Uncertainty of glottal airflow estimation during continuous speech using impedance-based inverse filtering of the neck-surface acceleration signal

Víctor M. Espinoza[1,4], Daryush D. Mehta[2,3], Jarrad H. Van Stan[2], Robert E. Hillman[2,3], Matías Zañartu[1]

[1]Universidad Técnica Federico Santa María, Valparaiso, Chile.
[2]Massachusetts General Hospital, Center for Laryngeal Surgery and Voice Rehabilitation, Boston, USA.
[3]Harvard University, Harvard Medical School, Boston, USA.
[4]Universidad de Chile, Santiago, Chile.

Introduction

- Currently, we are studying the vocal behavior of individuals with typical voices and voice disorders by analyzing weeklong recordings using a smartphone-based ambulatory voice monitor. An Impedance-based Inverse Filtering (IBIF) algorithm is used to estimate the glottal (airflow) volume velocity (GVV) from a neck-surface acceleration signal. However, the IBIF estimation is obtained performing a sustained vowel /a/, which is a different scenario than continuous speech. In this work we explore the performance of IBIF in continuous speech and his impact in vocal measures of clinical relevance.

Goal

- **To determine the uncertainty of non-invasive glottal aerodynamic measures that are obtained using subglottal impedance-based inverse filtering (IBIF) of the signal from a neck-placed accelerometer during continuous speech.**

Methods

- Simultaneous and synchronous recordings of oral volume velocity (OVV) and neck skin acceleration (ACC) signals were performed by two adult females, one with vocal hyperfunction (polyp) and her matched control with normal vocal status.
- A band-pass filtered version of OVV signal was inverse filtered based on minimizing the formant ripple [1].
- Q factors related to the IBIF model [2] were selected to assess a statistical estimation of IBIF parameters based on 1) the *Maximum Likelihood Method (ML-model)*, and *Bootstrap (BT-Model)* [3] re-sampling technique. For ML-model we fit a Gamma distribution, and for BT-model a trimmed mean at 20%. From the derived statistical model, Monte Carlo random simulations were performed to estimate the Maximum Flow Declination Rate (MFDR) with their uncertainties.

Results

- In Figure 1 and 2, probability density distributions are presented for each Q parameter. Table 1 are shown the estimated Q parameters for both subjects and models as well, including an estimation using sustained vowel /a/.

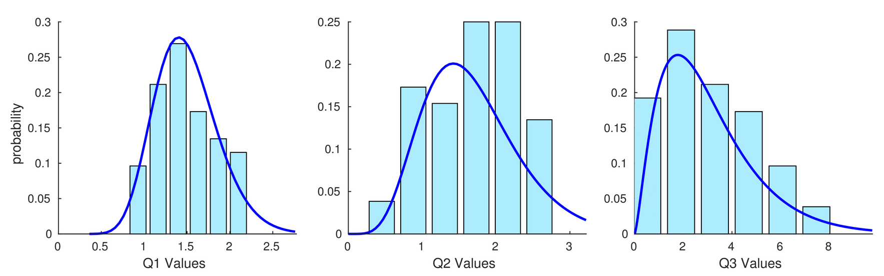

Figure 1: Q parameters distribution for subject NF026. Bar cyan: Histograms. Solid blue: Gamma distribution fit.

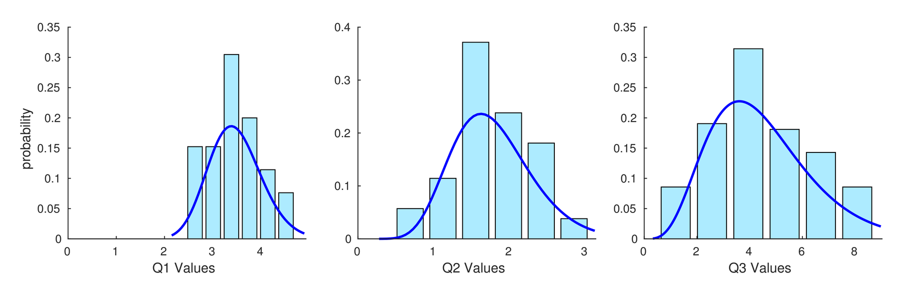

Figure 2: Same description as figure 1 for a subject PF026.

