [1] is used. The method makes use of two microphones placed along the normal to the surface of interest in order to separate the incident and reflected components of the field. This is done under the assumption of specular reflection, such that the reflected sound field arises from the image source at $\mathbf{r}_q'$ as a spherical wave [3]. Furthermore, the reflection coefficient is assumed to be that of plane waves. The sound pressure at the microphones is thus

$$p(\mathbf{r}_i) = \frac{\mathrm{e}^{-\mathrm{j}k|\mathbf{r}_i-\mathbf{r}_q|}}{|\mathbf{r}_i-\mathbf{r}_q|} + R(f)\frac{\mathrm{e}^{-\mathrm{j}k|\mathbf{r}_i-\mathbf{r}_q'|}}{|\mathbf{r}_i-\mathbf{r}_q'|}. \tag{3}$$

where $\mathbf{r}_i,\ i = 1, 2$ are the positions of the two microphones. The reflection coefficient of the sample at frequency f is thus estimated as

$$R(f) = \frac{\dfrac{\mathrm{e}^{-\mathrm{j}k|\mathbf{r}_2-\mathbf{r}_q|}}{|\mathbf{r}_2-\mathbf{r}_q|} - \dfrac{p(r_2)}{p(r_1)}\dfrac{\mathrm{e}^{-\mathrm{j}k|\mathbf{r}_1-\mathbf{r}_q|}}{|\mathbf{r}_1-\mathbf{r}_q|}}{\dfrac{p(r_2)}{p(r_1)}\dfrac{\mathrm{e}^{-\mathrm{j}k|\mathbf{r}_1-\mathbf{r}_q'|}}{|\mathbf{r}_1-\mathbf{r}_q'|} - \dfrac{\mathrm{e}^{-\mathrm{j}k|\mathbf{r}_2-\mathbf{r}_q'|}}{|\mathbf{r}_2-\mathbf{r}_q'|}}, \tag{4}$$

and the sound absorption coefficient is

$$\alpha(f) = 1 - |R(f)|^2. \tag{5}$$

In an experimental context, the distance between the microphones must be large enough for a phase difference to be observed [1], but small enough to avoid spatial aliasing. For the purposes of numerical illustrations in the present paper, the microphones are respectively placed at 1 cm and 3 cm from the sample.

2.3 Convolutional neural network

We adopt an architecture based on two convolutional layers and two fully connected layers. A schematic of the network can be seen in Fig. 2. The input data consists of the 12×12 array predictions of the absolute sound pressure at 5 frequencies; which means that the network has 5 channels, each of size 12×12, as input. The

first convolutional layer Conv1 has 16 filter channels, each with kernel size 2×2 and stride equal to one, and rectified linear unit (ReLU) activation functions. The second convolutional layer Conv2 has 32 filters and the remaining properties same as those of the first layer. At the output of Conv2 there is a max-pooling operator, which downsamples the image by a factor of two and takes the maximum value over patches of size 2×2. These values are then flattened into a vector of size 800 and fed into a fully connected layer FC1, containing 100 neurons with ReLU activation functions. Lastly, the output is another fully connected layer with 5 neurons, each corresponding to the absorption coefficient at each frequency, and using a sigmoid activation function to constrain the output to the interval $[0, 1]$.

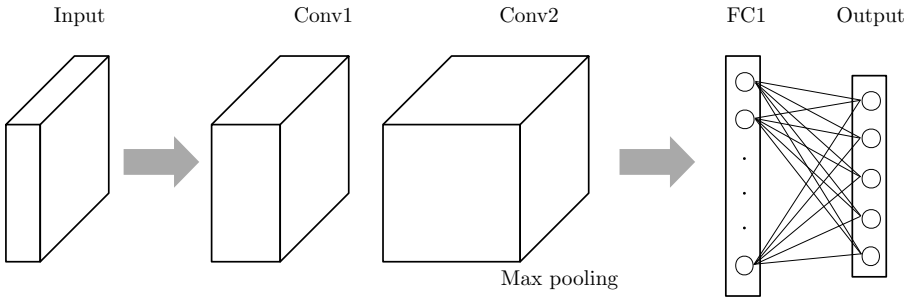

Figure 2 – Schematic of the CNN architecture. The input data is an array of size $12 \times 12 \times 5$, where the 5 channels correspond to the 5 frequencies. Conv1 and Conv2 correspond to the first and second convolutional layers, respectively, while FC1 and the output layer are fully connected layers. The output is a vector of size 1×5, containing the absorption coefficient spectrum.

