AES EUROPE 2024

MADRID

June 15-17 | Universidad Politécnica de Madrid, Spain

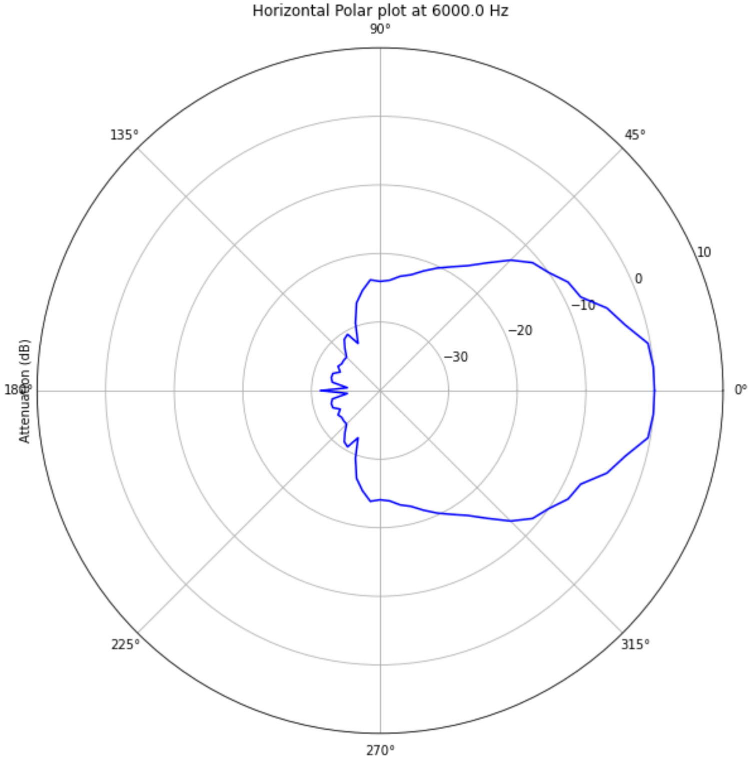
Interpolation of loudspeaker level balloons from polar measurements by using deep learning

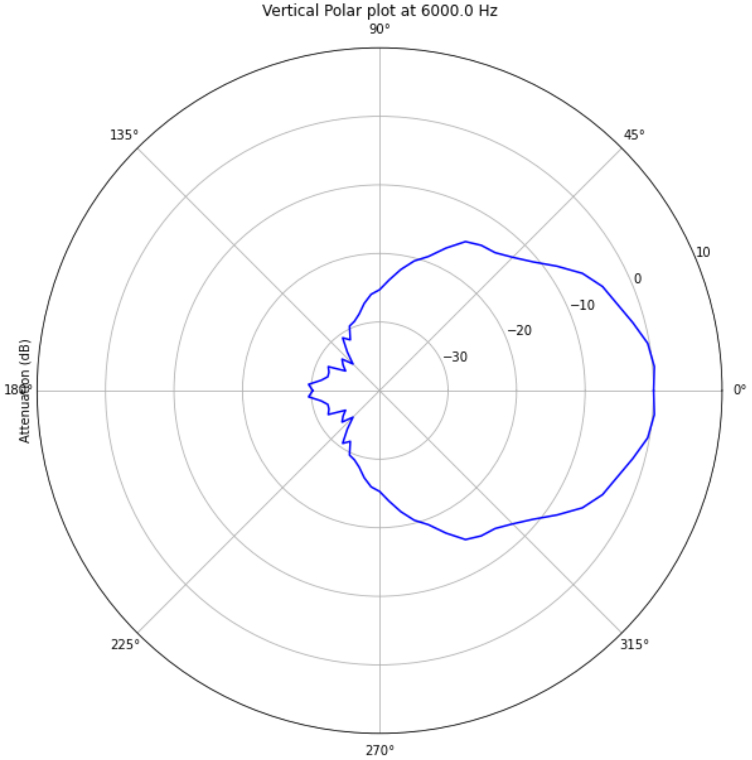
Víctor Manuel Catalá Iborra[1]