Table 1: Estimated Q parameters from both models. $\sqrt{\text{Variance}}$ in parentheses

	Subject	NF026	Subject	PF026
IBIF Parameters	ML model	BT model	ML model	BT model
Q1	1.4 (0.4)	1.5 (0.1)	3.4 (0.55)	3.5 (0.12)
Q2	1.4 (0.7)	1.7 (0.2)	1.6 (0.53)	1.8 (0.1)
Q3	1.8 (2.0)	2.9 (0.6)	3.6 (2)	4.4 (0.38)

Results

- In Figures 3 and 4 Glottal airflow waveforms (solid blue) from **ACC filtered signal closely follows the estimated glottal airflow** from OVV signal (solid black). **The uncertainties in glottal waveforms appears to be time-dependent** (dashed red line in figure 3 and 4). Over the closing phase is greater than the opening phase of the glottal pulse.

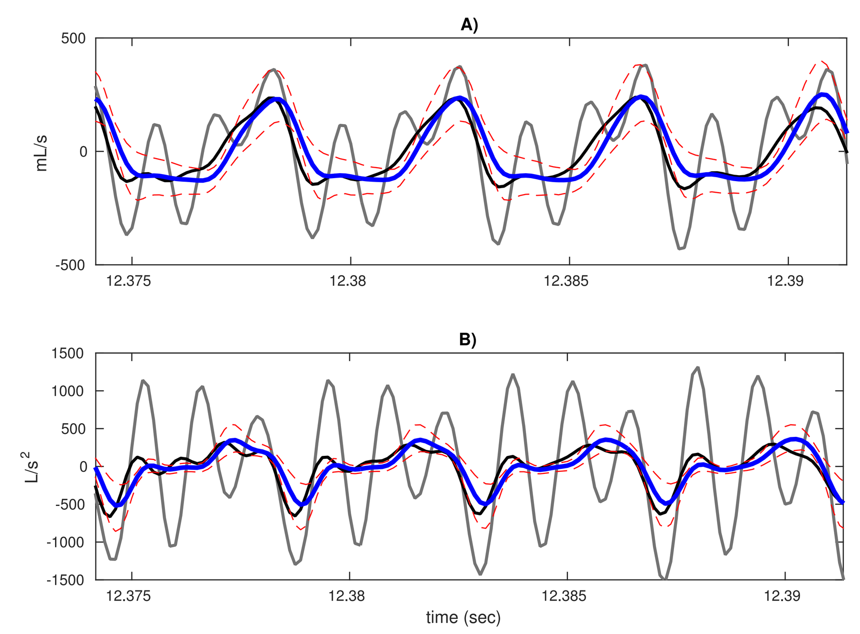

Figure 3: **Results of ML-model for subject NF026.** A) Glottal airflow estimation (GVV) from ACC signal. B) Time-derivative of GVV.