Table 1 – Summary of network architecture. Total trainable parameters: 83021.

Operation	Kernel size	No. channels	Activation	Output size
Input		5		$12 \times 12 \times 5$
Conv2D	2×2	16	ReLU	$11 \times 11 \times 16$
Conv2D	2×2	32	ReLU	$10 \times 10 \times 32$
Max pooling				$5 \times 5 \times 32$
Flatten				1×800
Fully connected			ReLU	1×100
Fully connected			Sigmoid	1×5

3 Results

3.1 Training and validation

The dataset used for training and validation consists of 285120 cases generated with the BEM model. Each case contains the magnitude of the 12×12 pressure fields at five frequencies: 125 Hz, 250 Hz, 500 Hz, 1 kHz and 2 kHz. These cases are obtained with the BEM model given the parameter combinations indicated in Table 2. Examples of three of these input cases can be seen in Fig. 3. The source distance is set to 1.5 m, the sound speed is 343 m/s, and the air density is 1.21 kg/m^3.

Note that we use a single-layer pressure array, in contrast to acoustical holography-based methods (e.g., [11, 12, 25, 26]) that use multiple layers to separate incident from reflected fields. Additionally, we take the absolute pressure instead of the complex-valued pressure. The motivation behind these choices is two-fold. First, it reduces the complexity of the network and the number of sensor positions. Second, the direct field is the same for all cases of the training set that have the same incidence angle – regardless of the other BEM model parameters. This enables the network to exploit this invariance of the direct field to predict the sound absorption; however, verification of this is out of scope for the current work.

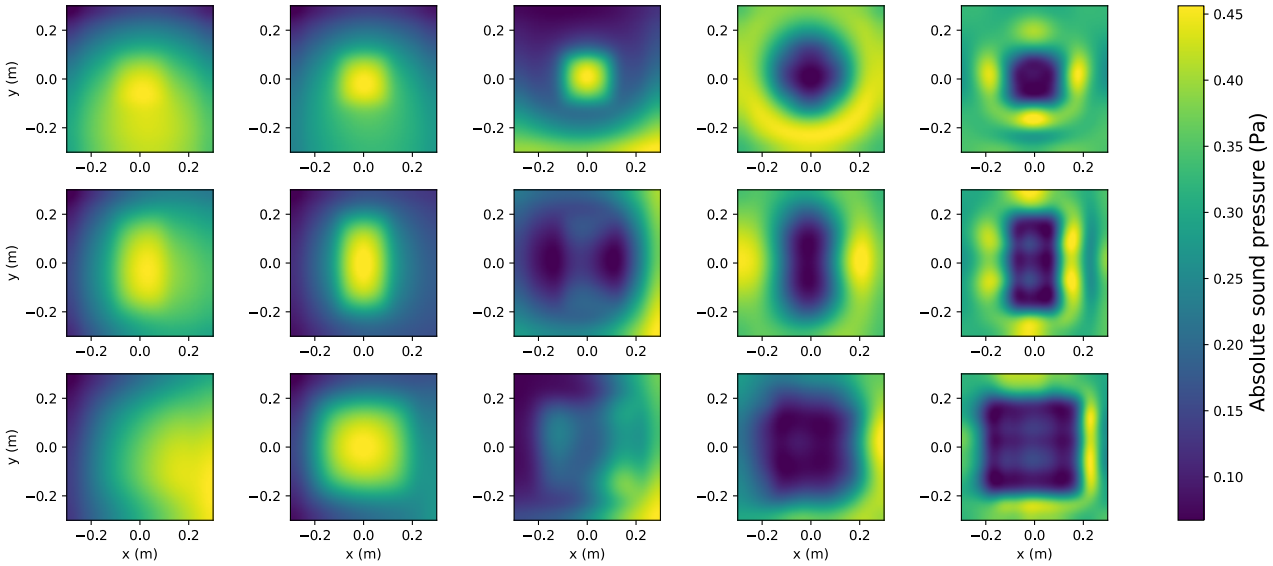

Figure 3 – Input examples from the training set. Absolute sound pressure (Pa) at the 12×12 array obtained with the BEM model. Colormaps are interpolated for easier visualisation. Rows: three different samples. Columns: frequencies 125 Hz, 250 Hz, 500 Hz, 1 kHz and 2 kHz.