[1] *DAS AUDIO GROUP, Fuente del Jarro, Spain*

AES EUROPE 2024

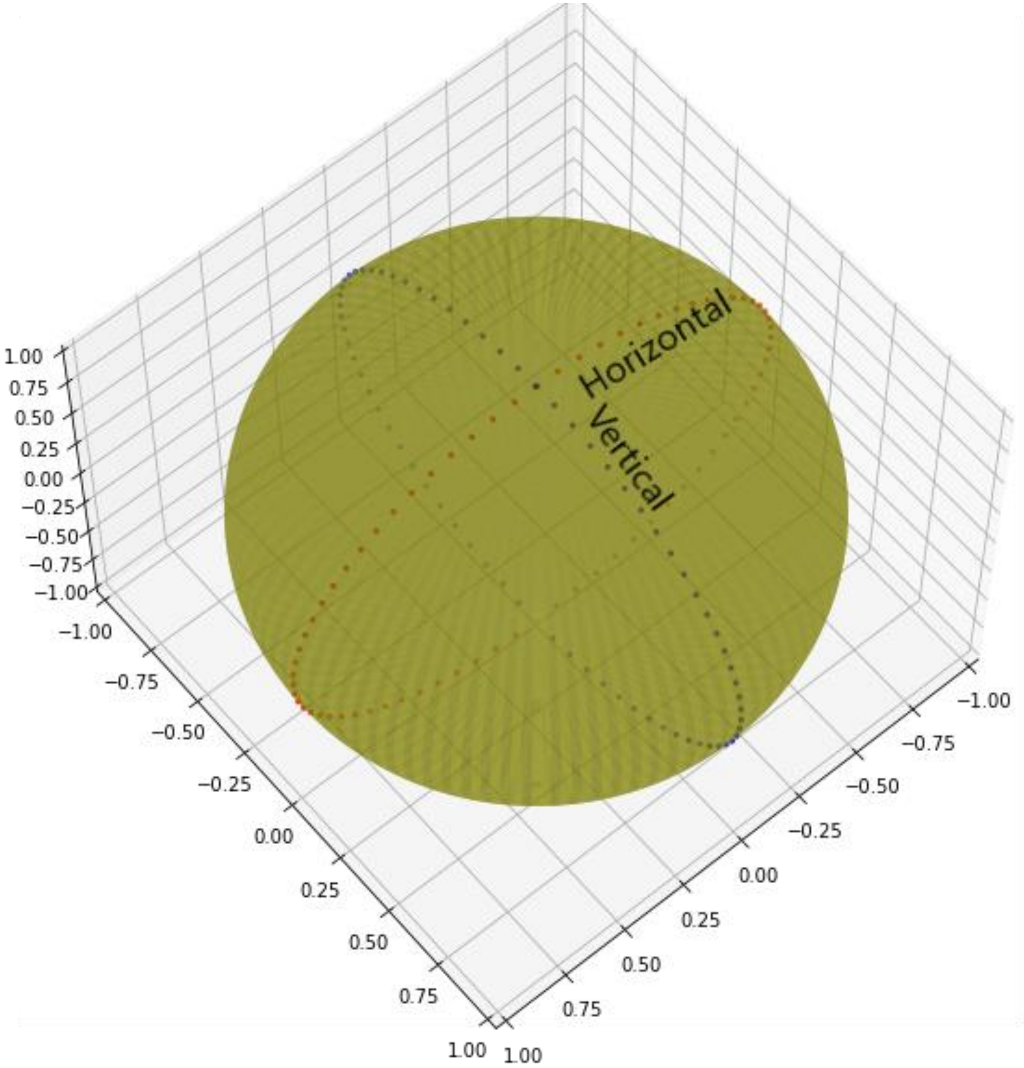
Motivation





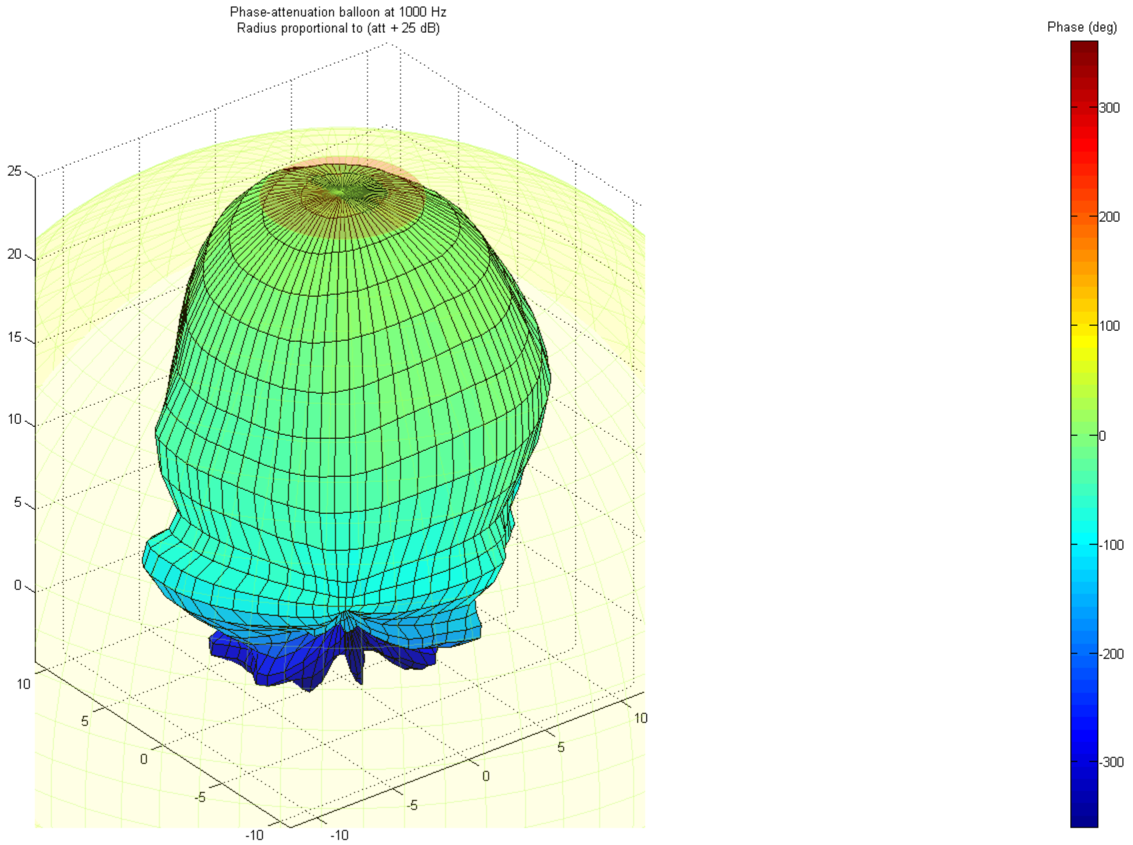
Polar diagrams.

Motivation



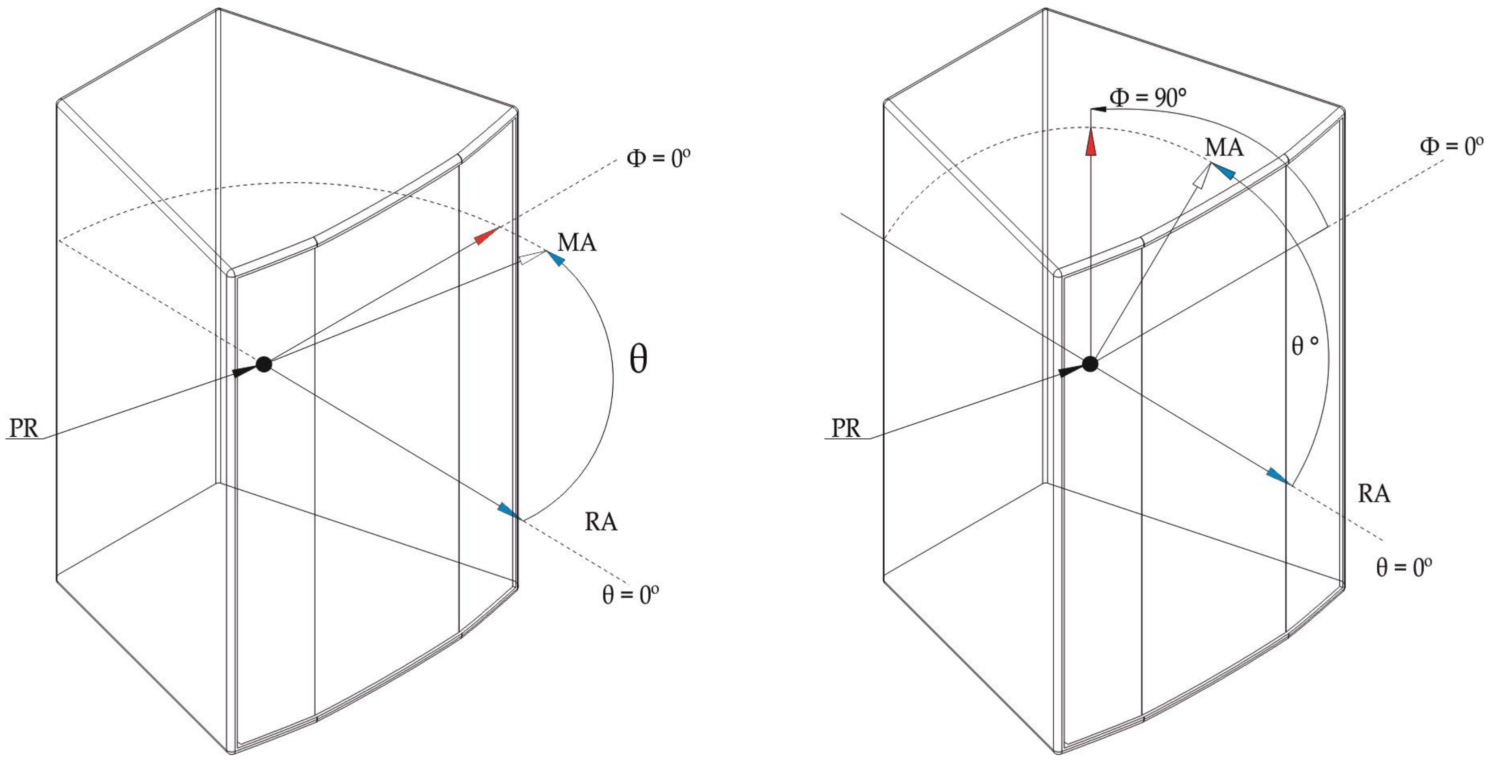
Horizontal (red) and vertical (blue) polar measurement points.

