




PredictUS: A Method to Extend the Resolution-Precision Trade-Off in Quantitative Ultrasound Image Reconstruction

Farah Deeba¹  and Robert Rohling^{1,2}

¹ Department of Electrical and Computer Engineering, The University of British Columbia, Vancouver, BC, Canada

{farahdeeba,rohling}@ece.ubc.ca

² Department of Mechanical Engineering, The University of British Columbia, Vancouver, BC, Canada

Abstract. We present PredictUS, a novel Quantitative Ultrasound (QUS) parameter estimation technique with improved resolution and precision using augmented ultrasound data. The ultrasound data is generated using a sequence-to-sequence convolutional neural network based on WaveNet. The spectral-based QUS techniques are limited by the well-studied trade-off between the precision of the estimated QUS parameters and the window size used in estimation, limiting the practical utility of the QUS techniques. In this paper, we present a method to increase the window size by predicting the next data points of a given window. The method provides better estimates of local tissue properties with high resolution by virtually extending the property to a larger region. Our proof-of-concept study based on attenuation coefficient estimate (ACE), an important QUS parameter, attains a resolution reduction up to 50% while maintaining comparable estimation precision. This result shows the promise to extend the precision-resolution trade-off, which, in turn, would have implications in small lesion detection or heterogeneous tissue characterization.

Keywords: Quantitative Ultrasound · Attenuation coefficient estimate · WaveNet · Sequence-to-sequence neural network

1 Introduction

Quantitative Ultrasound (QUS) Imaging has introduced a paradigm shift in the field of biomedical imaging. Extending beyond qualitative B-mode ultrasound imaging, QUS presents clinically significant parametric images, which are descriptive of underlying tissue microstructure. Recent studies show that QUS potentially provides effective, non-invasive, and system independent biomarkers for non-alcoholic fatty liver disease (NAFLD) detection and monitoring, cervical

ripening detection, placenta characterization, and breast lesion characterization [1–3].

QUS extracts acoustic scattering and attenuating properties using algorithms based on estimates of power spectra of ultrasound radiofrequency (RF) signal backscattered from the interrogated tissue. The spectral based QUS techniques allow the normalization of the backscattered RF signal and thus filter out the system-dependent factors such as focusing, diffraction and transducer electromechanical response [1, 4]. Unfortunately, the power spectral estimation, typically obtained from FFT based periodogram of windowed RF signal, imposes a fundamental trade-off on QUS between image resolution and estimation precision. Smaller windows provide high spatial resolution, a desirable property for many imaging applications such as characterization of thin (e.g. human skin) or heterogeneous (e.g. placenta) tissue. However, smaller windows yield noisy and inaccurate power spectra estimates due to limited spectral resolution and spatial variation noise inherent in ultrasonic scattering. Larger windows improve accuracy and precision of power spectra estimates and therefore the estimation of QUS parameters, with an expense of reduced spatial resolution [5, 6]. One study found that the trade-off between spatial resolution and the variance of QUS parameter is optimized with a window size of 10 independent scanlines laterally and 10 times the wavelength axially [6].

To expand the precision-resolution trade-off, different modifications of periodogram have been investigated. Welch method was found to yield the most accurate and precise spectral estimate with reasonable computational cost [6]. Alternatively, autoregressive (AR) techniques have been reported to exceed the performance of FFT based periodogram, especially for smaller windows [7]. However, AR techniques show degraded performance with increasing depth in higher attenuating media due to violation of the stationarity assumption. More recently, deep neural networks such as WaveNet [8], have significantly improved state-of-the-art performance in fields of forecasting non-stationary and non-linear processes, such as speech and financial time-series [8, 9].

In this work, we present PredictUS, a spectral based QUS technique based on US RF signal prediction using a sequence-to-sequence convolutional neural network (CNN) modelled with a WaveNet inspired architecture. Given a small windowed RF signal, this method predicts the next data points, resulting in a larger window. Therefore, the method yields better estimates of power spectra as well as can characterize local tissue properties with high resolution by essentially extending the property to a larger region. We demonstrate the applicability of PredictUS for improved measurement of attenuation coefficient estimate (ACE), a QUS parameter.

