Stochastic processes evaluation with dyadic filter utilization in clinical stabilometry

Ing, Miroslav Kopecký

Vedoucí práce: Doc. Ing. Josef Kokeš, CSc.

Abstrakt

Studium lidského postoje se řadilo vždy mezi velmi aktivní oblasti boimedicínského výzkumu, při hledání porozumění a jeho spojitosti s centrálním nerovovým systémem. Samotný postoj jedince představuje multi-senzorovou regulaci s odezvou od tří základních typů čidel (chodidla, střední ucho a oční vjemy). Tyto informace mají zásadní vliv na stabilitu subjektu a v případě poškození jedno ze signálů mohou mít nepříjemné následky, kdy může dojít i k ohrožení na životě. Cílem článku je úvod do problematiky analýzy dat získaných za pomoci zařízení Synapsys Posturography System (SPS) a S.O.T testu. Soubor dat bude v článku považován za realizaci Markovského procesu (Random Walk). Dále je představen způsob modelování chování pacienta na základě znalostí rozdělení distribuční funkce. Za pomoci této informace a metody MCMC (Markov Chin Monte Carlo) je vytvořen model průběhu testu S.O.T.

Klíčová slova

Ensemble Empirical Mode Decomposition, Markov Chain Monte Carlo, Metropolis Hastings algorithm, Posturography, Stabilometry.

1. Introduction

Data analysis goes hands in hands with science since its beginning. What changes is the sophistication of the models on which the analysis is based. Stationary, linear methods fail to capture some vital mathematical and physical properties of complex dynamic data such as those from evaluation of CoP [10]. Data used in current article are taken from standard postulography checkups during movement disorders investigation on real subjects.

Other issue is also individual signal analysis where are mostly used Fourier analysis based techniques to bring out data into energy-frequency domain for better visibility of intrinsic frequencies. This approach brings to us several issues we have to deal with.

The first one to be mention is in statistical point. Any similar processes are in theirs simplicity stochastic ones not deterministic as is commonly mentioned in literature. In that meaning is reliable to use appropriate technique to get true results from them. Displayed amount of frequencies during using only Fourier based analysis could cause any misunderstanding.

Before continue in reading let’s point out what is the definition of stochastic process. Stochastic process [20] is defined in following manner, a stochastic process F is a collection,

$$F_t : t \in T, \tag{1}$$

where each F_t is an X-valued random variable and T is time. Stochastic process is more reliable to represent CoP measurement [11]. The data could be defined as Markov chain [19] which is a subset of stochastic processes [5][6]. Markov chain, named after Andrey Markov, is a discrete random process with the Markov property which has following meaning.

A stochastic process has the Markov property in the conditional probability distribution of future of the process depends only upon the present state; that is, given the present, the future does not depend on the past.
This statement corresponding with situation we have during subject checkups. Subject is asked only for quiet stand (sway) with closed or opened eyes in meaning of S.O.T. [15] Any of his CoP moment is up to his stability abilities which could be involved by any disorders.

The second issue is analytical technique. Although Fourier based techniques are widely accepted for its strong mathematical background, they have limitations in nonlinear, non-stationary and stochastic processes analysis. For example the Short-Time Fourier Transform could be taken as an instance of time-frequency analysis (STFT; see Allen and Rabiner 1977). This approach divides the time signal *f(t)* into a series of small overlapping pieces. Each piece is then windowed and individually transformed. The STFT of function *f(t)* is defined by

$$F_{ST}(t,\omega)=\frac{1}{\pi}\int_{-\infty}^{\infty} f(t)h(t-\tau)\mathrm{e}^{-i\pi\tau}d\tau \tag{2}$$

where h(τ) is the window function. Contour plots of the energy density function $|F_{ST}(t,\omega)|$ are typically presented. The approach is most useful when the physical process very close to linear, so then the superposition of the sinusoidal solution is valid and the time series is locally stationary, so that the Fourier coefficients are changing slowly. One of the drawbacks is fixed by adjusting window. The wavelet analysis seeks to address this defect by decomposing the time series into local time-dilated and time transferred wavelet components using time-frequency pieces of wavelets ψ. The wavelet transform of signal f(t) is then

