ICA 2022
24th INTERNATIONAL CONGRESS ON ACOUSTICS
October 24(Mon) - 28(Fri), 2022 Gyeongju, Korea

PROCEEDINGS of the
24th International Congress on Acoustics

October 24 to 28, 2022 in Gyeongju, Korea

A quantitative comparison of linear and nonlinear loudspeaker models

Franz M. Heuchel[*,1] and Finn T. Agerkvist[1]

[1]Technical University of Denmark, Acoustic Technology, Lyngby, Denmark

[*]fmheu@dtu.dk

ABSTRACT

We present a short review of models for the dynamics of the conventional moving-coil loudspeaker and a data set with which we quantitatively compare their predictive errors. The models include the effects of viscoelasticity and nonlinear stress-strain relation in the suspension, eddy-current losses in the magnet system, a finite air gap and nonlinear inductance. Vibrational wave propagation within the diaphragm is not considered. We present a model fitting technique and benchmark the loudspeaker models on the data set to establish baselines for future work. A Python implementation is made available at https://github.com/fhchl/quant-comp-ls-mod-ica22.

Keywords: electro-dynamic loudspeaker, modelling, dynamical systems, parameter estimation

1 INTRODUCTION

This paper reports work-in-progress on establishing baseline models that predict the nonlinear dynamics of electro-dynamic moving-coil loudspeaker drivers. We focus on time-invariant lumped-element models that predict the voice-coil displacement and velocity, as well as the current passing through it, from the voltage signal applied to the driver terminals. These models are typically used when attempting to (input-output) linearize the loudspeaker dynamics in order to reduce nonlinear distortion in audio reproduction [1, 2, 3, 4]. Arguably, the more accurate these models are in predicting the dynamics, the more effectively the linearization approaches will reduce nonlinear distortion. This is especially relevant if the linearization is based on a feedforward estimation of the state (as proposed e.g. by Klippel [1]) in contrast to state-feedback (as is custom in the control literature [5]).

The simplest loudspeaker model includes an electric resistance R_e and inductance L of the voice coil of length l, coupled via a magnetic field of strength B to the mechanical system of total moving mass M, suspension stiffness K and mechanical resistance R_m. An externally applied voltage u induces a current i through the coil leading to a change in displacement x. The dynamics of this system are described by the linear differential equations

$$u = Bl\dot{x} + R_e i + \frac{\mathrm{d}}{\mathrm{d}t}[iL] \qquad Bli = Kx + R_m\dot{x} + M\ddot{x}, \tag{1a,b}$$

where $\dot{x}$ and $\ddot{x}$ indicate velocity and acceleration. This model has since been extended to include many other effects and properties including viscoelasticity of the suspension [6, 7, 8], lossy inductance due to eddy currents [9], the magnetic reluctance force [10, 11], nonlinear stress-strain of suspension and spider, inhomogeneous magnetic field strength in the air-gap, and an inductance that varies with displacement [12] and current [13].

To measure research progress in the area of loudspeaker modelling, it is important to quantitatively compare the predictive errors of these models on a wide spread of loudspeaker drivers. This leads us to the main contributions of the present work: 1) We compare loudspeaker models on data gathered from 6 moving-coil loudspeaker drivers. Our analysis gives insight into the impact of including above effects and highlights some of the shortcoming of existing models. 2) We review how loudspeaker models are fitted to data and provide the data set and an implementation of the procedures in Python.

The paper is structured as follows: Section 2 reviews the implemented models. Section 3 presents the model fitting procedure. Section 4 presents the data set and section 5 the results of the model predictions on that data.

Model	Description
linear	Equations (1)
linear SLS	linear with standard linear solid element term
linear L2R2	linear with L2R2 term
linear L2R2+SLS	union of two above
full w/o reluctance	above with nonlinear terms (4, 5) for $d = 4$ and $f = 2$
full	above with reluctance force term of Risbo et al. [11]
full reluctance as [13]	above instead with reluctance term by Klippel [13]
full w/o $L(i)$	full without current-dependent inductance term (5)
full w/o SLS	full without standard linear solid element term
full w/o L2R2	full without L2R2 term

Table 1. Electro-dynamic loudspeaker models compared in this study.

2 LOUDSPEAKER MODELS

This section reviews extensions of the basic linear differential equation model (1) of the electro-dynamic loudspeaker driver. An overview over the tested models is given in table 1.