2 Method

The first signal processing step in spectral-based QUS techniques is the estimation of power spectra from a limited-length RF signal window. The proposed PredictUS method adds a WaveNet inspired deep neural network before the

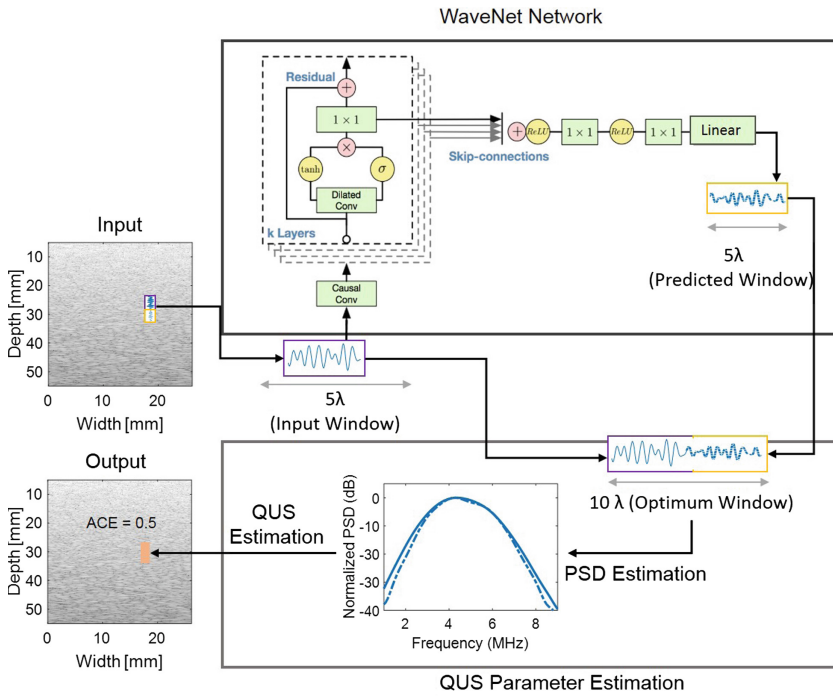


Fig. 1. Overview of the proposed PredictUS method. Blue solid line indicates the original input window, whereas the blue dash-dotted line indicates the predicted signal in the “WaveNet Network” block and the power spectrum obtained from the optimum (i.e. PredictUS) window in the “QUS Parameter Estimation” block. (Color figure online)

power spectra estimation step (Fig. 1). The network results in a larger window by sequentially predicting the next samples. The power spectra estimated from the larger RF signal windows are fed for the subsequent QUS parameter estimation.

2.1 ACE Computation

The ultrasound ACE is a measure of ultrasound amplitude dissipation due to the combined effect of scattering and absorption. ACE can be measured using the reference phantom method [4], a standard method to account for the system dependent factors.

According to this method, the RF data are acquired from both the tissue sample, s and a reference phantom, r with known properties using the same transducer and system settings. For a RF signal window centered at depth z from the transducer surface, the natural logarithm of the ratio of the power spectrum S from the sample to the reference phantom at frequency $f \in (f_1, f_k)$

can be written as [2, 4]:

$$\ln \frac{S_r(f, z)}{S_s(f, z)} = -4(\alpha_s(z) - \alpha_r(z))fz + \ln \frac{B_s}{B_r}. \quad (1)$$

where α is the effective ACE for the total ultrasound propagation path z and B is the backscatter coefficient (BSC). Substituting the following variables in Eq. 1 as: $\ln \frac{S_r(f, z)}{S_s(f, z)} = Y(f, z)$, $\alpha_r - \alpha_s = \alpha$, $\ln \frac{B_s}{B_r} = \beta$, we get,

$$Y(f, z) = -4\alpha(z)fz + \beta. \quad (2)$$

The above equation can be written in a matrix form: $\mathbf{y} = \mathbf{A}\mathbf{x} + \mathcal{N}(0, \sigma_N \mathbf{I})$, where

$$\mathbf{A} = \begin{bmatrix} 4zf_1 & 1 \\ \vdots & \vdots \\ 4zf_k & 1 \end{bmatrix}_{k \times 2}, \quad \mathbf{y} = \begin{bmatrix} Y(f_1, z) \\ \vdots \\ Y(f_k, z) \end{bmatrix}_{k \times 1}, \quad \mathbf{x} = \begin{bmatrix} \alpha \\ \beta \end{bmatrix}_{2 \times 1}.$$