Motivation



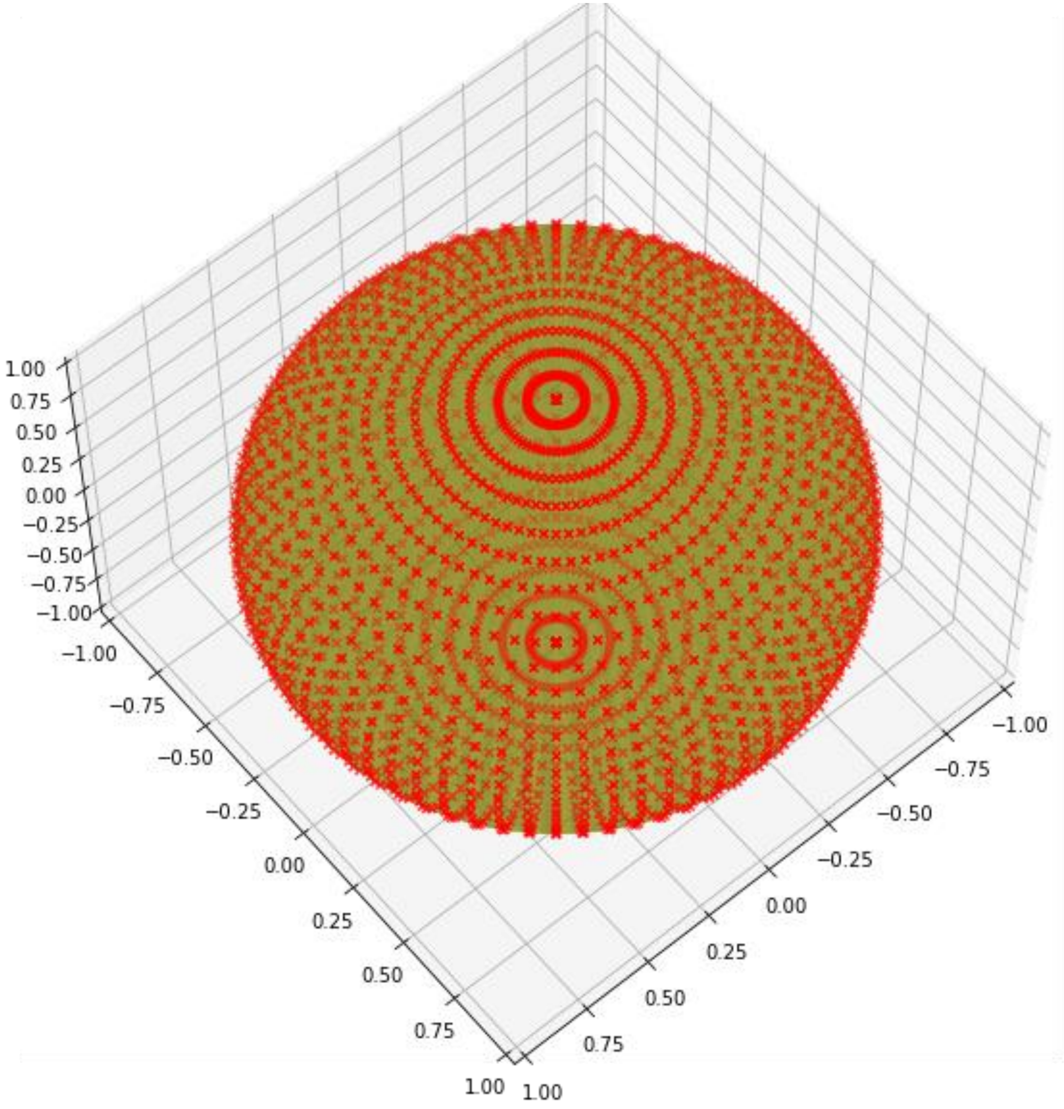
Loudspeaker radiation balloon (level and phase)

Motivation



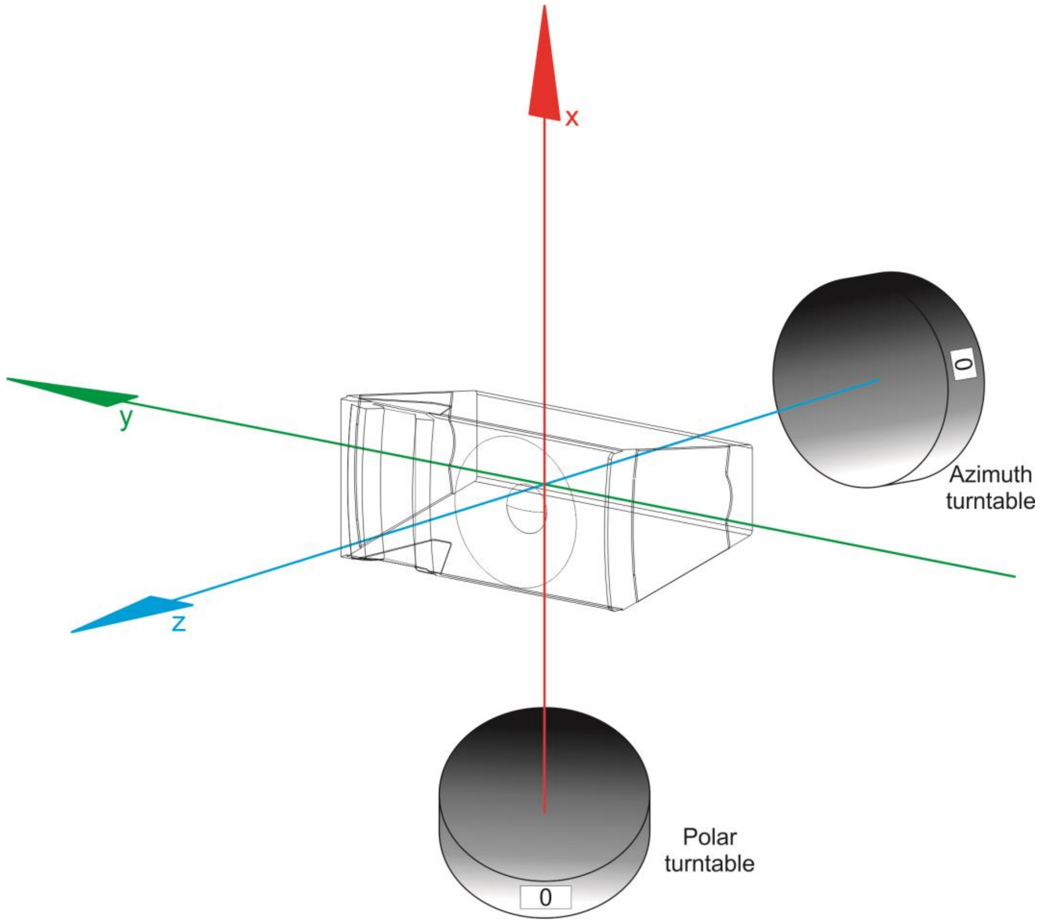
Polar system coordinates per AES, 2008 (reaffirmed 2014).