$$F_{WT}(a,b)=\frac{1}{\sqrt{a}}\int_{-\infty}^{\infty} f(t)\psi\left(\frac{t-b}{a}\right)d\tau \tag{3}$$

where a is the scale and b is the time shift. The wavelet transform represents the energy in the signal of the temporal scale a at $t = b$. Wavelet transform is nice; however is maladjusted because of the same basic wavelet usage for all data. Although we received data displayed in frequency domain this fact needs not provide us correct true answer. Honestly, we have seen only frequencies which are hidden in data chain.
Situation is dramatically changed after dyadic filtering usage which brings data to reliable format before Fourier or other techniques are employed.

2. Dyadic data filtering of centre of pressure data set

2.1 Centre of pressure analysis (CoP)

Method and mathematical background of human sway (upright stand) evaluation is proposed as Centre of Pressure (CoP) [13] [14]. During any sway observation are evaluated data from X and Y directions. In clinical medical dictionary are denoted as anteroposterior and lateral directions. CoP calculation have to deal with gravity constant as $g=9.81\ m/s$

$$CoP=\frac{F}{S}=\frac{mg}{S}, \tag{4}$$

where m is subjects mass (kg). Based on CoP Stabilograms (‘kinetizieogram’) evaluation can be seen many differences in both directions (Fig 1), anteroposterior (x) and lateral directions (y). These directions are treated to confirm subjects diagnose. Disadvantage of observed values is mostly impossibility to check a specific fluctuation in time-domain plane.

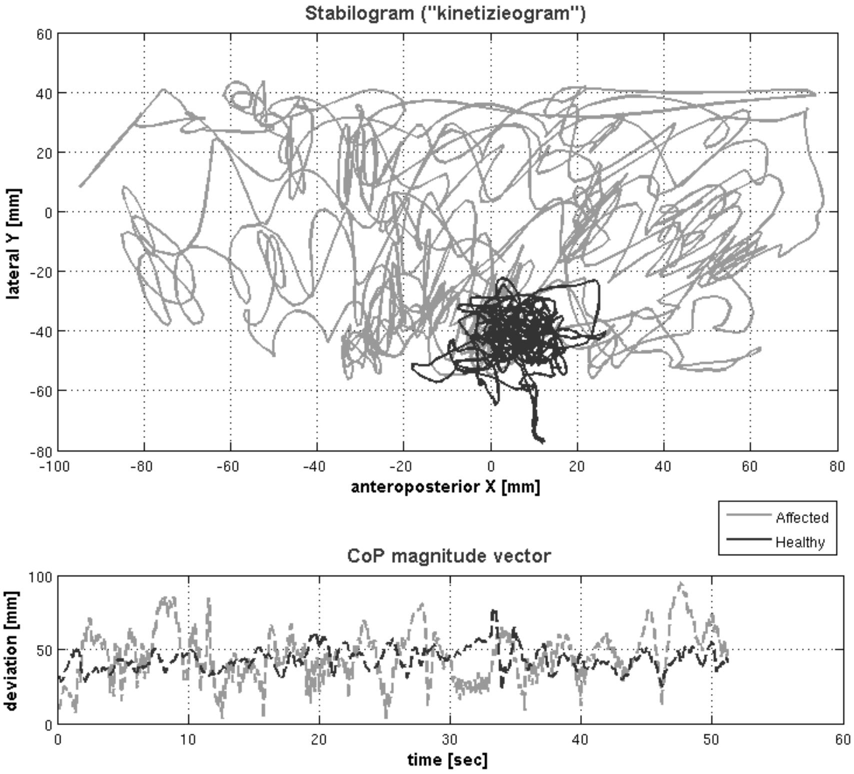

Fig. 1. CoP example with lateral direction flow of subject in the bottom

From such kinetizieogram is not possible create any detailed statement about kind of expected subject disorder. Status could be done in meaning of stability and instability which is mostly none impact to final diagnosis. As was noted before the key feature is dyadic data filtering. This could give us couple of independent intrinsic functions for next analysis.