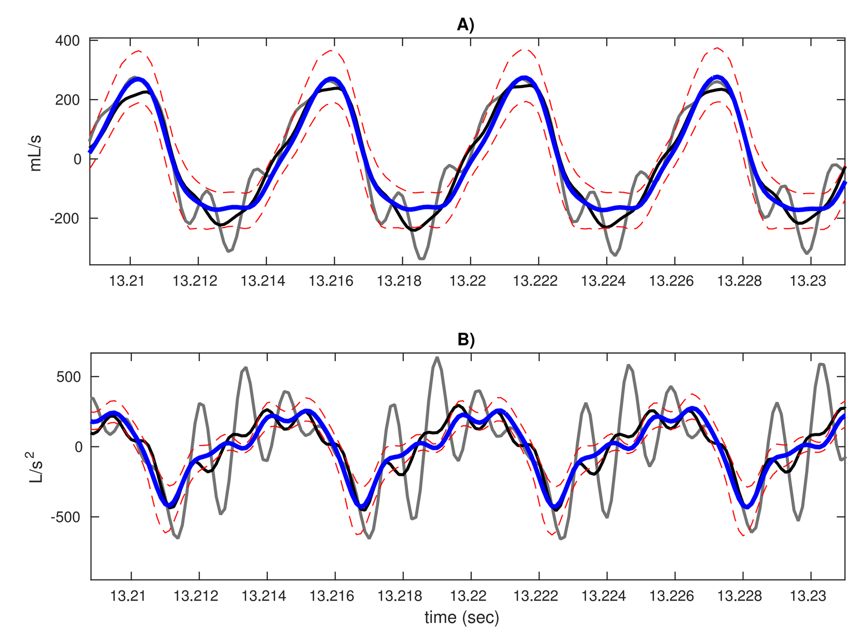

Figure 4: **Results of BT-model for subject PF026.** A) Glottal airflow estimation (GVV) from ACC signal. B) Time-derivative of GVV.

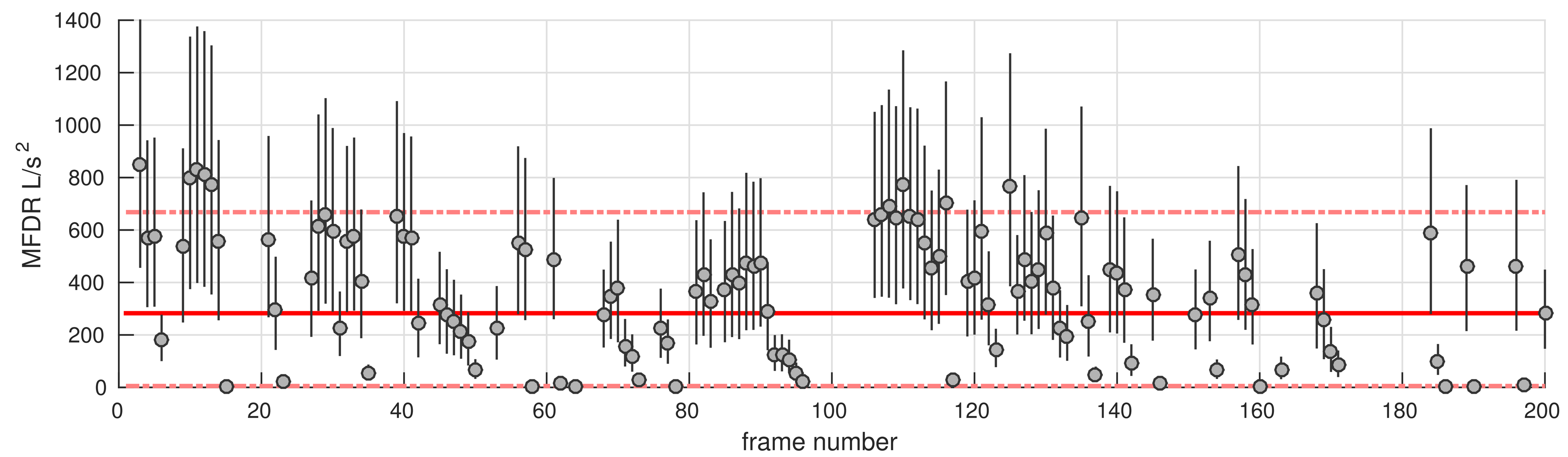

Figure 5: **MFDR mean values (bold circle) and percentile 5% and 95% (vertical black line behind bold circle) using the ML-model for subject NF026.** Red-solid: median of MFDR values (bold circle). Dashed: percentile 5% and 95% of MFDR values.