A least square fitting method [11] can be applied to solve for $\mathbf{x} = [\alpha, \beta]$ from the noisy estimation \mathbf{y} as follows:

$$\hat{\mathbf{x}} = \arg \min_{\mathbf{x}} \{ \|\mathbf{y} - \mathbf{A}\mathbf{x}\|_2^2 \}, \quad (3)$$

with the following constraints:

$$\alpha_{min} \leq \alpha \leq \alpha_{max}, \beta_{min} \leq \beta \leq \beta_{max} \quad (4)$$

Solving Eq. 2 gives us the effective ACE (α) for the total ultrasound propagation path. The local ACE at depth z_i can be computed as: $\alpha^{local}(z_i) = \frac{\alpha(z_i)z_i - \alpha(z_{i-1})z_{i-1}}{z_i - z_{i-1}}$.

2.2 Network Architecture

We employed a sequence-to-sequence CNN using WaveNet model. First introduced by the researchers from DeepMind, Wavenet is an autoregressive model, where each predicted sample is conditioned on the previous ones [8]. One key element of WaveNet is stacked layers of 1-dimensional dilated causal convolution. Causal convolutions are used to ensure that a prediction at time step t only depends on the previous time steps, whereas the use of a dilation rate increased as a factor of 2 results in an exponentially growing receptive field with depth. For the US RF signal prediction, a receptive field is required which is large enough to capture several wavelengths. For our application, one wavelength (λ) is approximately 20 samples for US transmission frequency of 5 MHz and sampling frequency of 50 MHz. We use 14 dilated causal convolution layers with a dilation rate of factor 2 with a reset ($2^0, 2^1, \dots, 2^6, 2^0, 2^1, \dots, 2^6$) and 64 filters with width of 2. As in original architecture, there are gated activation unit (combining a hyperbolic tangent and a sigmoid activation branch), residual, and skip connections (Fig. 1) in each of these layers to speed up the convergence and enable improved training for deeper models. Finally, the WaveNet output is passed through a ReLU activation followed by a linear projection.

3 Experiments and Results

3.1 Data

We used the k-Wave toolbox [10] to generate simulation US RF data. For the simulation, we use 96 element linear array transducer with 0.2 mm element pitch. The depth was set to 60 mm. Fixed focusing was used in transmission (focal depth 60 mm) and dynamic focusing was used in reception. We simulate 256 RF lines for each of the 32 different ACE values ranging from 0.1 to 1.65 dB/cm/MHz. The selected ACE range encompasses the observed ACE values in liver at different NAFLD stages and in placenta. Finally, we divide each RF line into 13 segments with 25% overlap, resulting in a training set of 106,496 RF line segments. We also created a separate test dataset of 550 examples, where 50 RF line segments are extracted for each of the 11 different ACE values (0.5–1.5 dB/cm/MHz).

3.2 Training and Testing

In the training stage, a *teacher-forcing* procedure is applied where the model performs a one-step ahead prediction. The model outputs a sequence of n steps, which is one time-step shifted version of the input sequence. Therefore, the model is trained using the correct output instead of the predicted output. We use a batch size of 64 and use Adam optimizer to minimize mean squared error loss with $\beta = 0.9$, $\beta_2 = 0.999$ and learning rate of 0.001.

In contrast to the training stage, the testing stage makes a n -step ahead prediction sequentially by feeding each prediction back into the network at the next time step.

3.3 RF Data Processing and Analysis

For ACE computation, the lateral dimension for the RF signal window was kept fixed at 5 scanlines. The axial dimension was varied from 5λ (100 samples) to 10λ (200 samples) to study the effect of PredictUS in improving the trade-off between resolution and precision. For the power spectrum computation, the Welch method has been found to yield more accurate and precise estimation compared to rectangular, Hanning, or Hamming windows [6]. Therefore, we used the Welch method to estimate the power spectrum from the RF scanlines within each RF window. According to the Welch method, each RF scanline within a window was subdivided into overlapping sections, with length equal to 67% of the original RF scanline and with 50% overlap. Each segment was then multiplied with a Hamming window. The power spectral density was obtained after averaging the periodograms obtained from the windowed segments. We considered the -20 dB bandwidth of the received power spectrum as the usable frequency range. To compute the ACE, we utilize the RF data with ACE of 0.5 dB/cm/MHz as the reference data.

