### A COMPARISON OF ESTIMATION METHODS FOR THE VHF VOICE RADIO CHANNEL

Mario GRUBER <sup>1,2</sup>, Konrad HOFBAUER <sup>1,3</sup>

<sup>1</sup> Innovative Studies (INS), Eurocontrol Experimental Centre, 91222 Brétigny sur Orge, France
 <sup>2</sup> Dept. of Aviation, University of Applied Sciences, 8020 Graz, Austria
 <sup>3</sup> Signal Processing and Speech Communication Laboratory, Graz University of Technology, 8010 Graz, Austria

mario.gruber.lav03@fh-joanneum.at, konrad.hofbauer@tugraz.at

Abstract. The aim of ongoing studies is to build a simulation model of the aeronautical VHF voice radio channel based on an analysis of previous measurements. A linear filter model is used to characterize the relation between measured input and output signals of a voice radio channel. Six well-established methods to estimate the channel impulse response are compared in terms of their accuracy and noise robustness. Deconvolution, discrete Fourier transform (DFT), least squares estimation and cross-correlation for different kinds of input signals are used for the estimations. Least squares estimation shows advantages over the other methods and is the only method feasible for all regarded scenarios. Particularly deconvolution fails to derive reliable estimates in the presence of noise. Using longer sequences shows superior performance for estimating noise, but the length is always limited by the assumed time-invariance within a segment. Estimating more filter parameters is more accurate only for a noiseless channel and always a tradeoff with the desired model.

### **Keywords**

Filter, identification, impulse response, estimation, voice radio channel.

# 1. INTRODUCTION

It is expected that analogue amplitude modulated radio will be applied in air traffic control (ATC) well beyond 2020. Current efforts in this field aim to embed inaudible digital data into the classical voice transmission, in order to enhance efficiency and flight safety [1]. In such communication systems the quality of the transmission channel has a strong influence.

This paper is related to an audio channel model using linear filters as system description. Different methods to estimate unknown impulse responses (IR), only knowing the sent and received audio signals, are compared. The objective is to prepare procedures to be applied in the analysis of the TUG-EEC-Channels Database [2], which provides recordings of audio signals transmitted with an aeronautical voice radio during measurement flights. The channel is considered to be time-invariant for short periods of time, where the conditions are regarded as constant. The simplified linear time-invariant (LTI) filter model (Figure 1) is characterized by the filter's impulse response h(n) or frequency response  $H(e^{j\omega})$  respectively, the input signal x(n), the output signal y(n) and additive white Gaussian noise (AWGN)  $s_n(n)$ .



FIG. 1: Filter model.

Section 2 provides a brief review of six different methods to estimate impulse responses. In section 3 their accuracy is compared.

# 2. ESTIMATION METHODS

The mathematical representation of linear digital filtering is convolution (1). The aim of this work is to identify appropriate filter parameters h(n), given x(n) and (1)

$$y(n) = s_n(n) + x(n) + h(n) = s_n(n) + \sum_{m=-\infty}^{\infty} x(m)h(n-m).$$

# 2.1 Deconvolution

Deconvolution is the inversion of convolution or filtering respectively. The algorithm used in the tests is based on filtering an impulse function  $\delta(n)$  by a filter which has y(n) as numerator coefficient vector and x(n) as denominator coefficient vector using a direct form II transposed implementation of the standard difference equation [3]. We assume that the support of h(n) is limited to p = 0...P - 1. To estimate the filter's impulse response the procedure is equivalent to

(2) 
$$\hat{h}(p) = \frac{1}{x(0)} \left( y(p) - \sum_{m=1}^{p} x(m) \hat{h}(p-m) \right).$$

Thus, the first element of the input vector x must be non-zero

for this method.

### 2.2 Spectral Division

The impulse response h(n) of a linear filter can be estimated by an inverse discrete Fourier transformation (DFT) of the frequency response H(k) [3], which results from a division of the output spectrum by the input spectrum.

(3) Time domain Frequency domain  
$$y(n) = x(n) * h(n) \longrightarrow Y(k) = X(k)H(k)$$

Using this theorem, a cyclic convolution<sup>1</sup> of x(n) and h(n) is inherently assumed.

### 2.3 Power Spectral Densities

Another way to use the DFT is to calculate the power spectral densities (PSD) based on correlation functions. A division in frequency domain again reveals the frequency response [4]:

(4) Time domain  $R_{xy}(n) = R_{xx}(n) * h(n)$   $\longrightarrow$  Frequency domain  $G_{xy}(k) = G_{xx}(k)H(k),$ 

whereas  $R_{xx}$  is the auto-correlation of the input signal and  $R_{xy}$  the cross-correlation between input and output. The corresponding PSD is  $G_{xx}$  and the cross-PSD is  $G_{xy}$  [3, 5].