2.1 Linear models

Eddy currents are known to decrease the inductance at high frequencies and introduce additional losses. The effect can be modelled by a fractional-order derivative model [14]. To avoid some of the computational challenges of such a model, it can be approximated over a limited bandwidth with the L2R2 model, which adds, in series to the inductance, an additional inductance L_2 in parallel with an electrical resistance $R_{e,2}$ [15]. This effectively introduces the additional electrical damping term $R_2(i - i_2)$ in the right hand side of (1a), where i_2 is the auxiliary current through L_2 which satisfies

$$\frac{\mathrm{d}}{\mathrm{d}t}[i_2 L_2] = (i - i_2)R_2. \tag{2}$$

The surround of many loudspeakers is made of rubber polymers that are viscoelastic: their stiffness and damping varies with frequency [16]. In loudspeakers, this linear effect is also well modelled by a fractional-order derivative model [6, 14]. This can be modelled in analogy to the L2R2 model, here called the standard linear solid (SLS) model, which consists of the original spring in series to an additional stiffness K_2 in parallel to a resistance $R_{m,2}$. It adds the additional force term $K_2 x_2$ to the right hand side of (1b), where x_2 is an auxiliary state with evolution

$$\dot{x}_2 = \dot{x} - x_2 K_2 / R_{m,2}. \tag{3}$$

2.2 Nonlinear models

A common way to introduce nonlinearity into (1) is to model Bl, K, and L as polynomials of the displacement x,

$$Bl(x) = \sum_{n=0}^{d} Bl_n x^n, \quad K(x) = \sum_{n=0}^{d} K'_n x^n, \quad L(x) = \sum_{n=0}^{d} L'_n x^n, \tag{4}$$

where Bl_n, K'_n and L'_n are the polynomial coefficients of the displacement dependence up to some order d.

The magnetic field generated by the voice coil can change the magnetic permeability (operating point) of the iron structure and thus the magnetic flux through the coil. This current-dependent inductance can be modelled via separable polynomials, i.e. the inductance in (4) is replaced by

$$L(x,i) = \left(\sum_{n=0}^{d} L_n x^n\right)\left(1 + \sum_{n=1}^{f} F_n i^n\right), \tag{5}$$

where F_n are the polynomial coefficients of the current dependence up to some order f.

Dodd, Klippel, and Oclee-Brown [17] show that the lossy inductance effect is also state dependent and model it by varying the parameters L_2 and $R_{e,2}$ as scaled functions of $L(x,i)$, i.e. $R_{e,2}(x,i) = R_{e,2} L(x,i)/L(0,0)$ and $L_2(x,i) = L_2 L(x,i)/L(0,0)$. Similar extensions of the SLS model may be useful but are not investigated here.

It is well known that a displacement dependent inductance implies the existence of an additional force term in the rhs of (1b), $-\frac{1}{2}i^2\partial L(x)/\partial x$, sometimes called reluctance force, flux-modulation or Cunningham force [10]. Risbo et al. [11] described the integration of the L2R2 model with the reluctance force, leading to the force term $-\frac{1}{2}i(i\partial L/\partial x + i_2\partial L_2/\partial x)$. This was modelled previously as $-\frac{1}{2}(i^2\partial L/\partial x + i_2^2\partial L_2/\partial x)$, e.g. by Klippel [13].

In summary, the "full" model including all the mentioned extensions is given by equations 3-5 and

$$u = Bl(x)\dot{x} + R_e i + R_{e,2}(i-i_2) + \frac{\mathrm{d}}{\mathrm{d}t}\left[iL(x,i)\right], \qquad Bl(x)i = K(x)x + K_2x_2 + R_m\dot{x} + M\ddot{x} - \frac{1}{2}i\left(i\frac{\partial L(x,i)}{\partial x} + i_2\frac{\partial L_2(x,i)}{\partial x}\right). \tag{6}$$

with the state variables $x' = [i, x, \dot{x}, i_2, x_2]$.