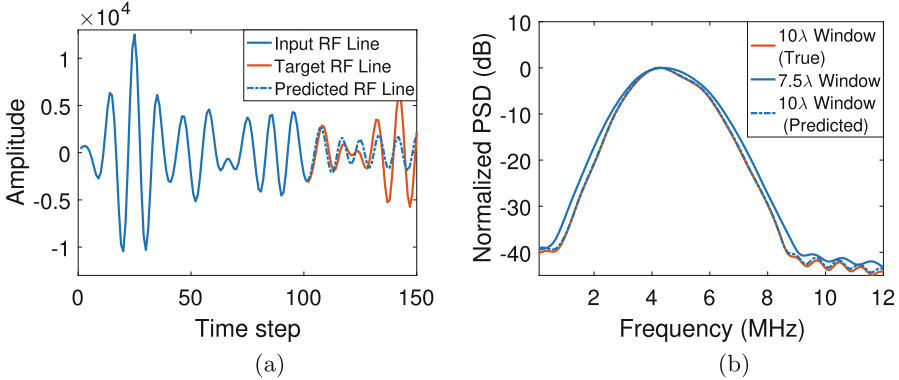


Fig. 2. (a) An example of RF line segment prediction using the proposed sequence-to-sequence CNN. (b) Comparison of power spectra estimation using original large window (10λ), original small window (7.5λ) and PredictUS large window (10λ) predicted from 7.5λ original window.

3.4 Performance Metrics

We use Mean Absolute Scaled Error (MASE) to measure the performance of sequence-to-sequence CNN to predict the larger RF window from the smaller one, where

$$MASE = \frac{\frac{1}{T} \sum_{t=1}^T |e_t|}{\frac{1}{T-1} \sum_{t=2}^T |Y_t - Y_{t-1}|} \quad (5)$$

Here, e_t is the prediction error, defined as the difference between the actual value and the predicted value and the denominator denotes the in-sample mean absolute error from the naive forecast method. A MASE value < 1 indicates a prediction performance better than the naive forecast method.

As a measure of precision of ACE, we report the standard deviation of the computed ACE as a percentage of the actual value. We also report the bias in the estimated ACE, as the difference between the estimated ACE and actual ACE presented as a percentage of the actual ACE.

3.5 Results

Performance of RF Data Prediction. We apply the trained sequence-to-sequence CNN on the test dataset. An example of RF data prediction is shown in Fig. 2a, where a 50-step (2.5λ) ahead prediction was made, given a RF line segment of 150 steps (7.5λ). We can see a precise prediction for the initial RF samples, which starts to degrade with increasing time step. We found that the power spectra estimate obtained from the PredictUS window of length 10λ (generated from 7.5λ segment) gives similar estimates as obtained from a larger window (10λ segment) and outperforms the estimates obtained from the original 7.5λ segment (Fig. 2b).

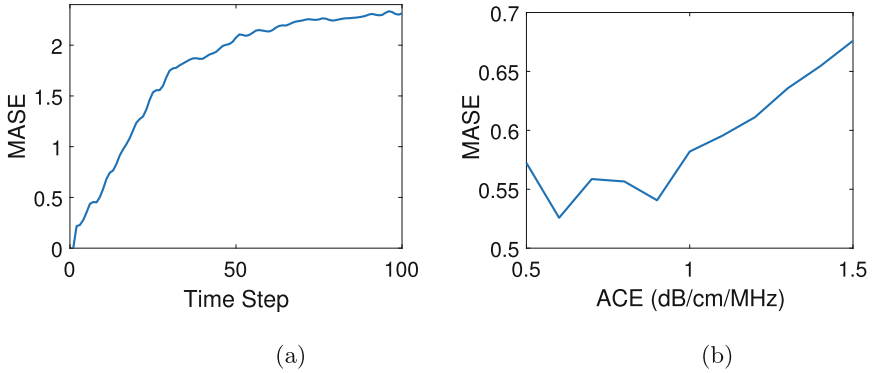


Fig. 3. RF data prediction performance: (a) for n -step ahead prediction with varying n , and (b) for varying ACE in term of mean absolute scaled error (MASE).

