Calibration of the QESTRAL model for the prediction of spatial quality

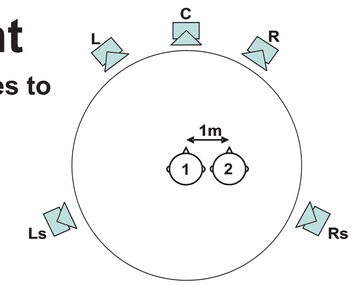
R. Conetta, M. Dewhirst, F. Rumsey, S. Zielinski, P.J.B Jackson, S. Bech, D. Meares and S. George

r.conetta@surrey.ac.uk

www.surrey.ac.uk/soundrec/QESTRAL

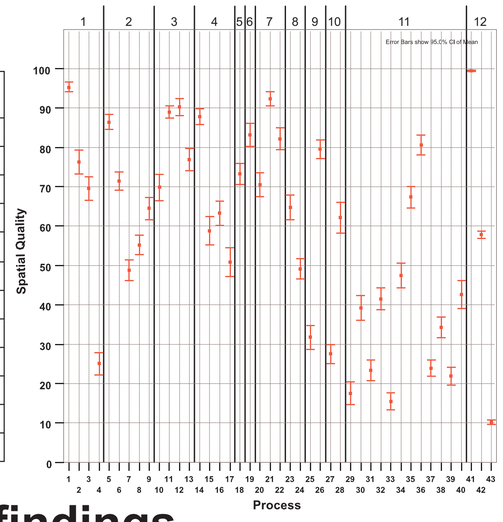
Listening experiment

- Designed to collect data on changes to spatial quality (SQ).
- 40 audio processes
- 3 x 5-channel programme items (e.g. Sport, Classical, Pop)
- 2 listening positions



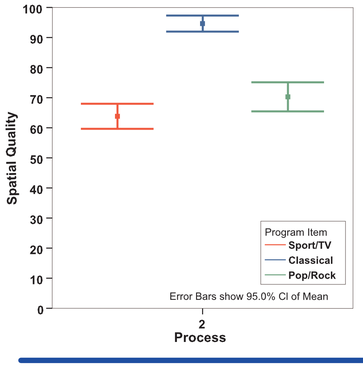
Results

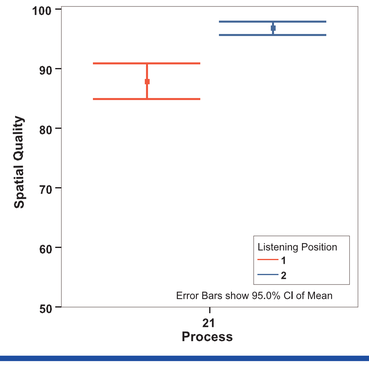
	Process groups
1	Down-mixing from 5 CH
2	Audio coding
3	Loudspeaker mis-placement
4	Channel routing errors
5	Inter-channel level mis-alignment
6	Inter-channel out-of-phase errors
7	Missing channels
8	Filtering
9	Inter-channel crosstalk
10	Virtual surround algorithms
11	Combinations of 1-10
12	Scale anchors



Summary of findings

- Audio process had the largest effect on perceived SQ.
- Programme item type had an effect on perceived SQ.
- Listening position also had an effect on perceived SQ.
- Wide and multi-modal distributions suggest that listeners found it difficult to assess some stimuli.





Prediction models

- Using the recommendations from the listening test analysis models were created for each listening position and programme material type.
- Stimuli with ambiguous means were removed.
- Some processes were also removed.

Probe signals

Probe signal	Scene	Description
1	Foreground	Sequence of 36 pink noise bursts panned at 10° Intervals.
2	Background	Decorrelated pink noise played simultaneously through each loudspeaker.

Objective metrics

Metric	Description	Probe signal
Mean_Ang	The mean absolute change to the angles calculated using the QESTRAL directional localisation model.	1
IACC0	The broadband mean value of IACC calculated with a 0° head orientation.	2
Card_KLT	The contribution in percent of the first eigenvector from a Karhunen-Loeve Transform (KLT) decomposition of four cardioid microphones placed at the listening position and facing in the following directions: 0°, 90°, 180° and 270°.	

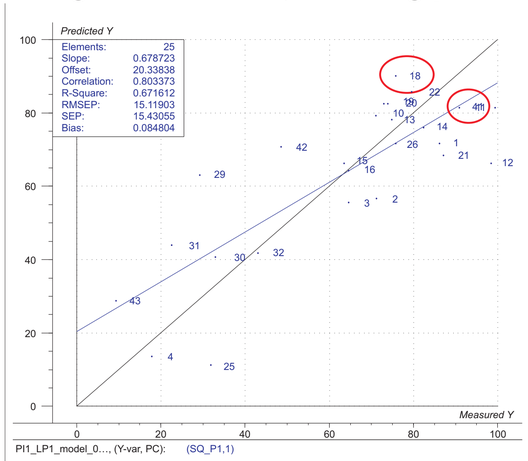
Results and discussion

- In some cases the prediction power shows an improvement over models created only for listening position.

Listening position	Prediction	
	R	Error (%)
1	0.73	17.05
2	0.80	15.80

Listening position	Programme Item	Prediction	
		R	Error (%)
1	1	0.80	15.12
	2	0.66	23.47
	3	0.76	17.25
2	1	0.84	14.87
	2	0.75	20.40
	3	0.83	14.79

Programme item 1, listening position 1

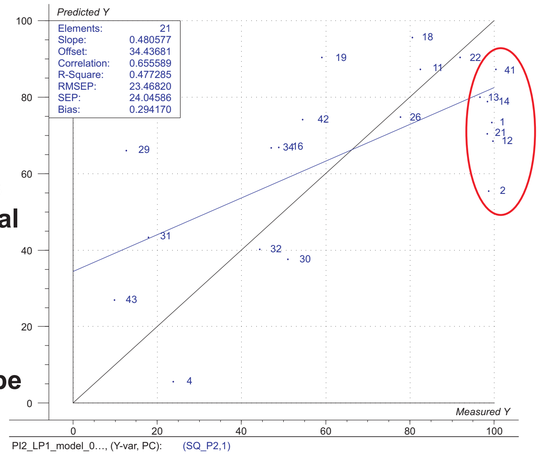


Metric	BW	B
IACC0	0.33	78.29
CardKLT	0.31	0.55
Mean_Ang	-0.31	-0.21
Constant	3.23	83.54

- Process 18 in which the front channels are 6dB lower in level than the rear channels, is predicted above the hidden reference (41).
- None of the metrics used in the model can measure level changes.

Programme item 2, listening position 1

Metric	BW	B
IACC0	0.33	78.29
CardKLT	0.31	0.55
Mean_Ang	-0.31	-0.21
Constant	3.23	83.54



- The circled processes are subjectively identical for this type of programme material.
- The metrics have measured a difference, due to the types of probe signal used.

Conclusions

- Creating models for different listening positions and programme material type improves the prediction in certain cases.
- Not fixing the selection of metrics used in the models may yield improved prediction.
- New metrics should be developed in order to predict processes that are not currently predicted.
- New probe signals could also be developed that show greater similarity to different types of programme material.

BANG & OLUFSEN

BBC

UNIVERSITY OF SURREY