### 2.4 Least Squares

The method of least squares selects estimates which minimize the sum of squared deviations between points generated by a function and corresponding points in the data [6, 7]. A matrix description is used to show the principle. Certain assumptions for  $\hat{h}(p)$ , x(n), y(n) and the Toeplitz structured convolution matrix **X** apply [8]. The estimate is given by the solution of the normal equation

(5) 
$$\hat{h} = (\mathbf{X}^{\mathbf{T}}\mathbf{X})^{-1}\mathbf{X}^{\mathbf{T}}y.$$

The principle is implemented by two different windowing methods [9]. The covariance method requires (6)

$$\mathbf{X_{cv}} = \begin{bmatrix} x(P-1) & \cdots & x(1) & x(0) \\ x(P) & \cdots & x(2) & x(1) \\ \vdots & \ddots & \vdots & \vdots \\ x(N-1) & \cdots & x(N-P+1) & x(N-P) \end{bmatrix}$$

and

(7) 
$$y_{cv} = \begin{bmatrix} y(\frac{P-1}{2}) \\ y(1 + \frac{P-1}{2}) \\ \vdots \\ y(N - 1 - \frac{P-1}{2}) \end{bmatrix}$$

Thus,  $N \ge 2P - 1$  samples are required to provide a meaningful estimate for P filter parameters, whereas P is assumed to be an odd number. For the auto-correlation method the entire convolution matrix

(8)  

$$\mathbf{X_{ac}} = \begin{bmatrix} x(0) & 0 & \cdots & 0 \\ x(1) & x(0) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ x(N-1) & x(N-2) & \cdots & x(N-P) \\ 0 & x(N-1) & \cdots & x(N-P+1) \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & x(N-1) \end{bmatrix}$$

is used. This implies that all values outside the observation window (n < 0 and n > N - 1) are considered to be zero. The same applies for the output

(9) 
$$y_{ac} = \begin{bmatrix} \vdots \\ 0 \\ y(0) \\ y(1) \\ \vdots \\ y(N-1) \\ 0 \\ \vdots \end{bmatrix}$$

symmetrically zero-padded to a vector of length N + P - 1.

### 2.5 Maximum Length Sequences

Maximum length sequences (MLS) are widely used as input signals to obtain the impulse response of linear systems [10, 11, 12]. One of the fundamental properties of MLS is that their auto-correlation is essentially an impulse  $\delta(n)$ . Hence, applied to a linear system, the result of the cross-correlation of input and output according to (4) is basically the impulse response.

(10) 
$$\hat{h}(p) = \delta(p) * h(p) \approx R_{xx}(p) * h(p) = R_{xy}(p)$$

The approximation is only valid for long sequences. To avoid the error induced with the assumption mentioned above, the following correction applies, assuming less filter parameters P than MLS-samples L [11]:

11) 
$$\hat{h}(p) = \frac{1}{L+1} \left( R_{xy}(p) + \sum_{m=0}^{P-1} \hat{h}(m) \right).$$

# 3. COMPARISON OF THE METHODS

The accuracy of the above methods to estimate impulse responses is compared. Either  $N=2^{14}$  samples of an arbitrarily selected audio signal<sup>2</sup> or of a repeated maximum length

(

<sup>&</sup>lt;sup>1</sup>Also circular or periodic convolution.

<sup>&</sup>lt;sup>2</sup>The audio file used in the examples is a recorded sound of a bee (16 Bit, PCM, Mono, 22.05 kHz). Thus it is not a periodic signal but still shows some repetitive structure.

sequence<sup>3</sup> are used as input signals for a predesigned testfilter (lowpass, finite impulse response (FIR), order=60, cutoff frequency  $f_c$ =4.41kHz, sample frequency  $f_s$ =22.05kHz). Estimation results are observed with and without letting the output y be affected by the noise  $s_n$  (SNR=10dB), whereas

(12) 
$$SNR = 20 \log_{10} \left( \frac{(x*h)_{\text{RMS}}}{s_{n, \text{RMS}}} \right) \text{ dB.}$$

The index RMS indicates the root mean square value of the corresponding signal.

Table 1 shows the notation for the estimation methods, which are in the following compared.

Description	Label	Reference	
Actual impulse response	h	predefined	
		test-filter	
Deconvolution	$\hat{h}_{Deconv}$	2.1 (2)	
Spectral division	$\hat{h}_{DFT}$	2.2 (3)	
Power spectral densities	$\hat{h}_{PSD}$	2.3 (4)	
Least squares,	$\hat{h}_{LS}$	2.4 (5)	
covariance method			
MLS	$\hat{h}_{Rxy}$	2.5 (10)	
MLS with correction	$\hat{h}_{Rxy\ corr.}$	2.5 (11)	

TAB. 1: Estimation methods.

