Model order selection criteria: comparative study and applications

Zbigniew Leonowicz *Wroclaw University of Technology* ***PL***

Juha Karvanen *Helsinki University of Technology* ***FINLAND***

Toshihisa Tanaka *Tokyo Univ. of Agricult. & Technology* ***JAPAN***

Jacek Rezmer *Wroclaw University of Technology* ***POLAND***

leonowicz@ieee.org

Outline

- Model order selection.
- Correlation matrix.
- Akaike Information Criterion – AIC. Minimum Description Length – MDL, Minka Bayesian Information Criterion – MIBS.
- Investigations – basic properties.
- Time-frequency parametric spectrum.
- Switching of the condenser banks – parameter estimation of the current.

Model order – why?

- Find the appropriate dimensionality which fits a given set of observations
  - Examples:
    - choice of degree for polynomial regression, order of multi-step Markov chain
    - estimating the number of signals in multi-channel time series
    - AR model order selection
- estimation of the signal and subspace dimension → correct estimates of the signal parameters

Correlation matrix

- Signal samples

$$\mathbf{x} = \begin{bmatrix} x(0) \\ x(1) \\ \vdots \\ x(N-1) \end{bmatrix}$$

- Correlation matrix

$$\mathbf{R_x} = \mathcal{E}\left\{\mathbf{x} \cdot \mathbf{x}^{*\mathbf{T}}\right\} = \begin{bmatrix} R_x(0,0) & R_x(0,1) & \cdots & R_x(0,N-1) \\ R_x(1,0) & R_x(1,1) & \cdots & R_x(1,N-1) \\ \vdots & \vdots & \ddots & \vdots \\ R_x(N-1,0) & R_x(N-1,1) & \cdots & R_x(N-1,N-1) \end{bmatrix}$$

- Eigenvalues

$$\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_p$$

AIC & MDL

second term makes AIC inconsistent estimator

$$\mathrm{AIC} = -2\log f(X|\hat{\Theta}) + 2k$$

$$\mathrm{MDL} = -\log f(X|\hat{\Theta}) + \frac{1}{2}k\log N$$

Derivation of AIC & MDL

$$\mathrm{AIC}(k) = -2\log\left(\frac{\prod_{i=k+1}^{p}\lambda_i^{\frac{1}{p-k}}}{\frac{1}{p-k}\sum_{i=k+1}^{p}\lambda_i}\right)^{(p-k)N} + 2k(2p-k)$$

$$\mathrm{MDL}(k) = -\log\left(\frac{\prod_{i=k+1}^{p}\lambda_i^{\frac{1}{p-k}}}{\frac{1}{p-k}\sum_{i=k+1}^{p}\lambda_i}\right)^{(p-k)N} + \frac{1}{2}k(2p-k)\boxed{\log}N$$

Criterion of Minka (MIBS)

- Based on Probabilistic PCA Model
- Operates on eigenvalues of the data correlation matrix
- Fast & accurate
- Scores models according to probability they assign the observed data
- Similar to **Bayesian Classification**
- Preference for simpler models