We analyse the performance of the network for n -step prediction for varying values of n . It was found that, on average, the network can make a 16-step prediction when $MASE < 1$. This performance is comparable to the previous work on time-series forecasting using WaveNet model, which reports MASE for a single-step ahead prediction [9]. Additionally, we investigated whether the ACE amplitude has an effect on the prediction performance. We computed MASE for a 15-step prediction for data with ACE varying from 0.5 to 1.5 dB/cm/MHz. From Fig. 3, we see that the MASE remains within a range of 0.52–0.58 for $ACE \leq 1$ dB/cm/MHz, and after that MASE starts to degrade with increasing ACE. This result agrees with previous finding from the AR based techniques where higher attenuating medium showed inferior QUS estimation performance [7].

Performance of ACE Computation. We investigate the performance of the proposed PredictUS method in extending the resolution-precision trade-off inherent in QUS parameter estimation. According to [6], an axial dimension of 10λ has been defined to be the optimum keeping the variance and bias of QUS parameter estimation within 10%. Taken these numbers as the baseline, in this study, we examined the limit to which the trade-off can be extended by simulating three cases with increasing difficulty as follows:

1. Case I: PredictUS window of length 10λ with 2.5λ predicted using 7.5λ window;
2. Case II: PredictUS window of length 10λ with 5λ predicted using 5λ window;
3. Case III: PredictUS window of length 10λ with 7.5λ predicted using 2.5λ window;

Here, PredictUS window refers to the window obtained by concatenating the original input window and the predicted window. For each case, the precision as well as the bias in ACE computation have been compared with the original ‘large’

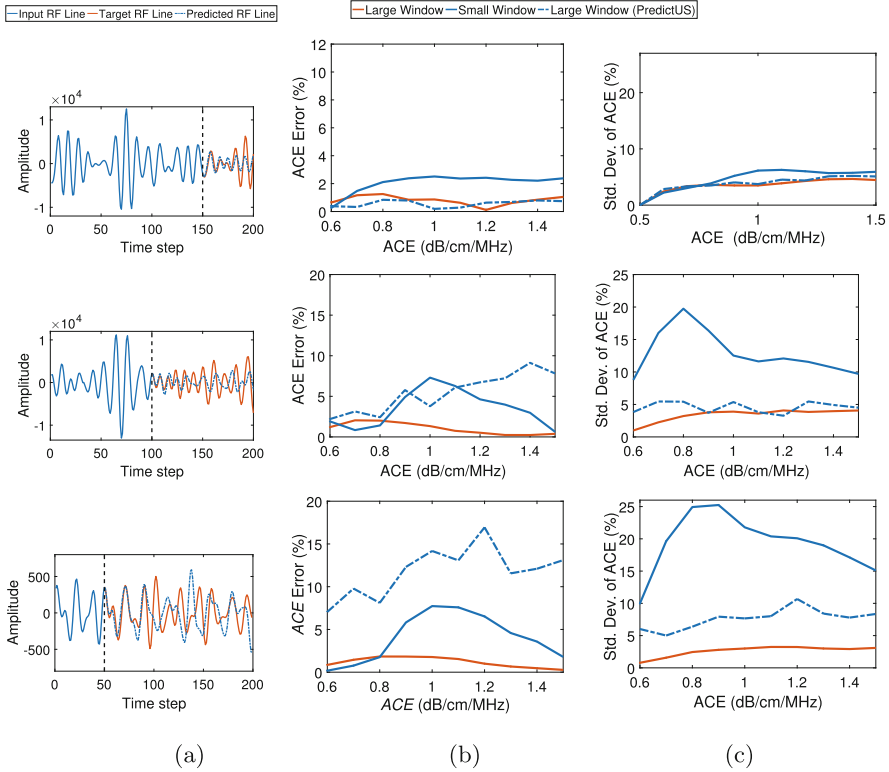


Fig. 4. PredictUS performance for case I (top), case II (middle), and case III (bottom). (a) Example RF line segments and their divisions; (b) ACE error for different ACE values; (c) Standard deviation of ACE for different ACE values.

RF window with length equal to the PredictUS window, and with the original ‘small’ RF window with length equal to the original data used in computing PredictUS window. The ACE computation results for these three cases have been demonstrated in Fig. 4.

For case I, the small window (7.5λ) with original data attains a precision and accuracy performance where ACE error and the standard deviation remain within 4% and 7%, respectively. The large window (10λ) with original data improves both the precision and the bias where ACE error and the standard deviation both remain within 2% and 6%, respectively. The ACE estimation using PredictUS window outperforms the estimation from the small window and achieves performance equivalent to that obtained from the large window.