Table 2 – BEM model parameters used to generate the training and validation sets.

Parameter	Min value	Max value	Step
Sample size L_x [cm]	20	80	20
Sample size L_y [cm]	20	80	20
Flow resistivity σ [Ns/m^4]	5000	10000	5000
Thickness d [mm]	5	200	20
Source elevation angle θ [°]	0	80	10
Source azimuth angle ϕ [°]	0	180	20

The 285120 cases are randomly split into 80% and 20% for training and validation, respectively. The loss function is the mean-squared error (MSE), a common choice for statistical regression. We use the Adam optimiser as it is computationally efficient and has low memory requirements [27]. Training is stopped after 200 epochs. Figure 4 shows the training and validation loss, as well as the mean absolute error (MAE) versus the number of epochs. It can be seen that the validation loss converges at about 100 epochs, and thereafter it improves marginally. For testing we use the network trained 200 epochs.

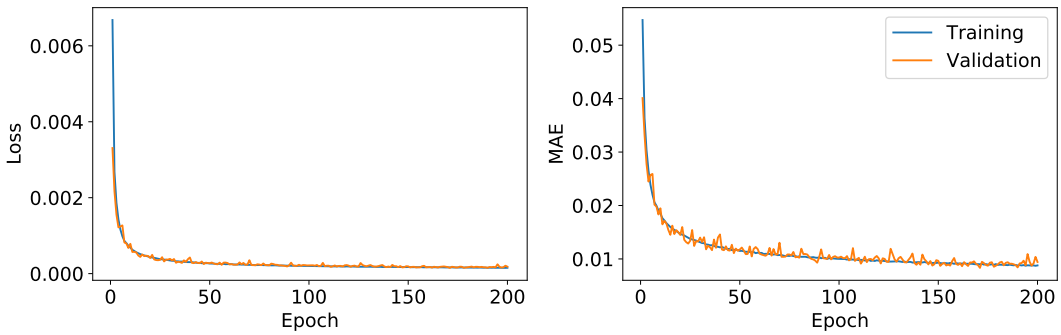

Figure 4 – Learning curves versus number of epochs. Left: Training and validation loss (MSE). Right: Mean absolute error (MAE) during training and validation.

3.2 Testing and benchmark comparisons

Two tests are done to assess the performance of the CNN to unseen data. In each of these tests, the CNN predictions are compared with those of the two-microphone method and the TMM reference.

3.2.1 Interpolation test

This test is done to evaluate the performance of the network against unseen cases whose parameters lie *inside* the range of parameters of the training set. The network is tested with an interpolation data set consisting of 15000 additional cases. These cases are generated with the BEM model, using 15 base cases (fixed sample sizes) and drawing random combinations of the remaining parameters in Table 2 with uniform distribution. The source distance is 1.5 m, the same as in the training set. The MSE and MAE for the entire data set are 0.002 and 0.03, respectively.

Figure 5 shows the results for four different cases. Two of which correspond to the smallest in size of the interpolation set (top row), and the others correspond to the largest (bottom row). The thickness and flow resistivity values lie in the lower (left column) and higher (right column) ends of the intervals in Table 2.