Motivation



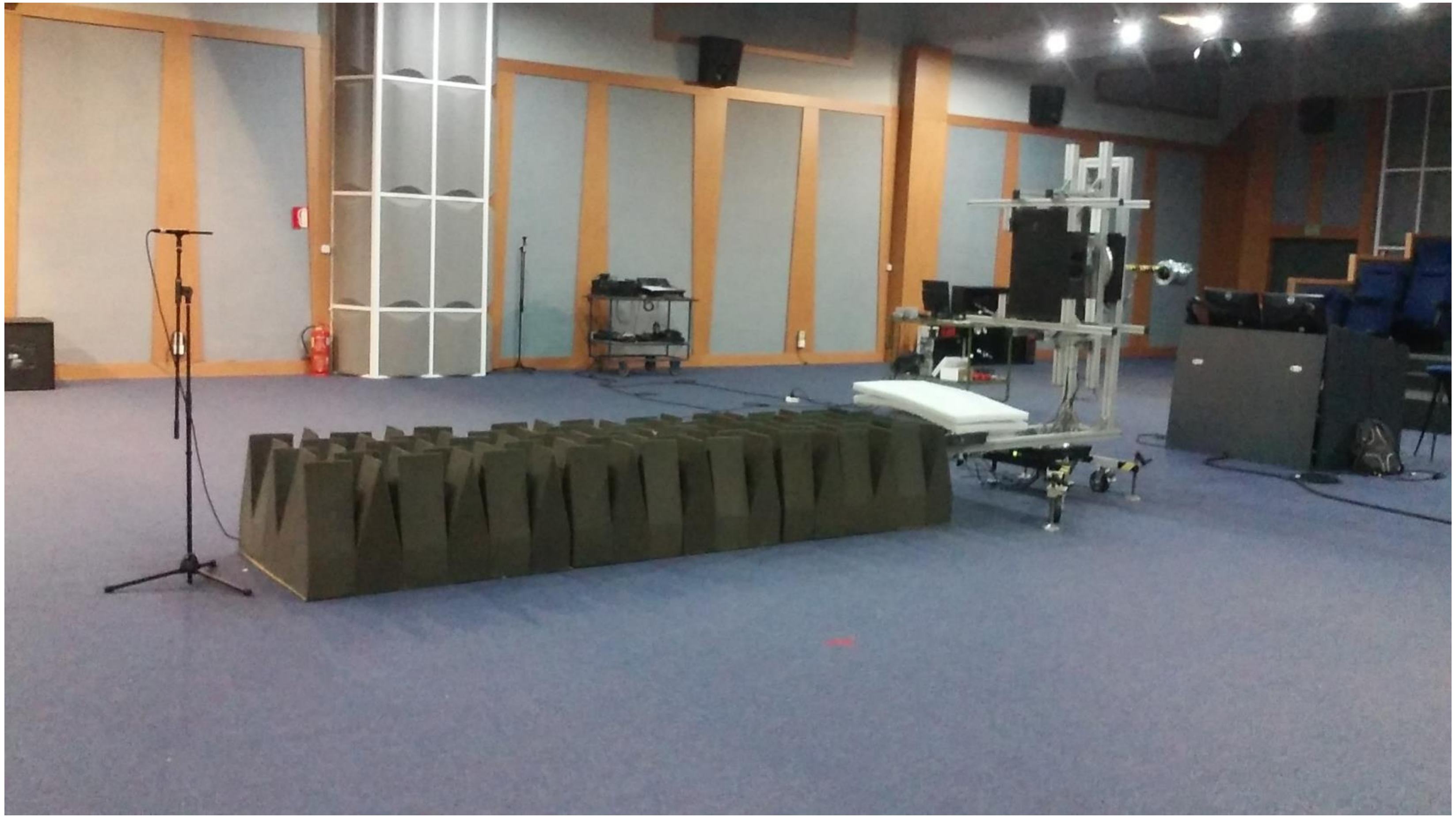
2664 measurement points which means 82584 SPL values with 1/3 oct resolution