Derivation of MIBS

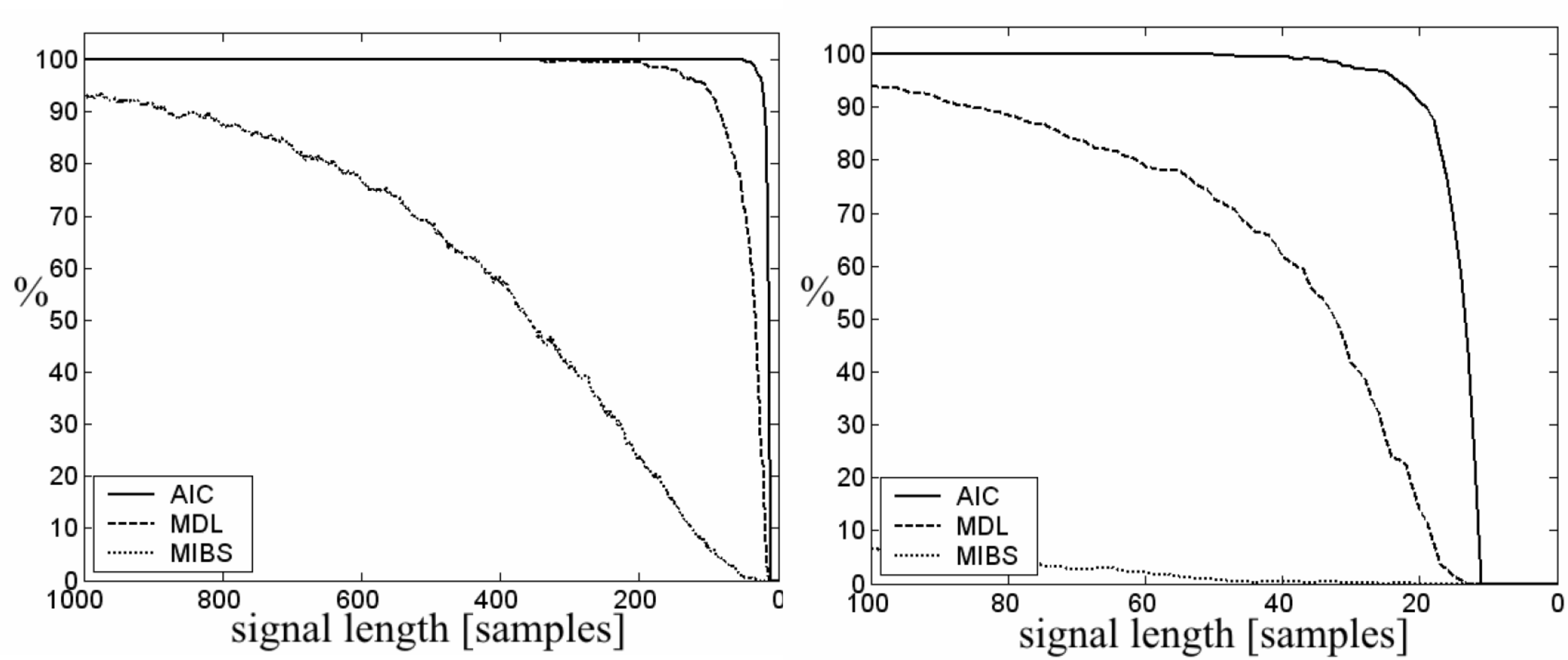
$$p(\mathbf{X}|q) = p(U)\left(\prod_{j=1}^{q}\lambda_j\right)^{-N/2} \hat{\sigma}_{\mathrm{ML}}^{-N(p-q)} \cdot (2\pi)^{(m+q)/2}|A_z|^{-1/2}N^{-q/2}$$

$$p(U) \quad = \quad 2^{-q}\prod_{j=1}^{q}\Gamma((p-j+1)/2)\pi^{\frac{-(p-i+1)}{2}}$$

$$|A_z| \quad = \quad \prod_{i=1}^{q}\prod_{j=i+1}^{p}N(\hat{\lambda}_j^{-1}-\hat{\lambda}_i^{-1})(\lambda_i-\lambda_j)$$

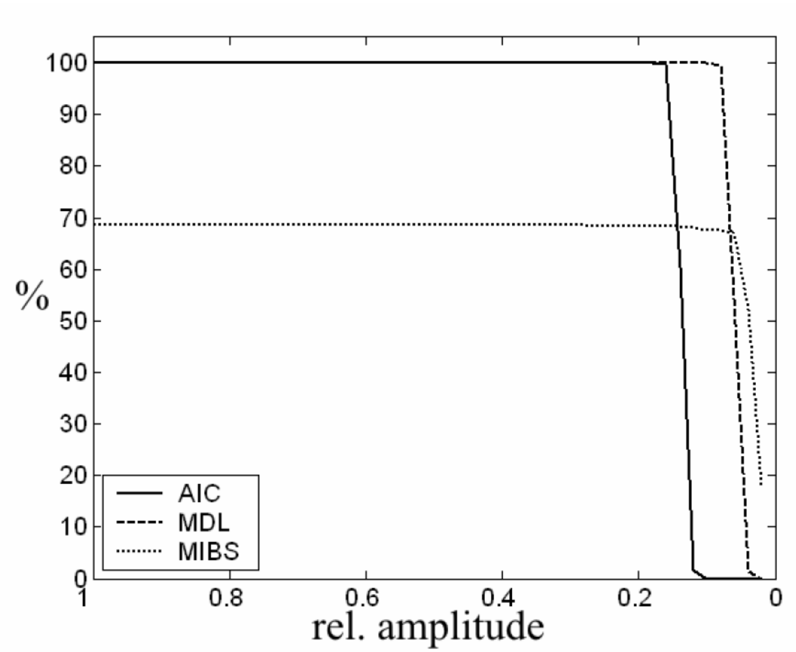
- Laplace method for approximation of integrals in Bayesian statistics