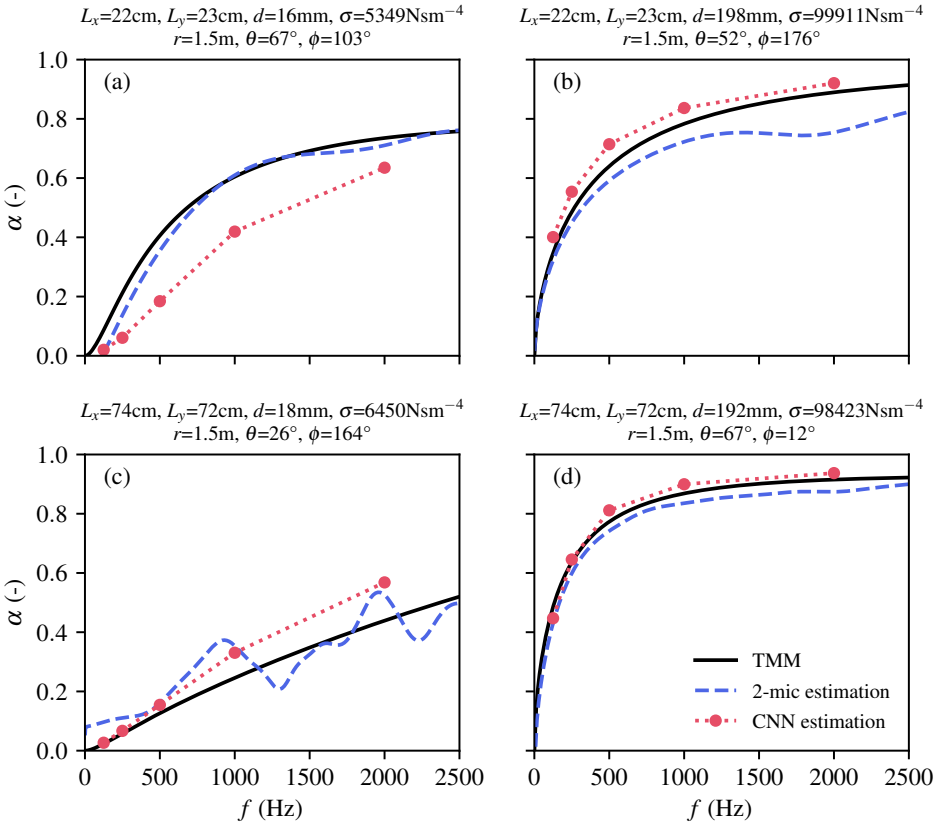

Figure 5 – Sound absorption coefficient spectra of four samples from the interpolation test set. Prediction with the two-microphone method (dashed), with the CNN (dotted circles), and TMM reference (solid).

It can be seen that the predictions with the two-microphone method contain spurious artefacts, most noticeably in Figs. 5(b) and (c). These oscillations are known in the literature [7, 8, 9] and can be attributed to the finiteness of the sample. On the other hand, the absorption predicted by the CNN agrees reasonably well with the reference for the cases in Figs. 5(b) and (d); cases with thicker, more resistive samples. For the case in Fig. 5(c) the CNN still provides a reasonable prediction, slightly overestimating the absorption at 1 kHz and 2 kHz.

As shown in Fig. 5(a), however, the CNN underestimates the absorption curve, most noticeably at frequencies above 250 Hz. A possible explanation for this is related to the proportion of training samples that have

similar BEM model parameters to those of Fig. 5(a). The joint distribution of absorption coefficient for the whole training set (approx. 1.4 million data points) is shown in Fig. 6. Most cases are highly absorptive (near 1), and this is reasonable since the absorption coefficient of such porous samples increases with frequency. This also means that non-absorptive cases are less common in the training set, which could pose difficulties at the moment of predicting the absorption of, e.g., thin and less resistive samples. We also suspect the model may be sensitive to the particular (and perhaps other) choices of parameters, and further work is needed to verify this.

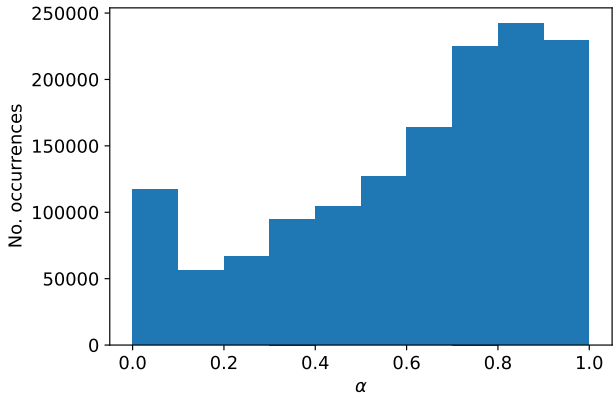

Figure 6 – Joint distribution histogram of the absorption coefficient α for the whole training set.