In case II, the difference between the ACE measures obtained from small window (5λ) and the large window (10λ) is more prominent, where small window results in a bias up to 7.3% and standard deviation as large as 19.7%. Interestingly, the PredictUS window, using original data of length 5λ only, can

attain a precision within 5.5%, slightly larger than that obtained from the large window (4%). The PredictUS window gives moderate bias (within 6%) for $ACE < 1.1$ dB/cm/MHz.

Finally, for case III, the PredictUS window only includes 2.5λ original data, however, achieves a standard deviation within 10%. Although, the increasingly accumulated error in the prediction of RF data affects the accuracy, resulting in a larger ACE error compared to the small original window.

In summary, compared to the optimum trade-off, the PredictUS method can reliably maintain similar precision and accuracy while improving the resolution to 75% of the optimum value. Moreover, the proposed method can still achieve comparable precision and accuracy for low ACE values with a resolution improvement of 50%. However, reducing the resolution to 25% of the optimum exhibit degraded ACE measurement, which can be attributed to the error accumulation in the n -step ahead prediction. Unlike the case of RF data prediction, high ACE does not have any distinct effect on the performance of ACE computation.

4 Conclusion

We propose a novel QUS parameter estimation method, PredictUS, utilizing ultrasound RF signal prediction. The method shows promising results by predicting larger RF windows from the smaller ones. We conduct a proof-of-concept study based on extensive simulation analysis. The proposed sequence-to-sequence convolutional neural network based on WaveNet model was able to estimate RF signal samples to a reasonable accuracy and therefore improve the power spectral estimate. A resolution reduction, as high as 50%, while maintaining comparable estimation precision introduces a paradigm shift by challenging the insistent trade-off between precision and resolution, inherent in ultrasound spectral estimation. Future research will address the issue of error accumulation in the n -step prediction by further improving the CNN structure.

References

1. Oelze, M.L., Mamou, J.: Review of quantitative ultrasound: envelope statistics and backscatter coefficient imaging and contributions to diagnostic ultrasound. *IEEE Trans. Ultrason. Ferroelectr. Freq. Control* **63**(2), 336–351 (2016)
2. Deeba, F., et al.: Attenuation coefficient estimation of normal placentas. *Ultrasound Med. Biol.* **45**(5), 1081–1093 (2019)
3. Deeba, F., et al.: SWTV-ACE: spatially weighted regularization based attenuation coefficient estimation method for hepatic steatosis detection. In: *International Conference on Medical Image Computing and Computer-Assisted Intervention* (2019)
4. Yao, L.X., et al.: Backscatter coefficient measurements using a reference phantom to extract depth-dependent instrumentation factors. *Ultrason. Imaging* **12**(1), 58–70 (1990)
5. Oelze, M.L., O'Brien Jr., W.D.O.: Defining optimal axial and lateral resolution for estimating scatterer properties from volumes using ultrasound backscatter. *J. Acoust. Soc. Am.* **115**(6), 3226–3234 (2004)

6. Liu, W., Zagzebski, J.A.: Trade-offs in data acquisition and processing parameters for backscatter and scatterer size estimations. *IEEE Trans. Ultrason. Ferroelectr. Freq. Control* **57**(2), 340–352 (2010)
7. Wear, K.A., Wagner, R.F.: A comparison of autoregressive spectral estimation algorithms and order determination methods in ultrasonic tissue characterization. *IEEE Trans. Ultrason. Ferroelectr. Freq. Control* **42**(4), 709–716 (1995)
8. Oord, A.V.D., et al.: WaveNet: a generative model for raw audio. *arXiv preprint [arXiv: 1609.03499](https://arxiv.org/abs/1609.03499)* (2016)
9. Borovykh, A., Bohte, S., Oosterlee, C.W.: Conditional time series forecasting with convolutional neural networks. *arXiv preprint [arXiv:1703.04691](https://arxiv.org/abs/1703.04691)* (2017)
10. Treeby, B.E., et al.: Modeling nonlinear ultrasound propagation in heterogeneous media with power law absorption using a k-space pseudospectral method. *J. Acoust. Soc. Amer.* **131**(6), 4324–4336 (2012). *Medical physics of CT and ultrasound: Tissue imaging and characterization*
11. Nam, K., Zagzebski, J.A., Hall, T.J.: Simultaneous backscatter and attenuation estimation using a least squares method with constraints. *Ultrasound Med. Biol.* **37**(12), 2096–2104 (2011)