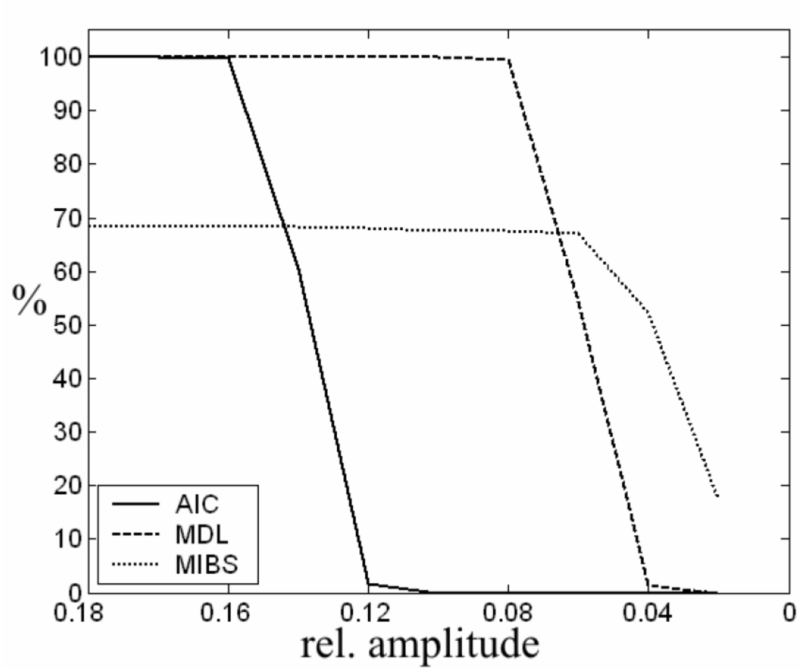
INVESTIGATIONS



- Accuracy ↔ window length
  - 2 sinusoids 50Hz & 150Hz, SNR=20dB, 1000 real.