3.2.2 Extrapolation test

The network is tested with an extrapolation set consisting of 15000 additional cases. As with the interpolation test set, 15 base cases are generated with the BEM model, drawing random combinations of the remaining parameters within the intervals shown in Table 2. The difference in this test is that the source distance is also varied randomly between 1 and 2 m. This test is done to evaluate the performance of the network against unseen cases whose parameters (in this case source distance) lie *outside* the range of parameters of the training set. The MSE and MAE for the entire data set is 0.003 and 0.04, respectively.

Figure 7 shows the results for four different cases. Two of which correspond to the smallest in size of the extra polation set (top row), and the others correspond to the largest (bottom row). The source distance is closer [7(a) and (d)] or farther [7(b) and (c)] than the one of the training set (1.5 m). It can be seen that the predictions with the two-microphones method only agree well with the reference in the last case, shown in Fig. 7(d); and it again contains spurious artefacts in the remaining cases. In contrast, the CNN predicts the absorption curves reasonably well, even when the source distance is different from the one in the training set. This indicates that the network has the potential to generalise to unseen cases.

Additionally, we conduct an analysis of the network performance versus source distance. For this analysis, the extrapolation data set is clustered into ten groups, dividing the source distance into intervals $(1, 1.1]$ m, $(1.1, 1.2]$ m, etc. The resulting MSE and MAE curves are shown in Fig. 8. It can be seen that the errors increase as the source distance moves away from 1.5 m (the distance in the training set), where it attains a minimum. However, the error increase is rather gentle, which suggests that the network is not overfitting too much. In particular, the errors increase faster when the source distance is decreased than when it is increased.

4 Conclusion

This paper introduces the use of convolutional neural networks (CNNs) to estimate the sound absorption coefficient of an infinitely large material sample from the knowledge of the behaviour of a finite rectangular specimen. The input data consists of the absolute acoustic pressure on a single layer of microphones over the finite sample of interest at five frequencies, and the output is the sound absorption coefficient at those frequencies. For convenience, the data for training, validation and testing is generated with a BEM model, comprising

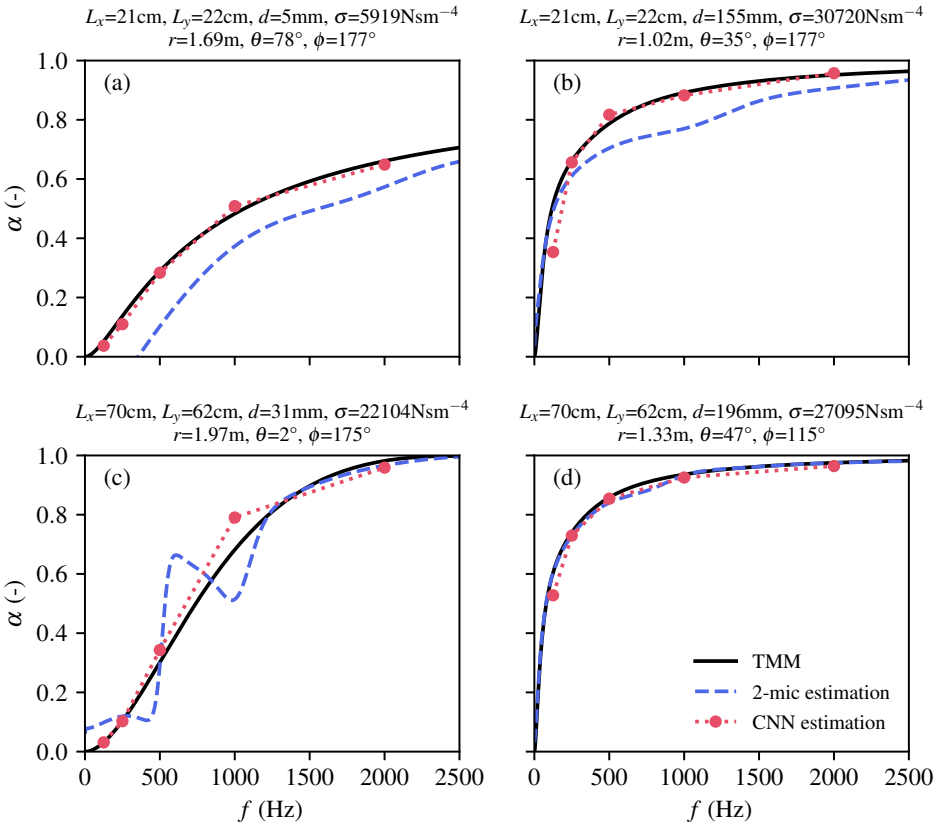

Figure 7 – Sound absorption coefficient spectra of four samples from the extrapolation test set. Prediction with the two-microphone method (dashed), with the CNN (dotted circles), and TMM reference (solid).