Motivation



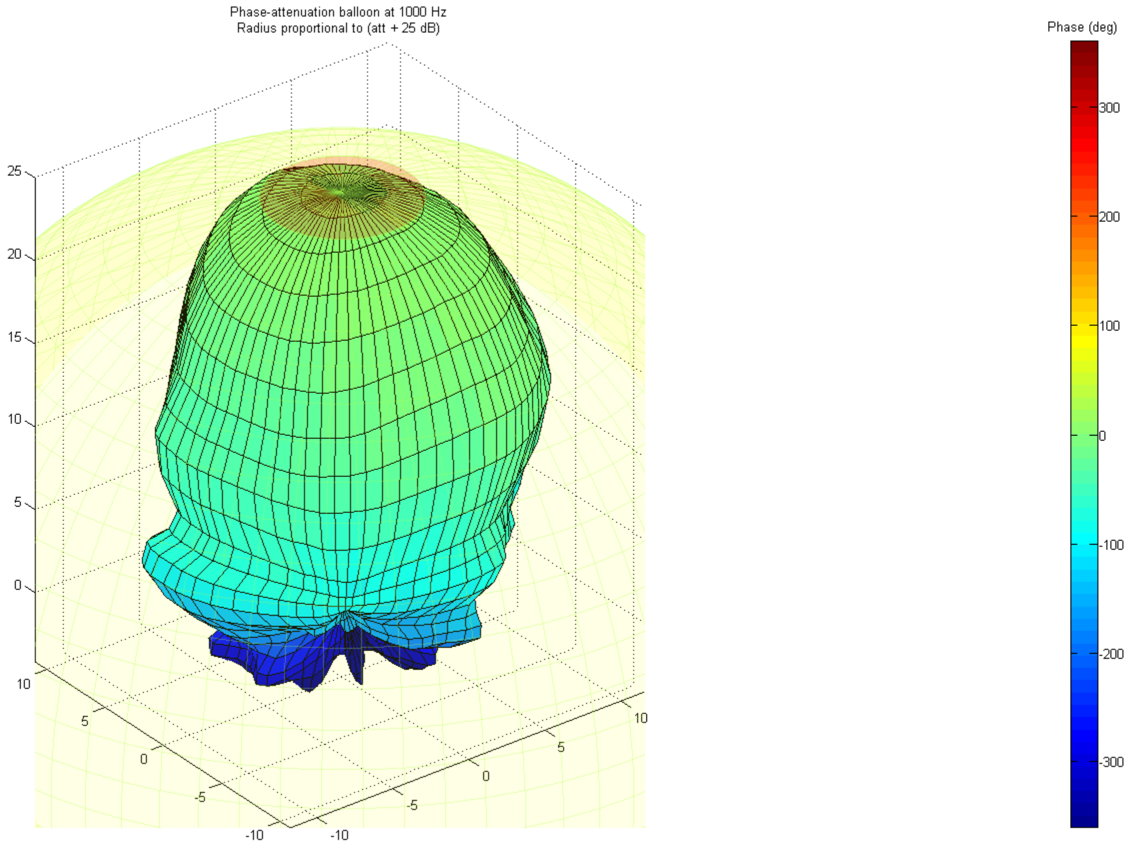
3D acquisition system with 2 turntables.

Motivation



Measurement set-up at DAS Audio Auditorium.

Motivation



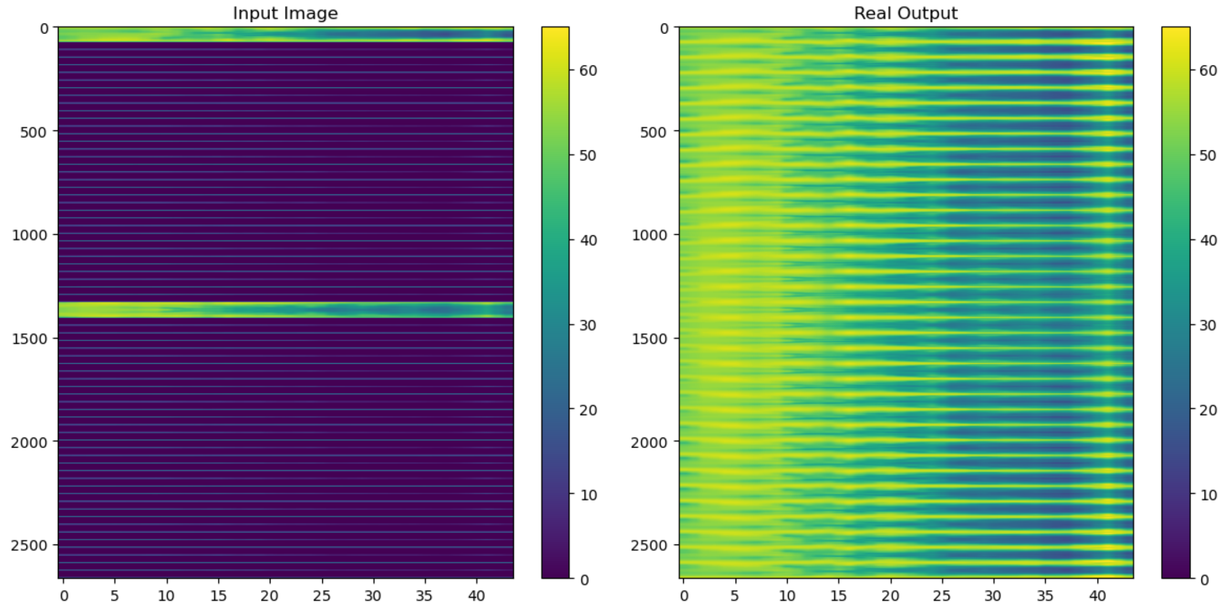
Loudspeaker radiation balloon (level and phase)

Method

- Traditional interpolation methods are based on the data of the loudspeaker being analyzed

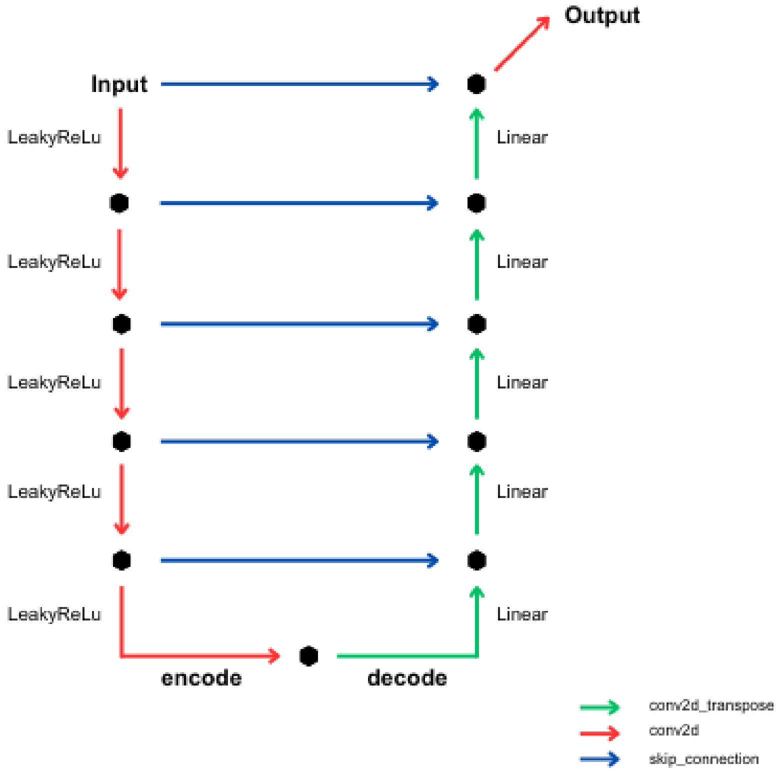
- Neural networks allow us to take profit of many loudspeaker data to build a model and then, apply it to a specific loudspeaker.