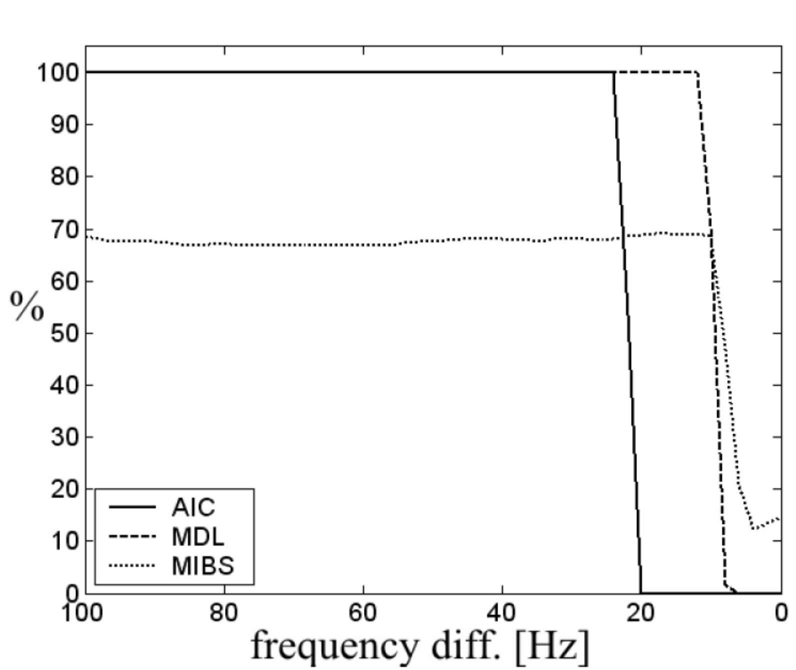
INVESTIGATIONS

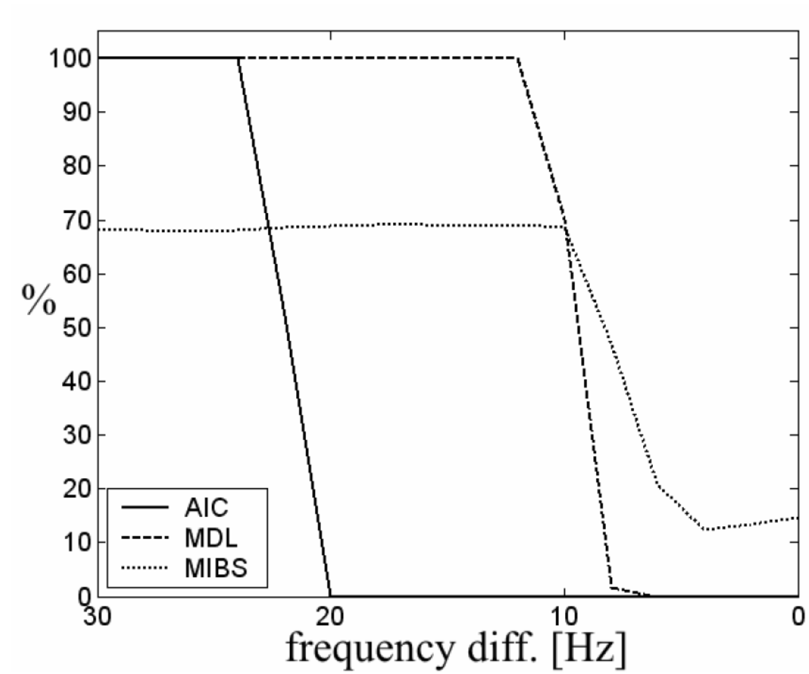




- Accuracy ↔ masking
  - Basic component 50Hz with amplitude one & 150Hz decreasing amplitude, SNR=20dB