3 PARAMETER ESTIMATION OF DIFFERENTIAL EQUATION MODELS

This section describes the estimation of the parameters of above differential equation models from data. The loudspeaker state trajectory for a given input voltage signal and a given model is the solution of the initial value problem

$$\begin{aligned} \dot{x}' &= f_\theta(t, x') \quad x'(t=0) = 0 \\ y &= h(x'), \end{aligned} \tag{7}$$

where f is the vector field (e.g. from the differential equations (1)) containing the time derivative of the state vector x' (e.g. $x' = [i, x, \dot{x}]$), θ are the parameters (e.g. $\theta = [Bl, K, L, R, R_e, M]$), y is the measurement vector of size Q (e.g. $y = [i, \dot{x}]$, $Q = 2$), and h selects the measured states from the state vector x'.

In practice, the input voltage and the measurement signals of duration T are only given at $N = f_s T$ discrete samples gathered at a rate f_s, i.e. $\tilde{u}_n = u(n\Delta t)$, $\tilde{y}_n = y(n\Delta t)$ for $n = 0, \ldots, N-1$. Let's define our parameterized forward model $F_\theta : \mathbb{R}^N \to \mathbb{R}^{N\times Q}$ that implements the map $\tilde{u} \mapsto \hat{y}$ from the discrete input voltage sequence to a prediction of the measured states. The forward model consists of three steps: 1) interpolate the sampled input sequence $\tilde{u}$ with cubic splines to get a continuous-time signal u, 2) solve the initial value problem (7) for the state sequence $\tilde{x}'_n$, and 3) compute the output sequence $\tilde{y}_n$ using h.

3.1 Nonlinear system identification via maximum likelihood

We assume that the state measurement is corrupted by zero-mean Gaussian noise and that all measurements have the same signal-to-noise ratio. The maximum likelihood estimate of the parameters for a given signal pair $(\tilde{u}, \tilde{y})$ is then the minimizer of the weighted sum of squared errors:

$$\hat{\theta} = \arg\min_{\theta\in\mathscr{C}} \sum_{i=1}^{M} \frac{1}{\mathrm{Var}[\tilde{y}]_i} ||[\tilde{y} - F_\theta(\tilde{u})]_i||^2, \tag{8}$$

where $[\cdot]_i$ indexes the components of the measurement vectors, the variance is taken over the time samples, and $\mathscr{C}$ denotes non-negative constraints on the linear parameters (the parameters that are not multiplied by a state variable in (4,5)). This is equivalent to minimizing the normalized mean-squared-error. Equation (8) is a box-constrained nonlinear least-squares problem that can be solved with a dedicated solver.

A good initial parameter guess for the above least-squares problem is important. Assuming that the loudspeaker is weakly nonlinear, we can use an estimate of the parameters of the model linearized around the state $x' = 0$ as a starting guess, while initiating all other parameters with zeros.

3.2 Linear system identification via matching of cross-spectral spectral densities

The parameters of the linearized systems, θ_{lin}, can be estimated by matching the transfer functions of the linearized model to an estimate of the transfer-functions from data [18]. For the persistently exciting input-output pair $(\tilde{u}, \tilde{y})$, the transfer function of the linearized system is computed from data at discrete frequencies ω via estimates of the corresponding cross- and auto-spectral densities:

$$\hat{G}_{iu}(\mathrm{j}\omega) = \frac{\hat{S}_{iu}(\mathrm{j}\omega)}{\hat{S}_{uu}(\mathrm{j}\omega)} \quad \hat{G}_{xu}(\mathrm{j}\omega) = \frac{\hat{S}_{xu}(\mathrm{j}\omega)}{\hat{S}_{uu}(\mathrm{j}\omega)} \quad \hat{G}_{\dot{x}u}(\mathrm{j}\omega) = \frac{\hat{S}_{\dot{x}u}(\mathrm{j}\omega)}{\hat{S}_{uu}(\mathrm{j}\omega)}. \tag{9}$$

Linear parameter	Bl_0 [N/A]	R_e [Ω]	R_m [Ns/m]	K_0 [N/m]	L_0 [H]	M [kg]	L_2 [H]	$R_{e,2}$ [Ω]	K_2 [N/m]	$R_{m,2}$ [Ns/m]
initial value	1	1	1	10^3	10^{-3}	10^{-3}	10^{-3}	1	10^3	10^3

Table 2. Initial guesses for the parameters θ_{lin} of the linearized systems.