2.2 Dyadic CoP data filtering

Ensemble empirical mode decomposition (EEMD) approach [1][2][3][4] is used for filtering purposes. The main principle of EEMD technique is to empirically decompose a signal $X(t)$ automatically into the set of the band-limited functions $IMF_{number}(t)$ called intrinsic mode functions (IMFs) with white noise addition (Fig 2). White noise plays major role (Eq.5) because of signal extraction at once and in the other hand any measurement is influenced by. In this meaning noise awareness is desired.

$$X(t) = CoP\,(t) + w_i(t) \tag{5}$$

Different white-noise realizations $w_i(t)$ are added over whole IMFs extraction procedure. Each IMF satisfies two basic conditions: (i) throughout the whole length of a single IMF, the number of extremes and the number of zero-crossings must either be equal or differ at most by one (although these numbers could differ significantly for the original data set), (ii) At any data location, the mean value of the envelope defined by the local maximum and the envelope defined by the local minimum is zero. Although the first condition is similar to the narrow-band requirement for a stationary Gaussian process, EEMD plays role of dyadic filter band.

The second one is noticed as a local requirement induced from certain exigency that the instantaneous frequency will not have redundant fluctuations. The EEMD algorithm for the signal X(t) can be summarized as follows:

1. white Gaussian noise series are added to the targeted data. Identify all the local extrema (the combination of both highs and lows) and connect all these local highs (lows) with a cubic spline or B-spline as the upper (lower) envelope;
2. decompose the data into IMFs; which means to obtain the first component $h(t)$ by taking the mean $m_n(t)$ of the upper and lower envelopes; The final $h(t)$ is designated as c_n any time;
3. repeat step 1 and 2 again based on different white noise series realization each time ;
4. do step 1 and 2 until IMFs of a data set is close to $\log_2(N)$ with N as number of total data points;
5. obtain (ensemble) the IMF means of the decompositions as the final true result

Whole process could be restored back by doing summary of individual IMF functions (Eq.6.),

$$CoP\ (t) = \sum_{j=1}^{n} c_n + r_n \tag{6}$$

where r_n is the residue of original data $X(t)$. The residue r_n could be treated as the signal trend which would be not necessary to know in next reading. After filtering procedure we gained more reliable data for next analysis as is visible form figure (Fig 2.). Main differences between healthy and affected subject are depicted here.