### 3.1 Audio Signal without Noise

Applying the audio signal without noise corruption of the output, all methods provide almost identical results, as shown in Figure 2 (a). Figure 2 (b) shows the absolute error between each estimate  $\hat{h}(p)$  and the reference h(p). Only an error for the deconvolution method is visible, even though within a range of  $3 \cdot 10^{-9}$  and therefore negligible. All other methods give perfect estimation.

### 3.2 Audio Signal and Noise

In the presence of noise the deconvolution method fails to estimate the filter's impulse response. It is therefore omitted in the related plots. The least squares method provides the most accurate result (Figure 3).

# 3.3 MLS without Noise

For the MLS without noise in the output, Figure 4 (a) again shows the original impulse response and the obtained estimates. The uncorrected cross-correlation method reveals a significant deviation from the other methods. The occurring error approximately looks like the impulse response itself (Figure 4 (b)), which can be explained by (11). All other methods provide almost identical estimates.

### 3.4 MLS and Noise

The deconvolution method, the DFT method and the PSD

<sup>3</sup>Length L=63 samples.



FIG. 2: IR estimation without noise, audio input.



FIG. 3: IR estimation with noise, audio input.



FIG. 4: IR estimation without noise, MLS input.



FIG. 5: IR estimation with noise, MLS input.

method do not provide reasonable estimates and are therefore omitted in the related Figure 5. The plots reveal that in the presence of noise the least squares method performs even better than using the MLS specific properties for the estimation. The correction for the cross-correlation method shows no significant improvement but in fact partially increases the estimation error.

# 3.5 Comparison in Terms of Estimation Error

As before, the six different methods are compared for the two different input signals. Estimations with the noise corrupted output and those without noise influence are contrasted in Table 2. The error measure used is defined by

(13) 
$$e_y(n) = \bar{y}(n) - \hat{y}(n) = x(n) * h(n) - x(n) * \hat{h}(n)$$

and the ratio

(

(14) 
$$E_y = 20 \log_{10} \left( \frac{e_{y, \text{ RMS}}}{y_{\text{RMS}}} \right) \text{ dB}.$$

Method	Audio	Audio	MLS	MLS
		+		+
		Noise		Noise
Deconv	-165	198	-195	171
DFT	-307	-16	-68	231
PSD	-249	-21	-179	8
LS	-294	-34	-279	-34
R <sub>xy</sub>	n/a	n/a	-27	-9.7
R <sub>xy corr.</sub>	n/a	n/a	-307	-10

TAB. 2: Estimation error to output ratio  $E_y$  in dB.

# 4. ANALYSIS OF ESTIMATION PARAMETERS

Additional influences on the estimations are discussed. The same definitions as made in the last section regarding the filter, the audio signal and the noise apply.

### 4.1 Signals of Different Lengths

In previous observations the *complete* sequence x and the corresponding output sequence y obtained by convolution were considered to demonstrate the methods. Now the influence of performing an estimation with sections of length N cut from those signals is discussed. Figure 6 (a) shows the accuracy of the estimation in terms of the error measure as defined in (14). No noise  $s_n$  is added, only the influence of the signal length is evaluated. Estimations with the covariance method (6, 7) and the auto-correlation method (8, 9) are compared. The covariance method is plotted starting with P samples. This means that the matrix (6) still contains input values from outside of the observation window, which are assumed to be zero. The actual condition is only met for  $N \ge 2P - 1$ .



FIG. 6: Estimation error to output ratio  $E_y$  considering different signal lengths N and two windowing methods.

Figure 6 (b) shows the same error measure  $E_y$  for a noise corrupted channel (SNR=10dB).

# 4.2 Noise Estimation

Observing an interval longer than the impulse response gives an opportunity to estimate the noise  $s_n$  through averaging. The least squares estimation error is considered as noise estimate

(15) 
$$\hat{s}_n(n) = y(n) - \hat{y}(n).$$

Similar to (13, 14), the estimation error of the noise is defined by

(16) 
$$e_s(n) = s_n(n) - \hat{s}_n(n)$$

and the ratio

(17) 
$$E_s = 20 \log_{10} \left( \frac{e_{s, \text{RMS}}}{s_{n, \text{RMS}}} \right) \text{ dB.}$$

The longer the sequence length N is chosen, the more accurate the actual white noise  $s_n$  is estimated by  $\hat{s}_n$ , as illustrated in Figure 7.



FIG. 7: Noise estimation error ratio  $E_s$  for different observation lengths N.

#### 4.3 Influence of Filter Order

For a real-world-system the length of the impulse response is not known. In order to observe the impact of the assumed filter length  $\hat{P}(\hat{h}(p), p=0...\hat{P}-1)$ , more and less filter parameters than the predefined FIR filter of order P actually has, are estimated. Figure 8 (a) shows the error measure (14) for  $(\hat{P}-1) = 2...2(P-1)$ . No noise  $s_n$  is added. The whole audio sequence with 2<sup>14</sup> samples is applied to the least squares estimation, using the covariance method.