Both densities are here estimated using Welch's method [19] with a window size of 2^{14} samples. Let's assume we have only measurements of current i, the applied voltage u and the velocity $\dot{x}$. The linear parameters can be estimated by solving the nonlinear least-squares problem

$$\min_{\theta_{\text{lin}}} \sum_k \left(\hat{G}_{iu}(\mathrm{j}\omega_k) - G_{iu}^{\theta_{\text{lin}}}(\mathrm{j}\omega_k)\right)^2 + \left(\hat{G}_{vu}(\mathrm{j}\omega_k) - G_{vu}^{\theta_{\text{lin}}}(\mathrm{j}\omega_k)\right)^2 \tag{10}$$

The transfer-function estimates in (9) have high variance at frequencies where $\hat{S}_{uu}$ is small. A more robust estimate of the parameters can be found by weighting the above residuals by $\hat{S}_{uu}$, thereby giving more emphasis to those frequencies where the input signal has high power. Thus, we use the estimate

$$\hat{\theta}_{\text{lin}} = \min_{\theta_{\text{lin}}} \sum_k \frac{1}{\mathrm{Var}_\omega[\hat{S}_{iu}]} \left(\hat{S}_{iu}(\mathrm{j}\omega_k) - G_{iu}^{\theta_{\text{lin}}}(\mathrm{j}\omega_k)\hat{S}_{uu}(\mathrm{j}\omega_k)\right)^2 + \frac{1}{\mathrm{Var}_\omega[\hat{S}_{vu}]} \left(\hat{S}_{vu}(\mathrm{j}\omega_k) - G_{vu}^{\theta_{\text{lin}}}(\mathrm{j}\omega_k)\hat{S}_{uu}(\mathrm{j}\omega_k)\right)^2, \tag{11}$$

where we weight by the variances of the cross-spectral densities over frequencies. The initial guesses for estimation of the linear parameters are given in table 2.

3.3 Implementation

A Python implementation of the parameter estimation routines and the data set of this study are made available at https://github.com/fhchl/quant-comp-lsmod-ica22. We use diffrax [20, 21] to solve the initial value problem (7) and JAX [22] to compute the Jacobians of the cost functions (8,11) through automatic differentiation. If not noted otherwise, the initial value problems are solved by a fourth-order Runge-Kutta scheme (Dopri5 with constant step size $1/f_s$). The least-squares problems are solved with the algorithm of [23] with adaptive parameter scaling [24] as implemented in SciPy [25].

4 DATA SET AND EXPERIMENTAL SETUP

The models are tested on data gathered from the six driver units displayed in figure 1. Each loudspeaker was driven by two different realizations of 10 s bandpassed (10 Hz-10 kHz) Gaussian-distributed pink-noise [26] generated at 4 different gains (0 dB, -6 dB, -12 dB, -18 dB). One realisation is used for model training, while the second is used for model validation. The maximum output RMS voltages (0 dB) for each driver are listed in the caption to figure 1 and were chosen manually to be reasonably large while not damaging the loudspeaker. For each excitation we recorded the input voltage $\tilde{u}$ at the loudspeaker terminals and the two system outputs $\tilde{y}$: the current via a $0.5\,\Omega$ shunt resistor and the velocity of the diaphragm at a single point on the diaphragm via a Doppler-laser vibrometer. Each measurement was repeated 5 times over which the median was taken. The signals were created and recorded via a digital interface (National Instruments USB-4431) at a sample rate of 96 kHz. The input voltage was amplified with a low-distortion power amplifier (Purify 1ET400A).

Every model was trained separately on each training realization and used to predict both the training and the validation realizations, leading to $6 \times 4 \times 4 = 96$ predictions in the training set and 96 predictions in the validation set for each of the 23 models (10 models + 7 other orders d + 6 other solvers for the initial value problem). Errors between measurements and predictions are given as relative errors, i.e. the normalized root-mean-squared-error, in percent. Total error refers to the value where each state is normalized by its variance and then the root-mean-squared error of the concatenated states is taken. The solution of the differential equations diverged for some combination of model parameters and input voltage signals (see below). These data points were removed from the results.