Method



The problem is reduced to an incomplete-image recovery problem

Method



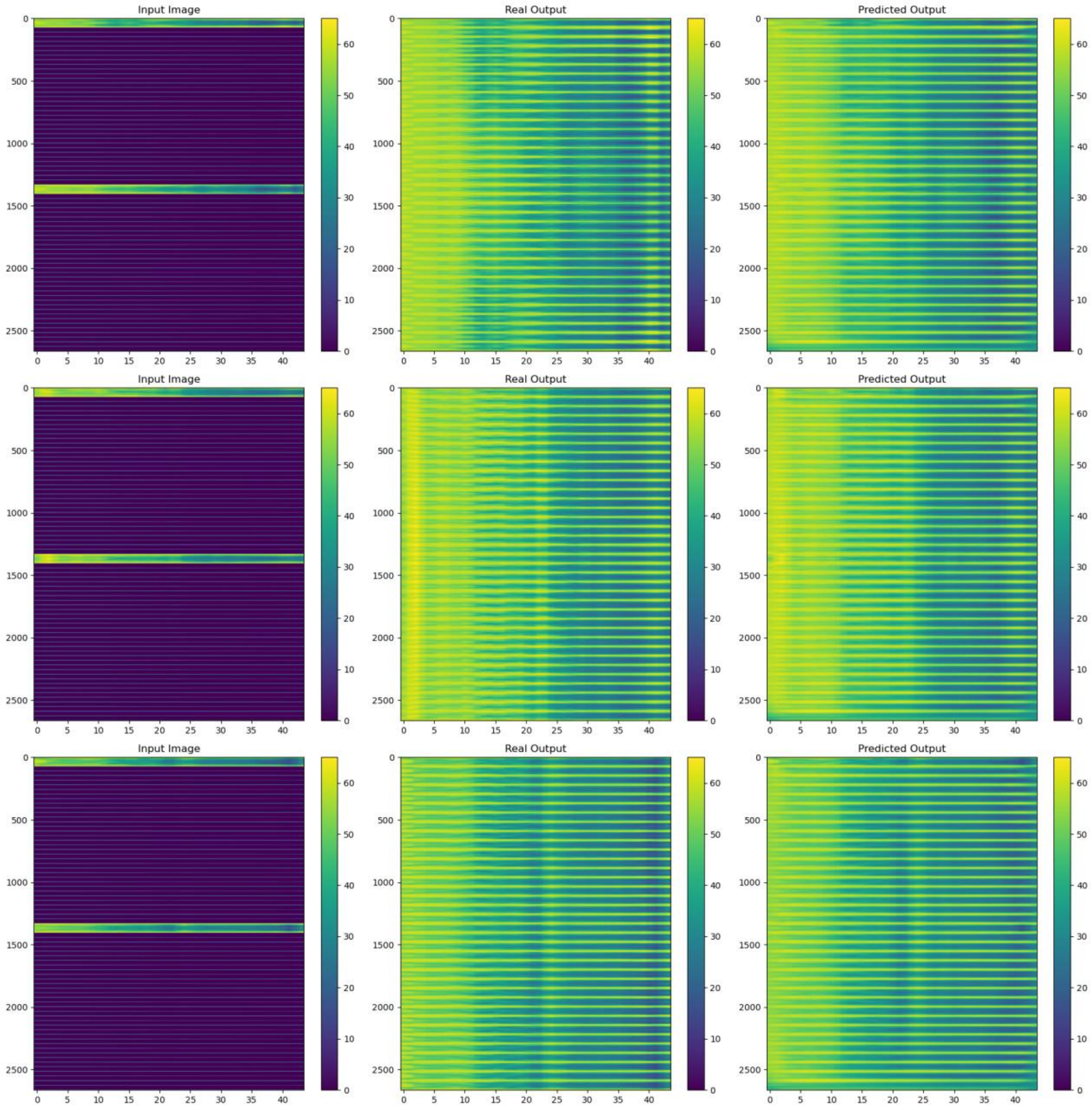
Proposed U-net architecture

Method

Layer	Input size	Stride	Padding	Kernel size	Output size
conv2d_44	(2664, 44, 1)	(2, 1)	valid	(75, 3)	(1295, 42, 16)
conv2d_45	(1295, 42, 16)	(2, 2)	valid	(75, 3)	(611, 20, 32)
conv2d_46	(611, 20, 32)	(2, 2)	valid	(75, 3)	(269, 9, 64)
conv2d_47	(269, 9, 64)	(2, 2)	valid	(75, 3)	(98, 4, 128)
conv2d_48	(98, 4, 128)	(2, 2)	valid	(75, 3)	(12, 1, 256)
conv2d_transpose_32	(12, 1, 256)	(2, 1)	valid	(76, 4)	(98, 4, 128)
conv2d_transpose_33	(98, 4, 256)	(2, 2)	valid	(75, 3)	(269, 9, 64)
conv2d_transpose_34	(269, 9, 128)	(2, 2)	valid	(75, 4)	(611, 20, 32)
conv2d_transpose_35	(611, 20, 64)	(2, 2)	valid	(75, 4)	(1295, 42, 16)
conv2d_transpose_36	(1295, 42, 32)	(2, 1)	valid	(76, 3)	(2664, 44, 16)
conv2d_49	(2664, 44, 17)	NA	same	(75, 4)	(2664, 44, 1)