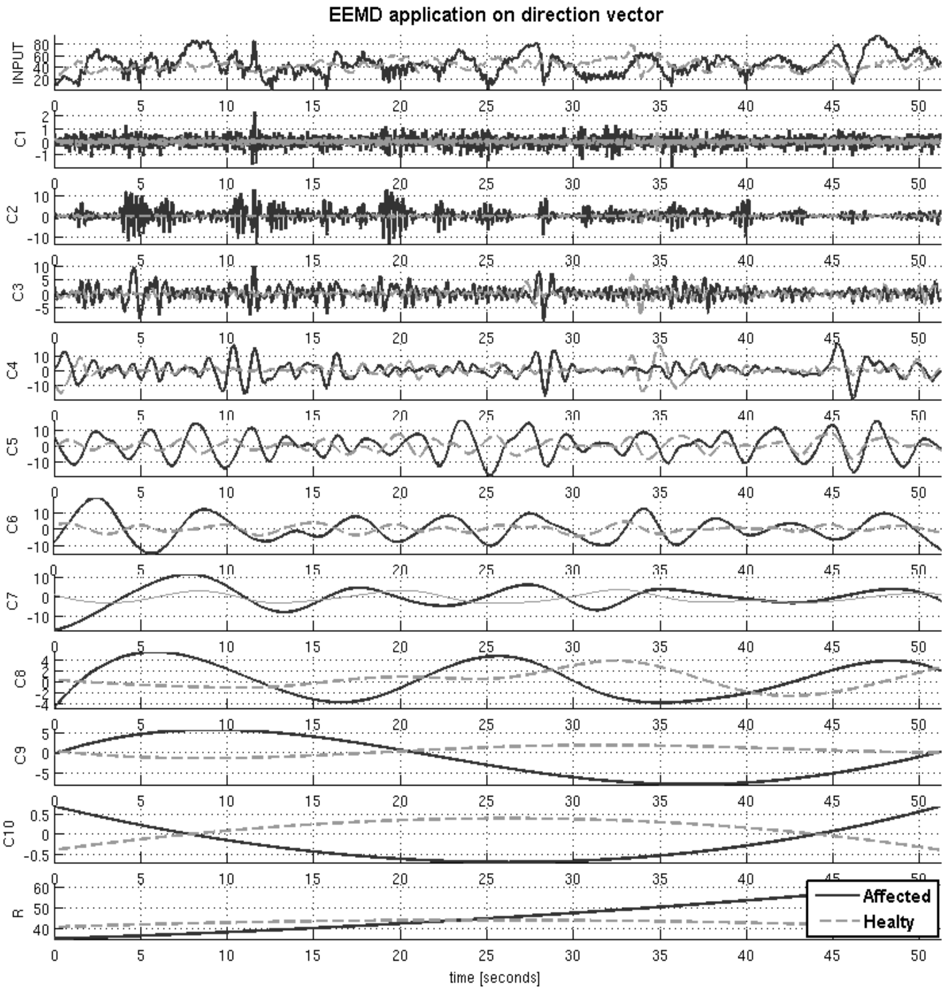

Fig. 2. extracted IMFs from the CoP lateral vector

3 CoP data normal distribution and step to human sway modeling

For majority cases and processes caused in the world is the most convenient distribution standard normal distribution. Normal distribution is sometimes called Gaussian, after Carl Friedrich Gauss. Gauss used it to analyses astronomical data sets.
The graph looks of probability density function is bell-shaped, with peak at the mean and corresponding to well known density function (Eq.7).

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{\frac{-(x-\mu)^2}{2\sigma^2}} \tag{7}$$

Where variable μ is called mean value and parameter σ^2 is the variance.
Knowledge of distribution function is helpful in finding additional information stored inside gained signal from measured CoP. Based on such knowledge are others hidden information highlighted and could be used for latter CoP modeling as will be described latter in the text.

3.1 Normal distribution of CoP data

Without dyadic data filtering has been created the set of histograms following previously noted equation (Eq.7). Even at this start moment are displayed differences between subjects. Affected subject movements are distributed in different ways over the force platform (Fig.3.).