5 RESULTS AND DISCUSSION

Table 3 compares the models introduced above at maximum input level. The results can be coarsely clustered into linear models with $11\% - 16\%$ error and polynomial models with around $6\% - 7\%$ error. Interestingly,

Figure 1. Loudspeakers used for data set. The maximum RMS voltages used for each driver from left to right are 4 V, 4 V, 6 V, 6 V, 6 V and 8 V.

	training error			validation error		
Model	i	$\dot{x}$	total	i	$\dot{x}$	total
full	5.8 ± 2.3	6.9 ± 2.7	6.5 ± 1.9	6.0 ± 2.7	6.2 ± 3.1	6.3 ± 2.5
full w/o $L(i)$	5.9 ± 2.5	6.9 ± 2.6	6.6 ± 2.0	6.2 ± 2.9	6.2 ± 3.1	6.4 ± 2.6
full reluctance as [13]	6.3 ± 2.2	6.8 ± 2.5	6.6 ± 2.0	6.5 ± 2.6	6.2 ± 3.0	6.4 ± 2.6
full w/o reluctance	6.6 ± 2.5	6.8 ± 2.5	6.8 ± 2.3	6.9 ± 2.8	6.1 ± 3.0	6.5 ± 2.8
full w/o SLS	5.9 ± 2.4	7.5 ± 2.2	6.9 ± 1.7	6.1 ± 2.7	6.9 ± 2.7	6.7 ± 2.3
linear SLS+L2R2	9.7 ± 0.7	11.9 ± 4.8	11.0 ± 2.9	10.0 ± 1.1	11.8 ± 4.2	11.1 ± 2.4
linear L2R2	9.3 ± 1.6	12.4 ± 4.6	11.1 ± 2.8	9.6 ± 2.0	12.3 ± 4.0	11.1 ± 2.4
full w/o L2R2	15.7 ± 1.9	7.0 ± 1.8	12.2 ± 1.3	16.0 ± 2.1	6.2 ± 2.3	12.2 ± 1.9
linear SLS	16.9 ± 1.5	12.6 ± 4.6	15.1 ± 1.9	17.2 ± 1.7	12.5 ± 4.0	15.2 ± 1.6
linear	17.1 ± 1.6	13.4 ± 4.1	15.5 ± 1.8	17.4 ± 1.8	13.3 ± 3.5	15.6 ± 1.5

Table 3. Comparison of model performance at maximum level (0 dB): relative error in percent (mean $\pm$ standard deviation over loudspeakers).

leaving out the L2R2 term leads to a much larger drop in accuracy in comparison to all other terms, which highlights the importance of modelling a lossy inductance. Including a current-dependent inductance, the reluctance force term or a simple model of viscoelasticity only has a small impact on the predictive accuracy of the full model. Figure 2 shows the averaged error spectra for the "full" model, which highlights error contributions from cone breakups at higher frequencies.

Table 4 compares the full model with varying polynomial orders d for the displacement dependency, again at maximum level. We observe a significant improvement in prediction accuracy for $d \geq 2$, but orders above $d = 2$ show practically the same error.

Table 5 compares the prediction errors over the levels of the input voltage (last column is the same as last column of table 3). For models trained at maximum level (0 dB), the relative error increases significantly when predicting at lower levels. This suggests that the current polynomial models do not generalize well over input

	training error			validation error		
Model	i	$\dot{x}$	total	i	$\dot{x}$	total
full with $d = 5$	5.8 ± 2.3	6.9 ± 2.7	6.5 ± 1.9	6.0 ± 2.7	6.2 ± 3.1	6.3 ± 2.5
full with $d = 6$	5.8 ± 2.3	6.9 ± 2.6	6.5 ± 1.9	6.0 ± 2.7	6.2 ± 3.1	6.3 ± 2.5
full ($d = 4$)	5.8 ± 2.3	6.9 ± 2.7	6.5 ± 1.9	6.0 ± 2.7	6.2 ± 3.1	6.3 ± 2.5
full with $d = 3$	5.8 ± 2.3	7.0 ± 2.6	6.6 ± 1.9	6.1 ± 2.7	6.4 ± 3.1	6.4 ± 2.5
full with $d = 2$	5.9 ± 2.3	7.1 ± 2.6	6.7 ± 1.9	6.2 ± 2.7	6.4 ± 3.1	6.4 ± 2.5
full with $d = 7$	6.6 ± 2.5	6.8 ± 2.5	6.8 ± 2.2	6.9 ± 2.8	6.2 ± 3.0	6.6 ± 2.8
full with $d = 1$	8.5 ± 1.5	11.6 ± 4.8	10.3 ± 3.1	8.7 ± 1.9	11.4 ± 4.1	10.3 ± 2.6
full with $d = 0$	9.2 ± 1.5	12.0 ± 4.8	10.8 ± 3.0	9.4 ± 1.9	11.9 ± 4.2	10.9 ± 2.5

Table 4. Comparison of polynomial order d of position dependence at maximum level (0 dB): relative error in percent (mean $\pm$ standard deviation over loudspeakers).