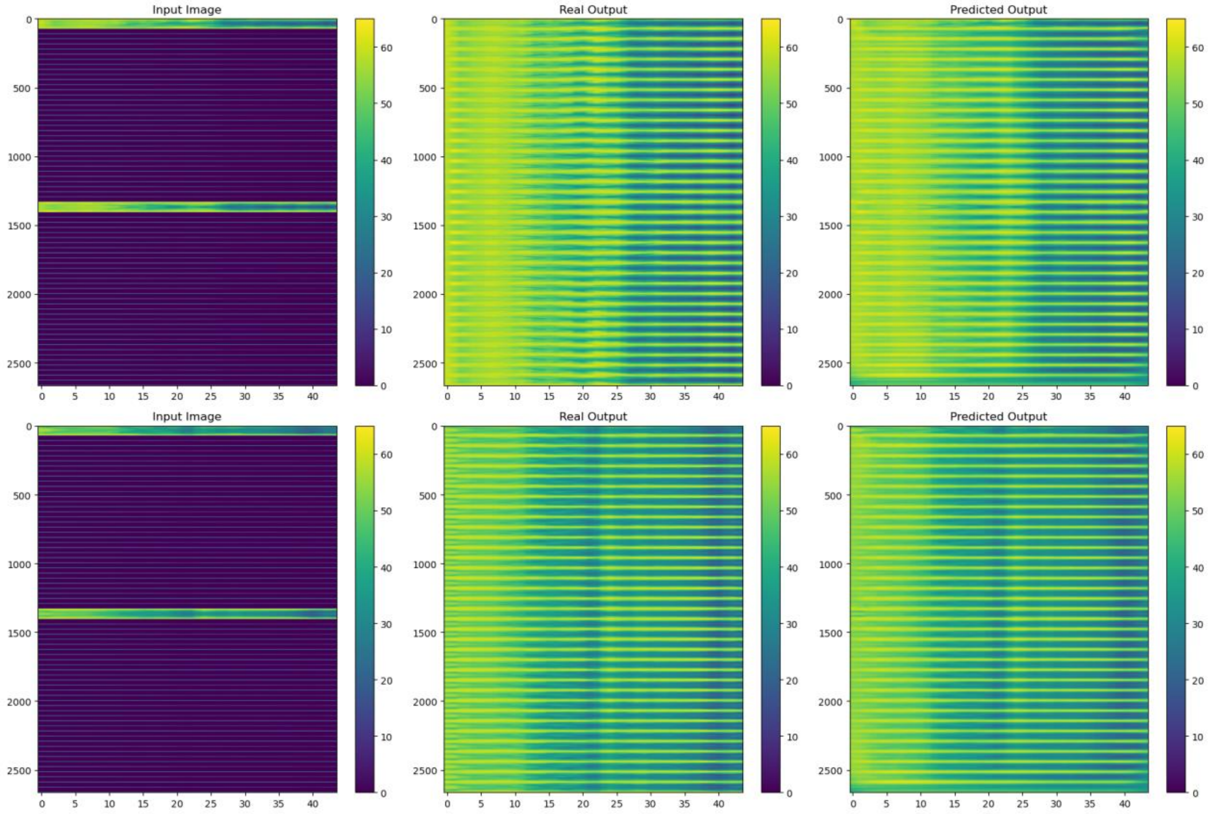
Parameters of the proposed U-net.

Results



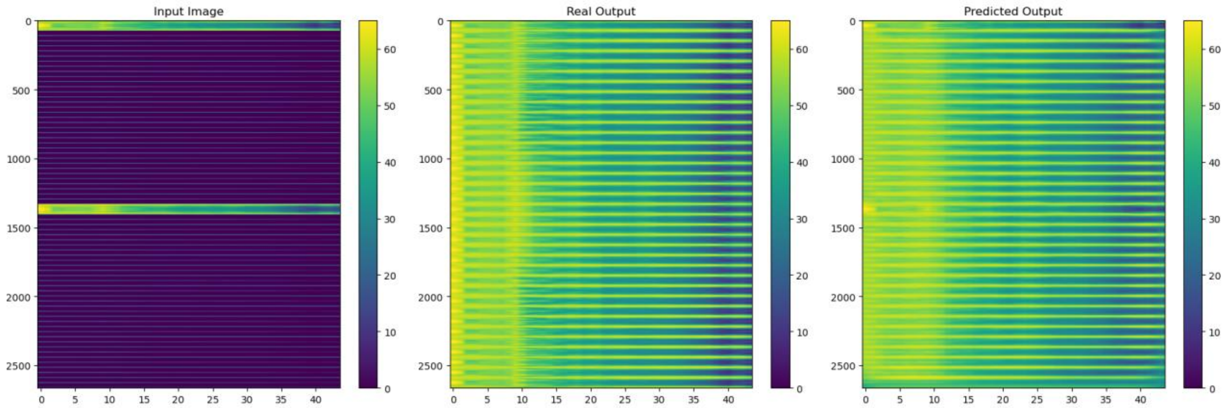
Training data images 1-2-3. Left column: only polar data. Central column: ground truth. Right: U-net prediction

Results



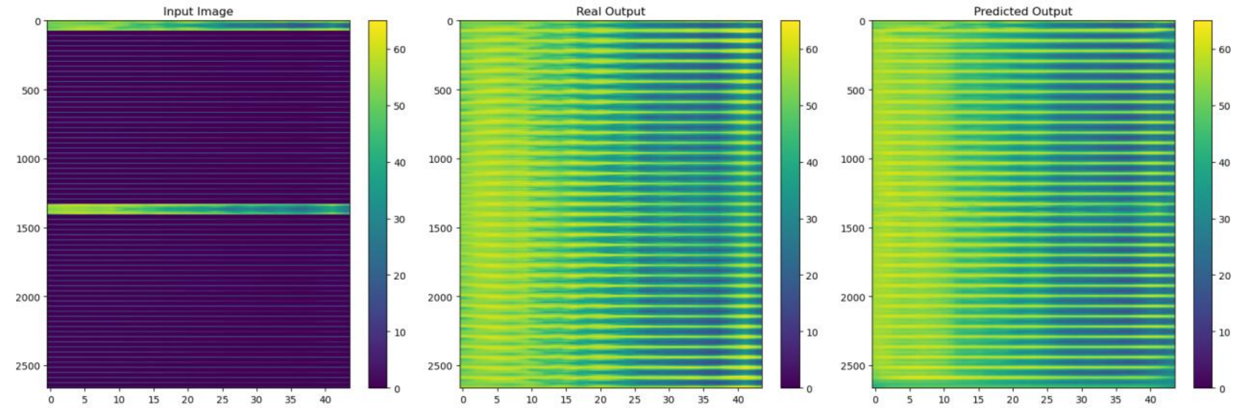
Training data images 4-5. Left column: only polar data. Central column: ground truth. Right: U-net prediction

Results



Validation data. Left column: only polar data. Central column: ground truth. Right: U-net prediction

Results

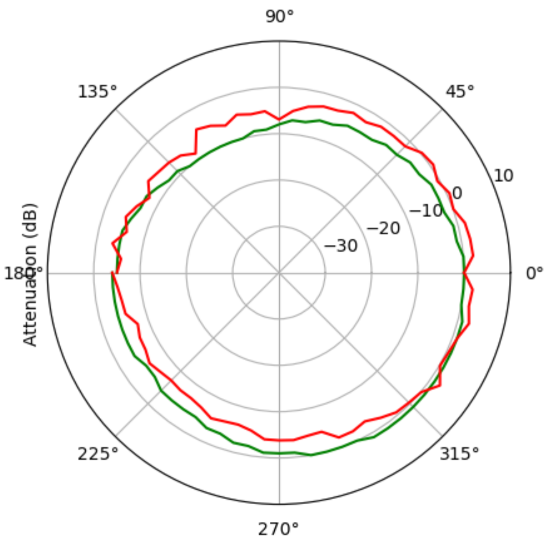


Test data. Left column: only polar data. Central column: ground truth. Right: U-net prediction