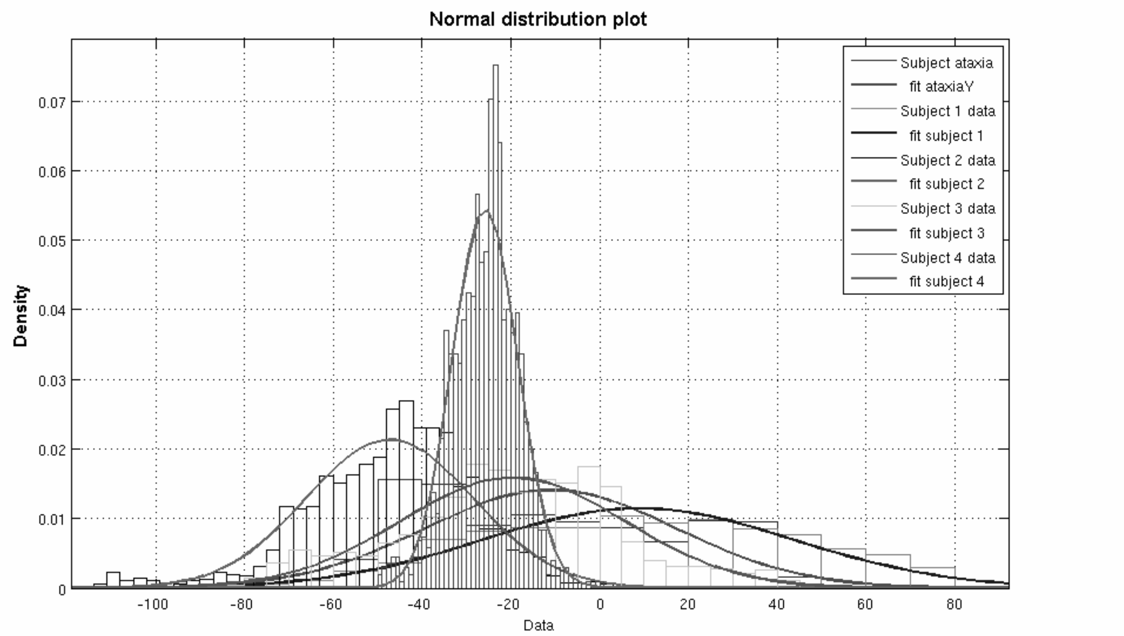

Fig. 3. *Distribution and fitting functions of one healthy subject 4 and the rest affected by any movement disorders. Subject 4 has created the highest peak around the mean value compare to affected ones*

Main point here is that healthy subject is able to control his centre of pressure around his center of gravity. This fact gives him stability and is strong signal about his health. Similar situation becomes to be visible after dyadic filtering procedure usage. Here are data of affected subjects more spread in the bottom of figure compare to healthy one (Fig.4.).

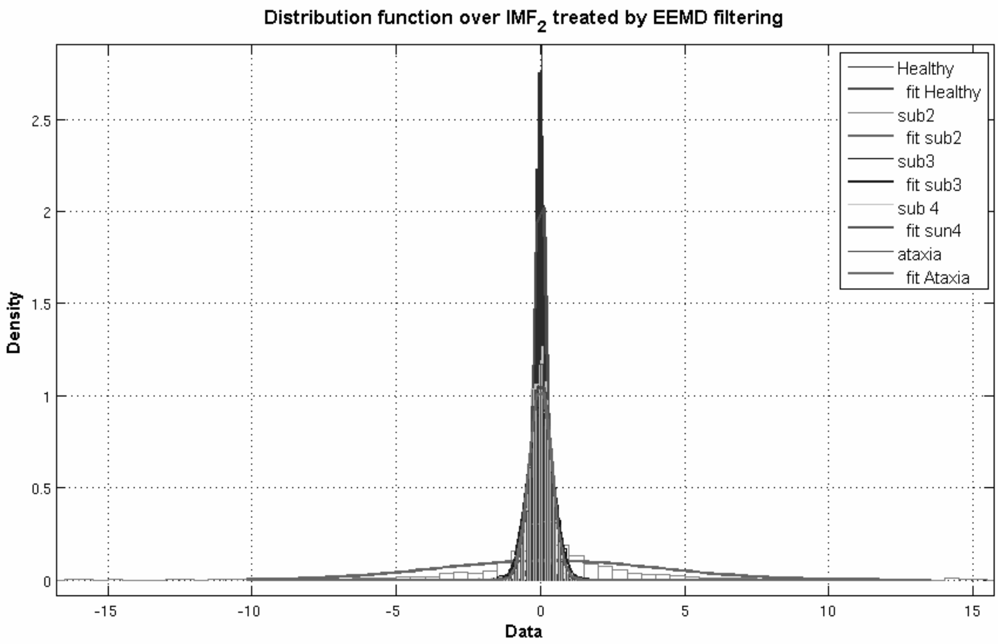

Fig.4. *Distribution and fitting functions of filtered data by EEMD method. For histograms are used only data from the second extracted IMF function. Others IMFs have similar behavior and are not necessary to display. Healthy subject has created the highest peak again. Special case is here subject affected by ataxia.*

Special case is here (Fig.4.) subject involved by proven ataxia and subject is before surgery operation. His distribution function behaves different than the rest of affected subjects but still very different to Healthy one. In non displayed IMF function histograms are distribution functions behaviors much closer to affected. Point here is that the mean value and variation are changing in dependency on kind of subject affection. This fact is still under strong research and factually brings more light to next human movement modeling.