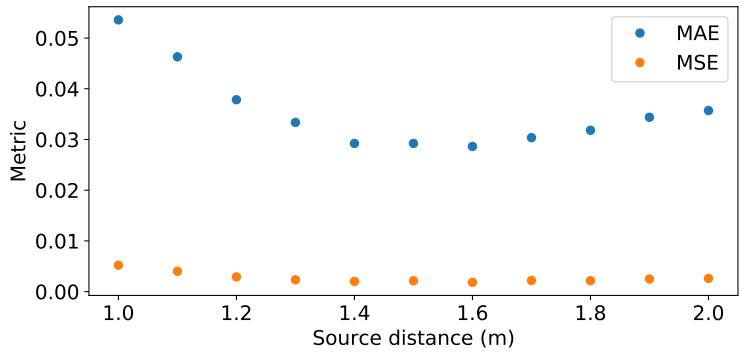

Figure 8 – Mean-squared error (MSE) and mean absolute error (MAE) of the CNN model on the extrapolation set as a function of source distance.

a baffled, locally-reacting porous absorber on a rigid backing. The BEM model is parametrized for varying sample sizes, thickness, flow resistivity, incidence angle, and frequency. A Delany–Bazley–Miki model is used to compute the wavenumber and the characteristic impedance of the sample. Due to the numerical modelling there are assumptions about the samples in question, such as homogeneity, local reaction, and the dependence of the impedance and wavenumber only on frequency and flow resistivity. However, we suspect the learning process could be extended to more complex cases. The network is trained for 200 epochs, and its performance is assessed against unseen data in two different tests and compared with benchmark solutions with the two-microphone method. In both tests the network outperforms the benchmark solution, and provides mean-squared errors and mean absolute errors in the order of 0.002 and 0.03, respectively. This is a preliminary

study, and future work includes training other architectures to account for single- or multiple-frequency maps, a thorough error analysis across the parameters of the BEM model, and validation with experimental data.

5 Acknowledgements

This work has been financed by Vetenskapsrådet under Grant Agreement 2020-04668.

References

[1] Allard, J.F., Sieben, B. Measurements of acoustic impedance in a free field with two microphones and a spectrum analyzer. J. Acoust. Soc. Am. 77(4), 1985, pp. 1617–1618.

[2] Mommertz, E. Angle-dependent in-situ measurements of reflection coefficients using a subtraction technique. Appl. Acoust. 46, 1995, pp. 251-263.

[3] Li, J.F., Hodgson, M. Use of pseudo-random sequences and a single microphone to measure surface impedance at oblique incidence. J. Acoust. Soc. Am. 102(4), 1997, pp. 2200–2210.

[4] Brandão, E., Lenzi, A., Paul, S. A review of the in situ impedance and sound absorption measurement techniques. Acta. Acust. United Ac. 101, 2015, pp. 443-463.

[5] de Bruijn, A. A mathematical analysis concerning the edge effect of sound absorbing materials. Acta. Acust. United Ac. 28, 1973, pp. 33-44.

[6] Thomasson, S. I. On the absorption coefficient. Acta. Acust. United Ac. 44, 1980, pp. 265-273.

[7] Otsuru, T., Tomiku, R., Din, N. B. C., Okamoto, N., Murakami, M. Ensemble averaged surface normal impedance of material using an in-situ technique: Preliminary study using boundary element method. J. Acoust. Soc. Am. 125(6), 2009, pp. 3784–3791.

[8] Hirosawa, K., Takashima, K., Nakagawa, H., Kon, M., Yamamoto, A., Lauriks, W. Comparison of three measurement techniques for the normal absorption coefficient of sound absorbing materials in the free field. J. Acoust. Soc. Am. 126(6), 2009, pp. 3020–3027.