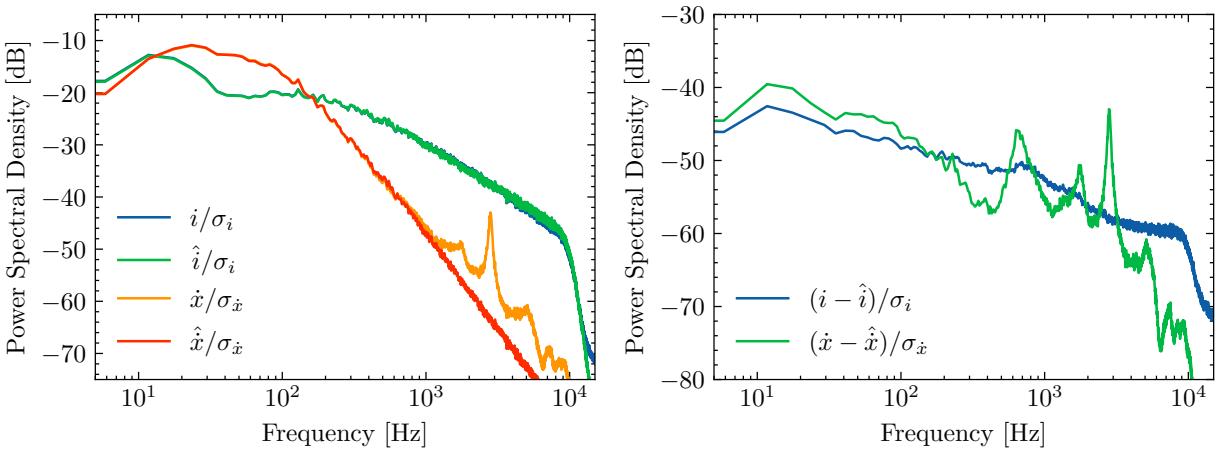

Figure 2. Power spectral densities of normalized measurements $(i,\dot{x})$, predictions $(\hat{i},\hat{\dot{x}})$ and errors $(i-\hat{i},\dot{x}-\hat{\dot{x}})$ for model "full" averaged over loudspeakers at maximum level (i.e. spectra on the left integrate to 1 and spectra on the right integrate to validation error of table 3).

	total validation error				
training level	-18 dB	0 dB			
validation level	-18 dB	-18 dB	-12 dB	-6 dB	0 dB
model					
full	5.4 ± 2.7	11.6 ± 3.6	8.7 ± 2.1	8.0 ± 3.1	6.3 ± 2.5
full w/o $L(i)$	5.9 ± 3.8	11.7 ± 3.6	8.8 ± 2.0	8.0 ± 3.1	6.4 ± 2.6
full reluctance as [13]	5.4 ± 2.7	11.7 ± 3.6	8.8 ± 2.1	8.1 ± 3.2	6.4 ± 2.6
full w/o reluctance	5.4 ± 2.7	11.8 ± 3.6	9.0 ± 2.2	8.3 ± 3.2	6.5 ± 2.8
full w/o SLS	6.2 ± 1.9	12.0 ± 3.5	9.1 ± 1.9	8.4 ± 3.0	6.7 ± 2.3
linear SLS+L2R2	5.6 ± 2.6	9.9 ± 2.5	8.9 ± 3.3	9.7 ± 5.7	11.1 ± 2.4
linear L2R2	6.3 ± 1.7	10.1 ± 2.4	9.1 ± 3.2	9.9 ± 5.6	11.1 ± 2.4
full w/o L2R2	11.9 ± 1.6	16.4 ± 3.4	14.1 ± 2.1	13.8 ± 1.6	12.2 ± 1.9
linear SLS	12.1 ± 1.4	14.7 ± 1.4	14.1 ± 1.7	14.8 ± 3.9	15.2 ± 1.6
linear	13.0 ± 1.4	15.2 ± 1.4	14.7 ± 1.8	15.3 ± 3.9	15.6 ± 1.5

Table 5. Comparison of input levels during training and prediction: relative error in percent (mean $\pm$ standard deviation over loudspeakers).

level ranges. Note that this does not imply an increase in absolute error. Low errors at low input levels can be obtained by training the models on low level data (leftmost column).