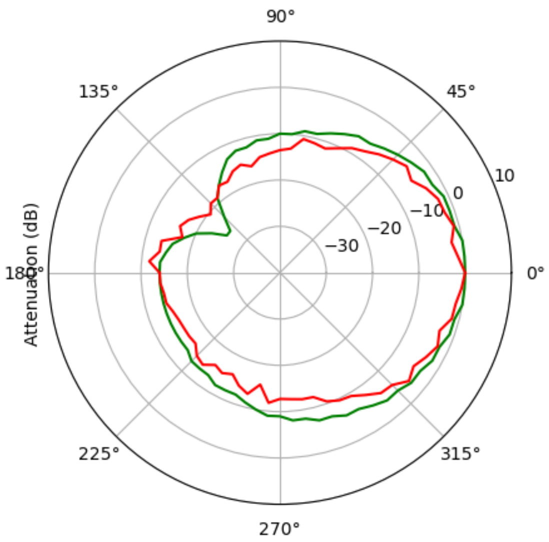
Results

	Mean Absolute Error (MAE)
Training data	2.8 dB
Validation data	4.0 dB
Test data	3.8 dB

Results

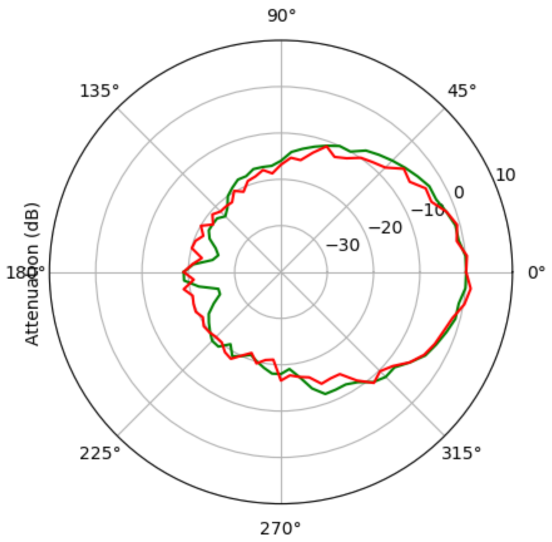


Polar plot at Phi=45º - 250 Hz

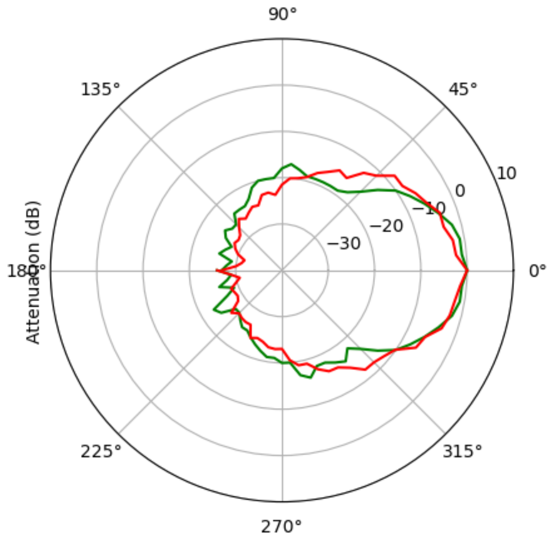


Polar plot at Phi=45º - 500 Hz

Results

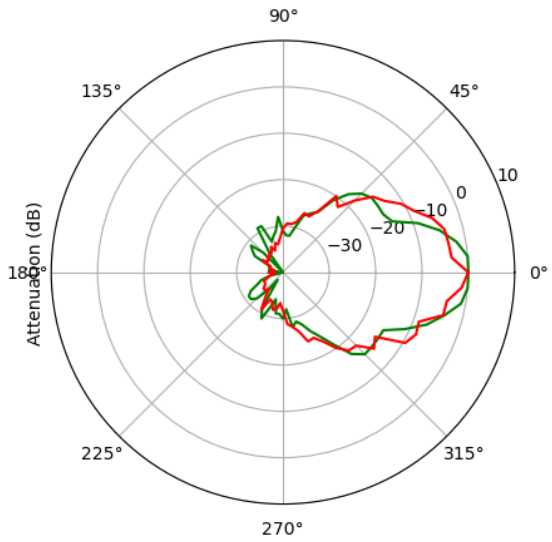


Polar plot at Phi=45º - 1 kHz

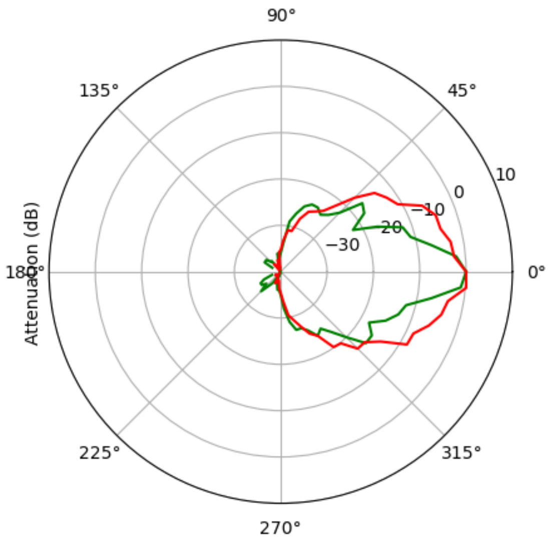


Polar plot at Phi=45º - 2 kHz

Results

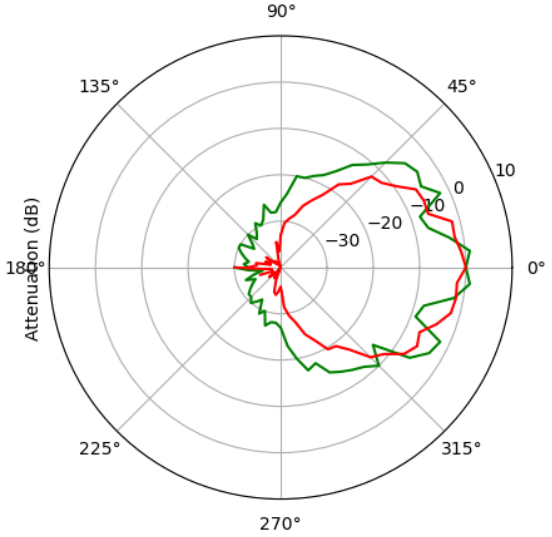


Polar plot at Phi=45º - 4 kHz



Polar plot at Phi=45º - 8 kHz

Results



Polar plot at Phi=45º - 16 kHz

Conclusions and further research

- Neural Networks can be used to approximate level loudspeaker radiation balloons from polar measurements. In particular, U-net deep learning architectures can provide MAE lower that 4 dB over full balloons.

- At high frequencies, with steeper variation of SPL against angles, the performance of the proposed network is lower. Refined architectures, more data and farther research are needed to improve accuracy.

- Similar methods could be applied to phase interpolation in order to get complex responses.

- More powerful models could be achieved by combining true polar measurements and full 3D simulations (from Finite Element Method software, for example) as input data, to obtain more accurate predictions. In this case, simple polar measurements would refine full 3D simulations.

Thanks!