Figure 8 (b) shows the same error measure  $E_y$  for a noise corrupted channel (SNR=10dB).



FIG. 8: Estimation error to output ratio  $E_y$  depending on the assumed filter order  $(\hat{P} - 1)$ .

(a)  $E_{y}$ , noiseless channel

### 5. DISCUSSION AND CONCLUSION

Six different approaches to estimate channel impulse responses are compared. Audio signals and MLS are applied to an LTI system. Estimations with a clean and a noise corrupted output signal are performed.

The deconvolution method fails to provide reasonable estimates in the presence of noise. The implementation only uses as many input and output samples as estimated filter parameters. Thus, no noise averaging effect can be used by applying longer sequences.

The least squares estimation outperforms all other methods. It is the only approach effectively working for each test case and provides accurate results in every regarded scenario. Hence for the analysis of three further estimation parameters (signal length, noise and filter order) only the least squares method is used.

Longer sequences lead to more accurate results when estimating the AWGN that corrupts the system output. The covariance method lags a bit behind the auto-correlation method, due to limited data available as a result of the windowing.

To estimate the impulse response in the absence of noise with least squares, the covariance method should be used once the length condition allows. However, it is not beneficial to exceed this criterion by much. Furthermore, the length of the segment is limited due to the assumed timeinvariance of the channel within the segment and is therefore always a trade-off. With insufficient sample availability, the auto-correlation method provides an alternative windowing, which still allows an estimation although with less accuracy.

Estimating half of the actual filter parameters already provides an estimation error to signal ratio of less than -30dB. It is not reasonable to choose a longer impulse response once the outer parts of the response are indistinguishable from the channel noise or once the number of parameters exceeds the actual filter length. Estimating more parameters means that also more noise influence is reproduced in the estimate. Moreover a model depending on as few parameters as possible is desirable.

The linear model using FIR filters is one possible model chosen and all investigations in this work are based on this particular approach. To verify the applicability of this model the coherence function (18) can be used to investigate the linear dependency between the input and output.

### **Appendix A - Coherence**

The coherence function (18) is defined using power spectral densities (compare section 2.3).

(18) 
$$\gamma_{xy}^2 = \frac{|G_{xy}|^2}{G_{xx}G_{yy}}$$

The signals x and y have to contain significant power concerning the frequency band in order to obtain a meaningful coherence value.

# References

- K. Hofbauer and H. Hering, "Noise Robust Speech Watermarking with Bit Synchronisation for the Aeronautical Radio," in *Pre-Proceedings of the 9th Information Hiding conference*, Saint Malo, France, June 2007.
- [2] K. Hofbauer, "The TUG-EEC-Channels Database," Retrieved from http://www.spsc.TUGraz.at/TUG-EEC-Channels/, 15 Jan 2007.
- [3] A. V. Oppenheim, R. W. Schafer, and J. R. Buck, *Discrete-Time Signal Processing*, 2nd ed. Prentice Hall, 1999.
- [4] S. Maré, "Detection of nonlinear distortion in audio signals," *IEEE Transactions on Broadcasting*, vol. 48, no. 2, pp. 76–80, June 2002.
- [5] R. G. Brown and P. Y. C. Hwang, Introduction to Random Signals and Applied Kalmann Filtering - With Matlab Exercises and Solutions, 3rd ed. John Wiley & Sons, 1997.
- [6] D. D. Wackerly, W. Mendenhall III, and R. L. Scheaffer, *Mathematical Statistics with Applications*, 6th ed. Duxbury, 2002.
- [7] S. V. Vaseghi, Advanced Digital Signal Processing and Noise Reduction, 3rd ed. Wiley, 2006.
- [8] M. J. Levin, "Optimum estimation of impulse response in the presence of noise," *IRE Transactions on Circuit Theory*, pp. 50–56, March 1960.
- [9] H. Thornburg, "Autoregressive Modeling: Elementary Least-Squares Methods," Retrieved from http://ccrma.stanford.edu/ jos/lattice/, 6 Jul 2007.
- [10] C. Dunn and M. Hawksford, "Distortion immunity of MLS-derived impulse response measurements," *J. Audio Eng. Soc.*, vol. 41, no. 5, pp. 314–335, May 1993.
- [11] D. D. Rife and J. Vanderkooy, "Transfer-function measurement with maximum-length sequences," J. Audio Eng. Soc., vol. 37, no. 6, pp. 419–444, June 1989.
- [12] S. Golomb, *Shift Register Sequences*. San Francisco: Holden-Day, 1967.