Training at low and predicting at high input levels can lead to divergent solutions of the system of differential equations. This observation is summarized in table 6 which shows the number of divergences over training and prediction level. Generally, for polynomial models, training at high and predicting at low input levels leads to stable predictions, while doing the opposite can lead to unstable predictions. Linear models never diverged in this study.

Table 7 compares different differential equation solvers implemented by [21] for the "full" model. They were used here with a constant step size $1/f_s$ due to computational reasons. Higher-order Runge-Kutta solvers tend to give better results compared to Euler discretization, but the difference is not large at a signal bandwidth of 10 Hz-10 kHz at a sampling rate of 96 kHz. At larger bandwidths or lower sampling rate, the Euler method gave considerably worse results (not shown) and we advice to use higher-order solvers in such cases.

level at model prediction level at model training	-18 dB	-12 dB	-6 dB	0 dB
-18 dB	1	19	59	76
-12 dB	0	1	22	50
-6 dB	0	0	0	24
0 dB	0	0	0	1

Table 6. Total number of diverging model predictions over input voltage level during training and prediction: models fitted at low input level can diverge when predicting at large input levels. In total, 311 of 2304 predictions diverged, all of them from nonlinear models. The model "full with $d = 7$" diverged most often (30 times) while "full with $d = 1$" diverged least often (3 times).

Solver for model "full"	total training error	total validation error
Dopri5	5.7 ± 2.1	8.4 ± 4.4
Tsit5	5.8 ± 2.4	8.6 ± 4.6
Heun	5.6 ± 2.5	8.6 ± 4.7
Ralston	5.7 ± 2.6	8.6 ± 4.7
Midpoint	5.8 ± 2.6	8.8 ± 4.7
Bosh3	6.0 ± 2.7	8.9 ± 4.7
Euler	6.8 ± 3.0	9.2 ± 4.2

Table 7. Comparison of solvers implemented in [21] for initial value problem (7) with *constant* step size: relative error in percent (mean $\pm$ standard deviation over loudspeakers, training and validation levels).

6 CONCLUSION

This work compared lumped-element models for loudspeakers on a small data set. For this data set, the results suggest that including simple models for the effect of current-dependent inductance, the reluctance force or viscoelasticity of the suspension only lead to small improvements of predictive accuracy. The generalized Cunningham force of [11] did not lead to a significant change in prediction error in comparison to the previous model. The L2R2 model of a lossy inductance, on the other hand, improves errors similarly well as including nonlinear stiffness, inductance, and force factor. A larger than quadratic order of the displacement dependence did not improve results significantly. Also, existing nonlinear loudspeaker models do not generalize well over changing input levels.

As a community, in order to make effective progress on the predictive performance of loudspeaker models, we need quantitative benchmarks on a shared and public data set. To this end, we have published implementations of the models and fitting procedures of this work together with the data set as an invitation for fellow researchers to collaborate on this enterprise.

ACKNOWLEDGEMENTS

The authors would like to thank Manuel Hahmann for giving valuable feedback to the manuscript, to Huawei for funding this research, and to the Korean and international acoustic communities for organizing this conference.

REFERENCES

[1] Wolfgang Klippel. “The Mirror Filter-A New Basis for Reducing Nonlinear Distortion and Equalizing Response in Woofer Systems”. In: *J. Audio Eng. Soc.* 40.9 (1992), pp. 675–691.

[2] Wolfgang Klippel. “Direct Feedback Linearization of Nonlinear Loudspeaker Systems”. In: *J. Audio Eng. Soc.* 46.6 (1998), pp. 499–507.

[3] Hans Schurer, Cornelis H. Slump, and Otto E. Herrmann. “Exact Input-Output Linearization of an Electrodynamical Loudspeaker”. In: Proc. Audio Eng. Soc. Conv. 101. 1996.

[4] Xing Tian et al. “Compensation of Nonlinear Distortion in Loudspeakers Considering Nonlinear Viscoelasticity of the Suspension”. In: *J. Audio Eng. Soc.* 69.3 (2021), pp. 204–210.