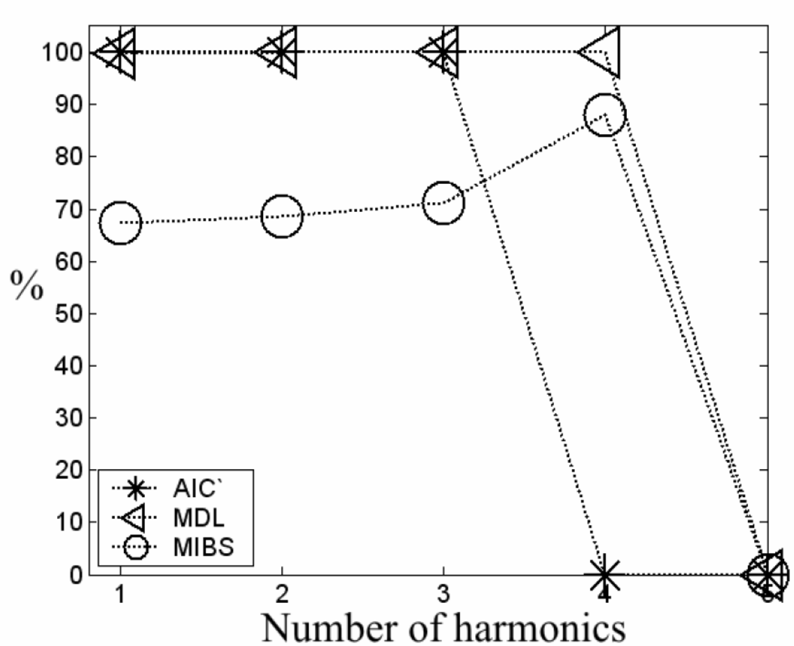
INVESTIGATIONS





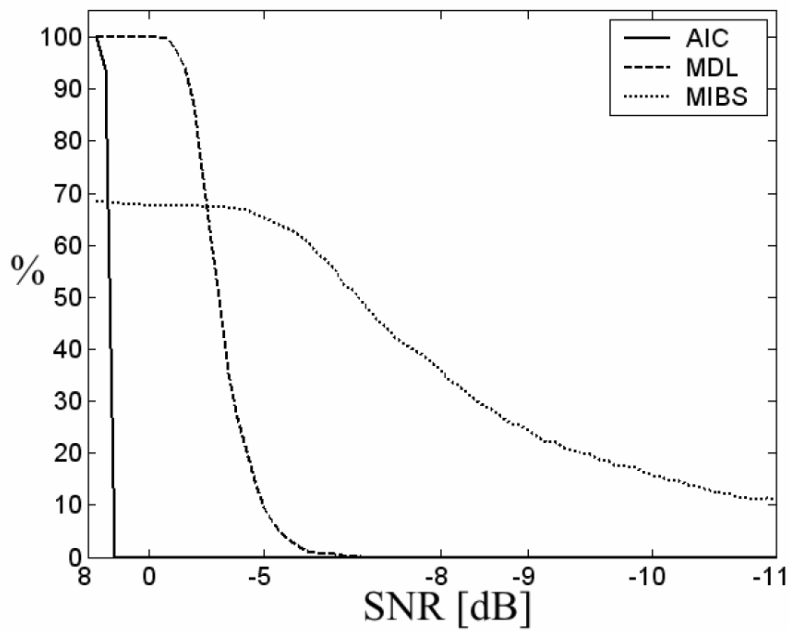
- Accuracy ↔ difference in frequency
  - Basic component 50Hz with amplitude one & second with frequency from 80 to 50 Hz, SNR=20dB

INVESTIGATIONS



- Accuracy ↔ number of components
  - 50, 100, 150, 200, 250 Hz, SNR=20dB

INVESTIGATIONS



- Accuracy ↔ Gaussian noise
  - Components 50, 150 Hz, SNR from 8 to -11 dB

Time-frequency parametric spectrum - MUSIC

- Eigenfilter $E_i(z) = e_i[0] + e_i[1]z^{-1} + \ldots + e_i[N-1]z^{-(N-1)}$

$$\mathbf{w} = \begin{bmatrix} 1 & e^{j\omega_i} & \cdots & e^{j(N-1)\omega_i} \end{bmatrix}^T$$

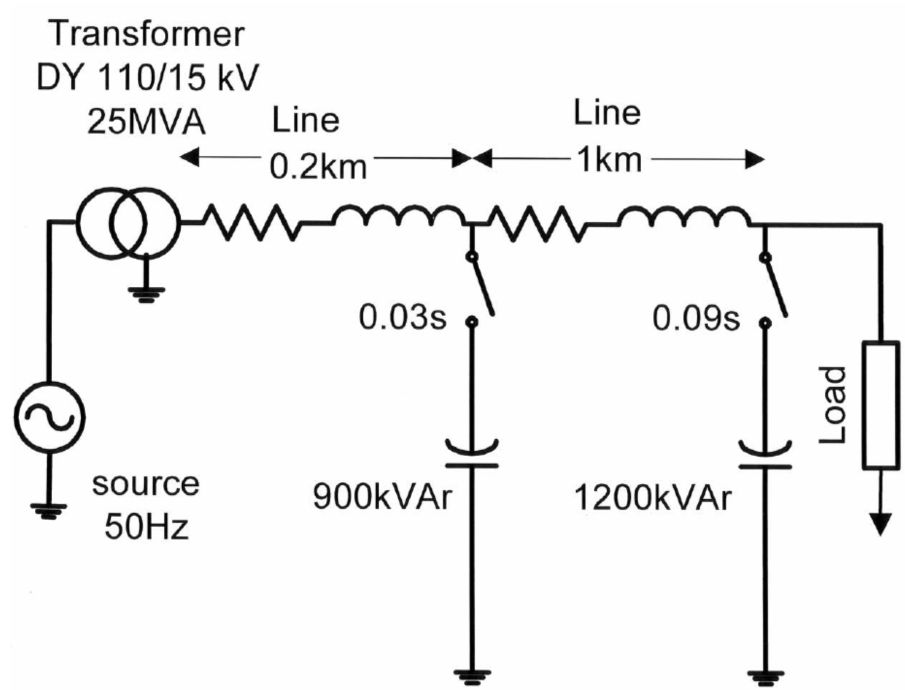
$$\mathbf{w}^{*T}\mathbf{E}_{noise}\mathbf{E}_{noise}^{*T}\mathbf{w} = \sum_{i=M+1}^{N} \mathbf{w}^{*T}\mathbf{e}_i\mathbf{e}_i^{*T}\mathbf{w}$$

- The polynomial has *M* double roots lying on the unit circle, correspond to the signal frequencies.