3.2 Human movements modeling by Markov Chain Monte Carlo method utilization

In previous text has been mentioned topic about human movement modeling in CoP movement investigation. CoP is measured by force platform utilization which is implemented in Synapys Posturography System (Fig.5.). Investigated subject goes through defined Sensory Organization Test (S.O.T.) [15]. The S.O.T. objectively identifies movement disorders by accessing the patient's ability to make effective use of visual, vestibular, and proprioceptive information. From checkup are received distribution functions of any individual measurements for next analysis and investigation. Data are filtering by EEMD to increase sensitivity of specific signal deviances which could provide part of true checkup result. For modeling similar data set is then used MCMC (Markov Chain Monte Carlo) method [16]. The Metropolis-Hastings algorithm [18] as one of MCMC algorithms is going to be described in next text.

3.2.1 Markov Chain Monte Carlo methods

MCMC methods, sometimes also call as random walk Monte Carlo, are a sets of algorithms for sampling from knowledge of probability distributions based on construction of Markov chain the has the desired distribution as its equilibrium distribution. The state of the chain after a large number of steps is then used as a sample from desired distribution which is known during CoP investigation. This fact has been mentioned before.

Fig.5. *Synapsys Posturography System with implemented force platform used for S.O.T test*

3.2.1 Metropolis-Hastings algorithm for MCMC utilization

This algorithm is used to employ MCMC and was named after Nicolas Metropolis and W. Keith Hastings who extended it to more general case in 70^{th}. The Gibbs sampling algorithm [17] is a special case of the M-H. A key idea here is to use rejection mechanism, with a "local proposal": We let the newly proposed X from Markov chain ($X^{(0)}, X^{(1)}, \ldots, X^{(n)}$) depends on previous state of the $X^{(t-1)}$ which is following our general idea of human CoP modeling. The Metropolis–Hastings algorithm can draw samples from any probability distribution $P(x)$, requiring only that a function dominating the density can be calculated at x. In Bayesian application, the normalization factor is often extremely difficult to compute, so the ability to generate a sample without knowing this constant of proportionality is a major virtue of the algorithm. The algorithm generated a next Markov chain in which each state x^{t+1} depends only on the previous state x^t. The algorithm uses a proposal density $Q(x^{'}; x^t)$, which depends only on the current state x^t. This proposal is "accepted" as the next value $x^{t+1} = x^{'}$ if α is drawn from $U(0,1)$ satisfies condition

$$\alpha < \min\{\frac{P(x^{'})Q(x^t; x^{'})}{P(x^t)Q(x^{'}; x^t)}, 1\} \qquad (8)$$

If the proposal is not accepted, then current value of x is retained $x^{t+1} = x^t$.
An M-H algorithm is smartly applicable on generating samples based on priori knowledge obtained from previous CoP density function investigation to model human movement (Fig.6.).