[9] Brandão, E., Lenzi, A., Cordioli, J. Estimation and minimization of errors caused by sample size effect in the measurement of the normal absorption coefficient of a locally reactive surface. Appl. Acoust. 73, 2012, pp. 543–556.

[10] Luo, Z.-W., Zheng, C.-J., Zhang, Y.-B., Bi, C.-X. Estimating the acoustical properties of locally reactive finite materials using the boundary element method. J. Acoust. Soc. Am. 147(6), 2020, pp. 3917-3931.

[11] Ottink, M., Brunskog, J., Jeong, C.-H., Fernandez-Grande, E., Trojgaard, P., Tiana-Roig, E. In situ measurements of the oblique incidence sound absorption coefficient for finite sized absorbers. J. Acoust. Soc. Am. 139(1), 2016, pp. 41–52.

[12] Richard, A., Fernandez-Grande, E., Brunskog, J., Jeong, C.-H. "Estimation of surface impedance at oblique incidence based on sparse array processing," J. Acoust. Soc. Am. 141(6), 2017, pp. 4115-4125

[13] Bianco, M. J., Gerstoft, P., Traer, J., Ozanich, E., Roch, M. A., Gannot, S., Deledalle, C.-A. Machine learning in acoustics: Theory and applications. J. Acoust. Soc. Am. 146(5), 2019, pp. 3590-3628.

[14] Goodfellow, I., Bengio, Y., Courville, A. Deep Learning, Ch. 9, MIT Press, 2016.

[15] Lähivaara, T., Kärkkäinen, L., Huttunen, J.M., Hesthaven, J.S. Deep convolutional neural networks for estimating porous material parameters with ultrasound tomography, J. Acoust. Soc. Am. 143(2), 2018, pp. 1148-1158.

[16] Yu, W., Kleijn, W.B. Room acoustical parameter estimation from room impulse responses using deep neural networks. IEEE/ACM Trans. Audio Speech Lang. Process. 29, 2021, pp. 436-447.

[17] S. Chakrabarty, S., Habets, E. A. Broadband DOA estimation using convolutional neural networks trained with noise signals. IEEE Workshop Appl. Signal Process. Audio Acoust., 2017, pp. 136–140.

[18] Cao, H., Wang, W., Su, L., Ni, H., Gerstoft, P., Ren, Q., Ma, L. Deep transfer learning for underwater direction of arrival using one vector sensor. J. Acoust. Soc. Am. 149(3), 2021, pp. 1699-1711.

[19] Olivieri, M., Pezzoli, M., Malvermi, R., Antonacci, F., Sarti, A. Near-field acoustic holography analysis with convolutional neural networks. Proc. Inter-noise, Seoul, 2020.

[20] Wang, J., Zhang, Z., Huang, Y., Li, Z., Huang, Q. A 3D convolutional neural network based near-field acoustical holography method with sparse sampling rate on measuring surface. Measurement 177, 2021, 109297.

[21] Miki, Y. Acoustical properties of porous materials: Modifications of Delany-Bazley models. J. Acoust. Soc. Jpn.(E) 11(1), 1990, pp. 19-24.

[22] Allard, J.F., Atalla, N. Propagation of Sound in Porous Media: Modeling Sound Absorbing Materials, 1 Ed., John Wiley & Sons, 2009.

[23] Atalla, N., Sgard, F. Finite Element and Boundary Element Methods in Structural Acoustics and Vibration, 1 Ed., CRC Press, 2015.

[24] Wu, T.W. Boundary Element Acoustics: Fundamentals and Computer Codes, 1 Ed., WIT Press, 2000.

[25] Tamura, M. Spatial Fourier transform method of measuring reflection coefficients at oblique incidence. I: Theory and numerical examples. J. Acoust. Soc. Am. 88(5), 1990, pp. 2259-2264.

[26] Nolan, M. Estimation of angle-dependent absorption coefficients from spatially distributed in situ measurements. J. Acoust. Soc. Am. 147(2), 2020, EL119-EL124.¨

[27] Kingma, D.P., Ba, J. Adam: A method for stochastic optimization. Proc. ICLR, 2015.