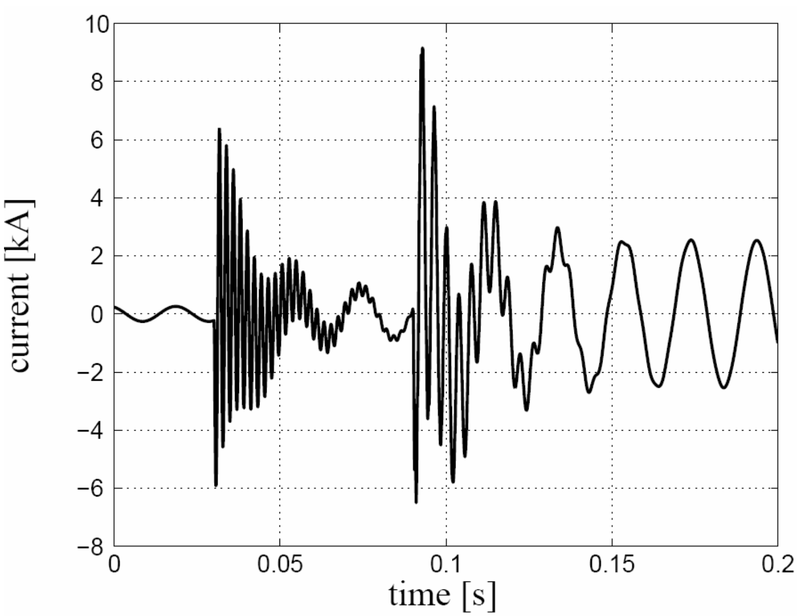
$$\widehat{P}\left(e^{j\omega}\right) = \left[\sum_{i=M+1}^{N} E_i(z)E_i^*\left(1/z^*\right)\right]^{-1}\Bigg|_{z=e^{j\omega}}$$

Switching of the condenser banks

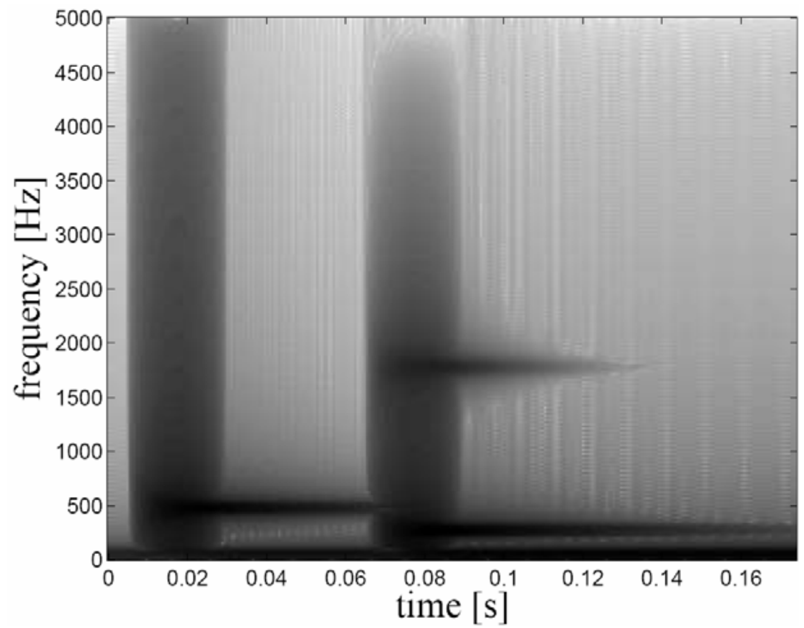
- sampling frequency 10 kHz
- analysis window 100 samples (0.01 s).
- The 1st condenser bank was switched on at t = 0.03 s and the 2nd at t=0.09 s.