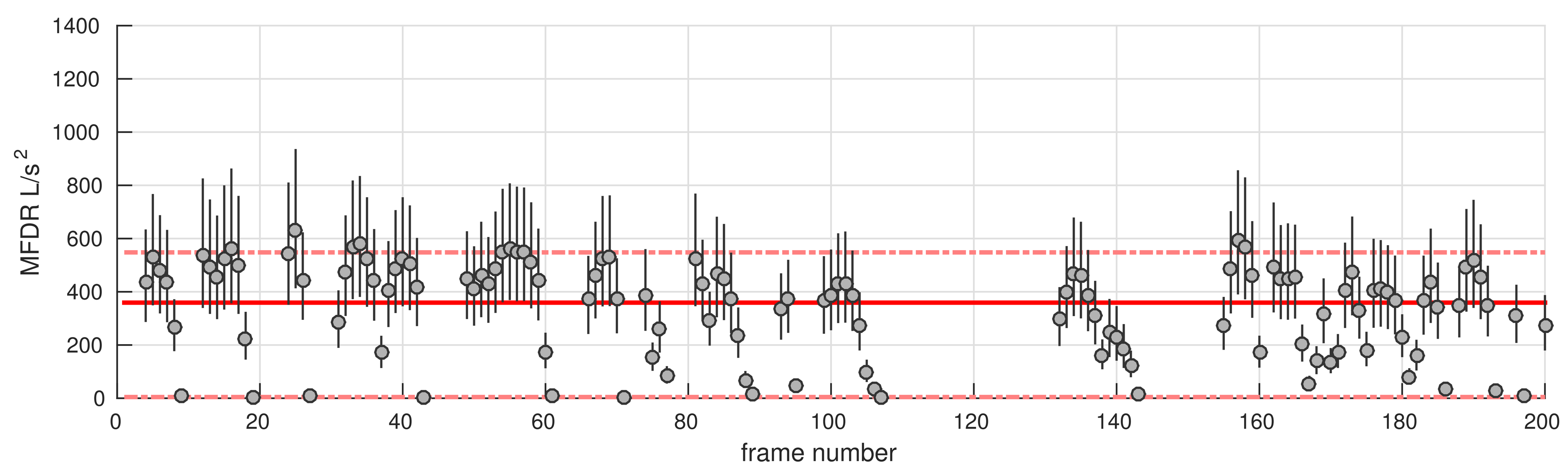

Figure 6: **MFDR mean values (bold circle) and percentile 5% and 95% (vertical black line behind bold circle) using the ML-model for subject PF026.** Red-solid: median of MFDR values (bold circle). Dashed: percentile 5% and 95% of MFDR values.

Results

- Figures 5 and 6 shows MFDR values for both subjects. The resulted **uncertainties of MFDR seems to be amplitude dependent**, as is expected for a Gamma distribution. Under previous observation, a relative error (standard error divided by mean value) is better suited to characterize the uncertainties for MFDR, which are presented in Table 2 for each subject and model as well. Both models perform consistently.

Table 2: Average uncertainty for MFDR measures

	Subject 01	- NF026	Subject 02	- PF026
	ML model	BT model	ML model	BT model
Relative error (%)	35.9	4.9	22.4	2.2

Conclusion

- The frame-based **Q parameters closely follow the proposed Gamma distribution**, showing a well defined central tendency and dispersion. This behavior suggests that the ensemble of **several frame-based Q parameters in continuous speech provide a wider overview of the underlying inverse filtering** mechanism compared to a single sustained vowel.
- The uncertainty of MFDR was described by a simple relative error which indicate the precision of the estimate.
- **The glottal airflow derived from our method, perform reasonably well** for continuous speech in a healthy and pathological subject as well.

References

- J. S. Perkell, E. B. Holmberg, and R. E. Hillman. *A system for signal processing and data extraction from aerodynamic, acoustic, and electroglottographic signals in the study of voice production.* The Journal of the Acoustical Society of America, 89(4):17771781, 1991.
- M. Zaartu, J. C. Ho, D. D. Mehta, R. E. Hillman, and G. R. Wodicka. *Subglottal impedance-based inverse filtering of voiced sounds using neck surface acceleration.* Audio, Speech, and Language Processing, IEEE Transactions on, 21(9):19291939, Sept 2013.
- L. Wassermann. *All of Statistics: A concise Course in Statistical Inference.* Springer, 2010.

Acknowledgments

- This research was supported in part by grants from the Voice Health Institute, the National Institutes of Health, National Institute on Deafness and Other Communication Disorders (Grants R33 DC011588 and F31 DC014412), CONICYT (Grants FONDECYT 1151077), and UTFSM.

vespinoza@uchile.cl