[5] Jean-Jacques E. Slotine and Weiping Li. *Applied Nonlinear Controlhttps://*. Englewood Cliffs, N.J: Prentice Hall, 1991.

[6] Antonin Novak. “Modeling Viscoelastic Properties of Loudspeaker Suspensions Using Fractional Derivatives”. In: *J. Audio Eng. Soc.* 64.1/2 (2016), pp. 35–44.

[7] Finn T. Agerkvist. “Non-Linear Viscoelastic Models”. In: Proc. Audio Eng. Soc. Conv. 131. 2011.

[8] Tobias Ritter and Finn T. Agerkvist. “Modelling Viscoelasticity of Loudspeaker Suspensions Using Retardation Spectra”. In: Proc. Audio Eng. Soc. Conv. 129. 2010, p. 11.

[9] W. Marshall Leach Jr. “Loudspeaker Voice-Coil Inductance Losses: Circuit Models, Parameter Estimation, and Effect on Frequency Response”. In: *J. Audio Eng. Soc.* 50.6 (2002), pp. 442–450.

[10] W. J. Cunningham. “Non-Linear Distortion in Dynamic Loudspeakers Due to Magnetic Effects”. In: *J. Acoust. Soc. Am.* 21.3 (1949), pp. 202–207.

[11] Lars Risbo et al. “Force Factor Modulation in Electro Dynamic Loudspeakers”. In: Proc. Audio Eng. Soc. Conv. 141. 2016.

[12] Alexander W. King and Finn T. Agerkvist. “Position Dependence of Fractional Derivative Models for Loudspeaker Voice Coils with Lossy Inductance”. In: Proc. Audio Eng. Soc. Conv. 142. 2017.

[13] Wolfgang Klippel. “Tutorial: Loudspeaker Nonlinearities - Causes, Parameters, Symptoms”. In: *J. Audio Eng. Soc.* 54 (2006), pp. 907–939.

[14] Alexander W. King and Finn T. Agerkvist. “Fractional Derivative Loudspeaker Models for Nonlinear Suspensions and Voice Coils”. In: *J. Audio Eng. Soc.* 66.7/8 (7/8 2018), pp. 525–536.

[15] Alexander W. King. “Nonlinear Fractional Order Derivative Models of Components and Materials in Hearing Aids and Transducers”. PhD thesis. Lyngby: Technical University of Denmark, 2019.

[16] R. M. Christensen. *Theory of Viscoelasticity*. 2nd ed. New York: Academic Press, 1982.

[17] Mark Dodd, Wolfgang Klippel, and J. Oclee-Brown. “Voice Coil Impedance as a Function of Frequency and Displacement”. In: *Proc. Audio Eng. Soc. Conv. 119* (2004), pp. 1–17.

[18] Lennart Ljung. *System Identification*. 2nd ed. Upper Saddle River, NJ: Prentice Hall, 1999.

[19] P. Welch. “The Use of Fast Fourier Transform for the Estimation of Power Spectra”. In: *IEEE Trans. Audio Electro.* 15 (2 1967), pp. 70–73.

[20] Patrick Kidger. “On Neural Differential Equations”. Feb. 4, 2022. arXiv: 2202.02435. URL: http://arxiv.org/abs/2202.02435 (visited on 03/03/2022).

[21] Patrick Kidger. *Diffrax*. URL: https://github.com/patrick-kidger/diffrax (visited on 08/13/2022).

[22] James Bradbury et al. *JAX: Composable Transformations of Python+NumPy Programs*. Version 0.2.5. 2018. URL: http://github.com/google/jax.

[23] Mary Ann Branch, Thomas F. Coleman, and Yuying Li. “A Subspace, Interior, and Conjugate Gradient Method for Large-Scale Bound-Constrained Minimization Problems”. In: *SIAM J. Sci. Comput.* 21.1 (Jan. 1999), pp. 1–23.

[24] Jorge J Moré. “The Levenberg-Marquardt Algorithm: Implementation and Theory”. In: *Numerical Analysis*. Springer, 1978, pp. 105–116.

[25] Pauli Virtanen et al. “SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python”. In: *Nat Methods* 17.3 (2020), pp. 261–272.

[26] Jens Timmer and M Koenig. “On Generating Power Law Noise.” In: *Astronomy and Astrophysics* 300 (1995), p. 707.