STFT

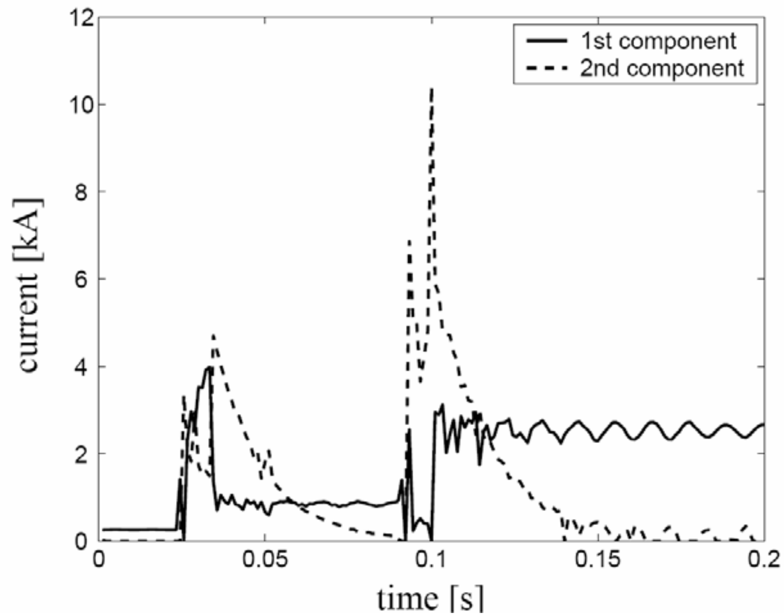


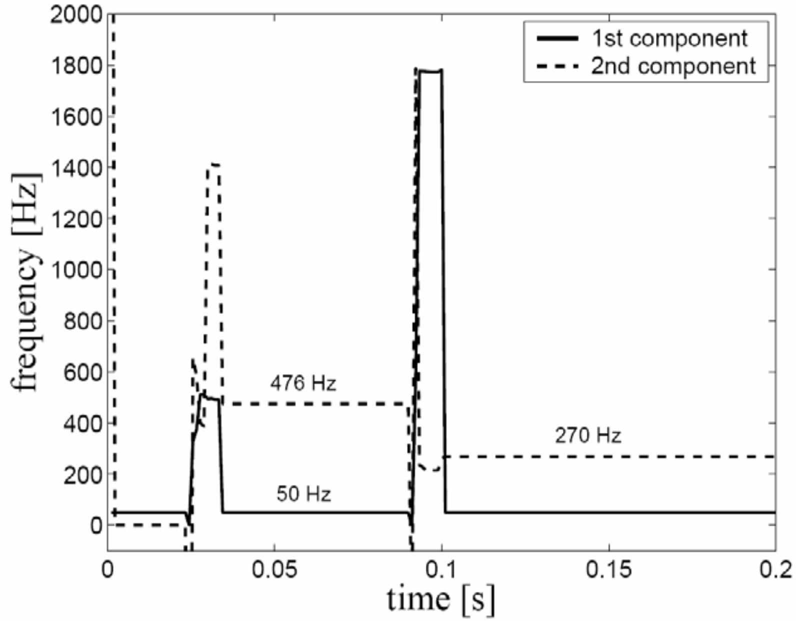
current [kA]
10
8
6
4
2
0
−2
−4
−6
−8
0
0.05
0.1
0.15
0.2
time [s]



frequency [Hz]
5000
4500
4000
3500
3000
2500
2000
1500
1000
500
0
0
0.02
0.04
0.06
0.08
0.1
0.12
0.14
0.16
time [s]

Results





- Instantaneous amplitude and frequency of two components

Conclusions

- The application of statistical model order selection (in this case - estimation of the number of sinusoidal components) allow to track on-line the parameters of the signal.
- The use of information-theoretic criterion like AIC, together with high-resolution parametric estimation method, like MUSIC, allows precise on-line estimation of the signal parameters by using the sliding window approach in the case when the parameters of the components are time-varying.