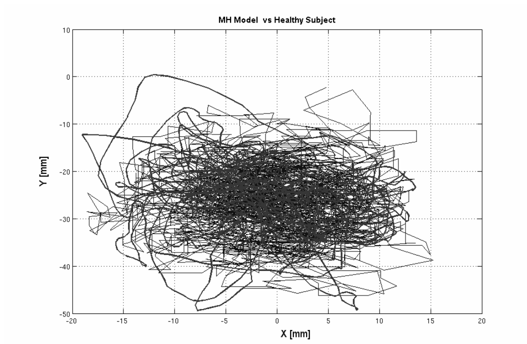

Fig. 6. *Stabilograms of Healthy subject vs. modeled ones by M-H usage. Both stabilograms are very similar and differences are not eye-catching*

4 Stochastic process analysis in frequency domain

In previous parts we have pointed out possibilities of data sorting and filtering to obtain valuable results from investigation. Next reading will touch Fourier transform techniques to check and discover intrinsic frequencies inside the signal. EEMD method is here employed as dyadic filter band which allow us to received smoothed signal. This signal includes highlighted parts which are necessary for analysis as you have seen before. Disordered frequencies are projected very sharply over individual IMF functions. Such signal could be treated then by Fourier transform. Fourier transform moves signal from time domain to frequency-energy domain.

Small subject perturbations are here projected as peaks of frequencies. It seems to be specific based on type of affection or movement disorders. EEMD plus FFT method allows us to see frequencies change over all extracted IMFs from the investigated data set (Fig.7.).

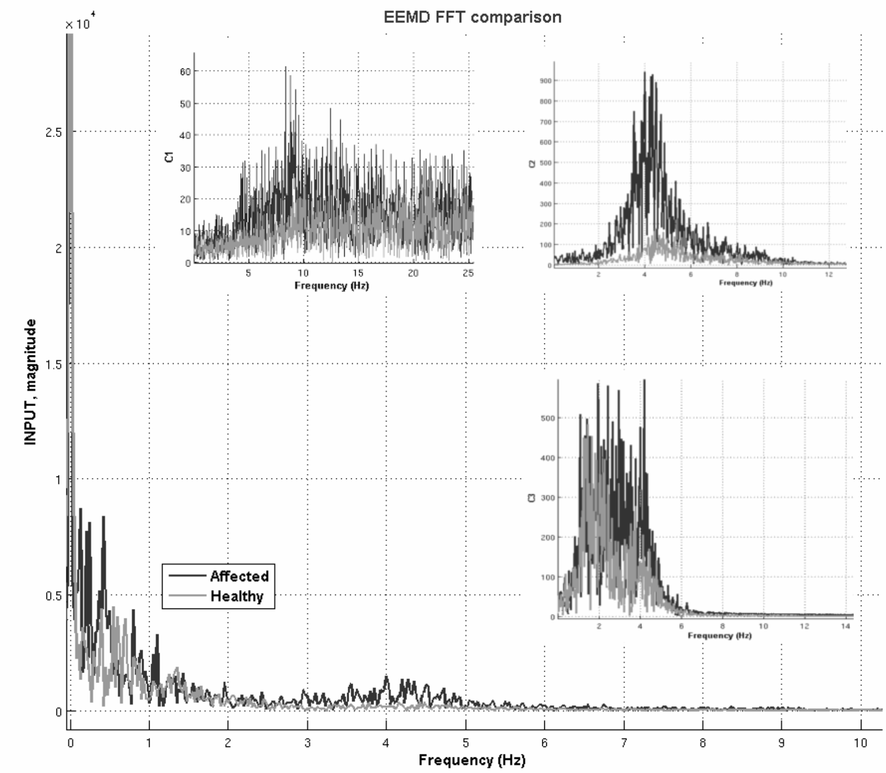

Fig. 7. *Comparison in energy-frequency domain; additional frequencies projection in signal based on type of movement disorder. C1, C2 and C3 are extracted IMFs from input signal treated by FFT*

4.1 M-H generated model comparison between healthy and affected subjects

Priori knowledge of subject distribution allows us in terms of Metropolis-Hasting generate set of expected subject's movements (Fig.8.). There seems to be no need to look through real subject. Although research in this way is not fully finished yet the received results are more then satisfactory. This could be other steps in human movement pattern research. The results are planed to be published in near future. Possibility of generating specific data set is necessary in coefficients searching. Better understanding of human movement is the one of other impact.

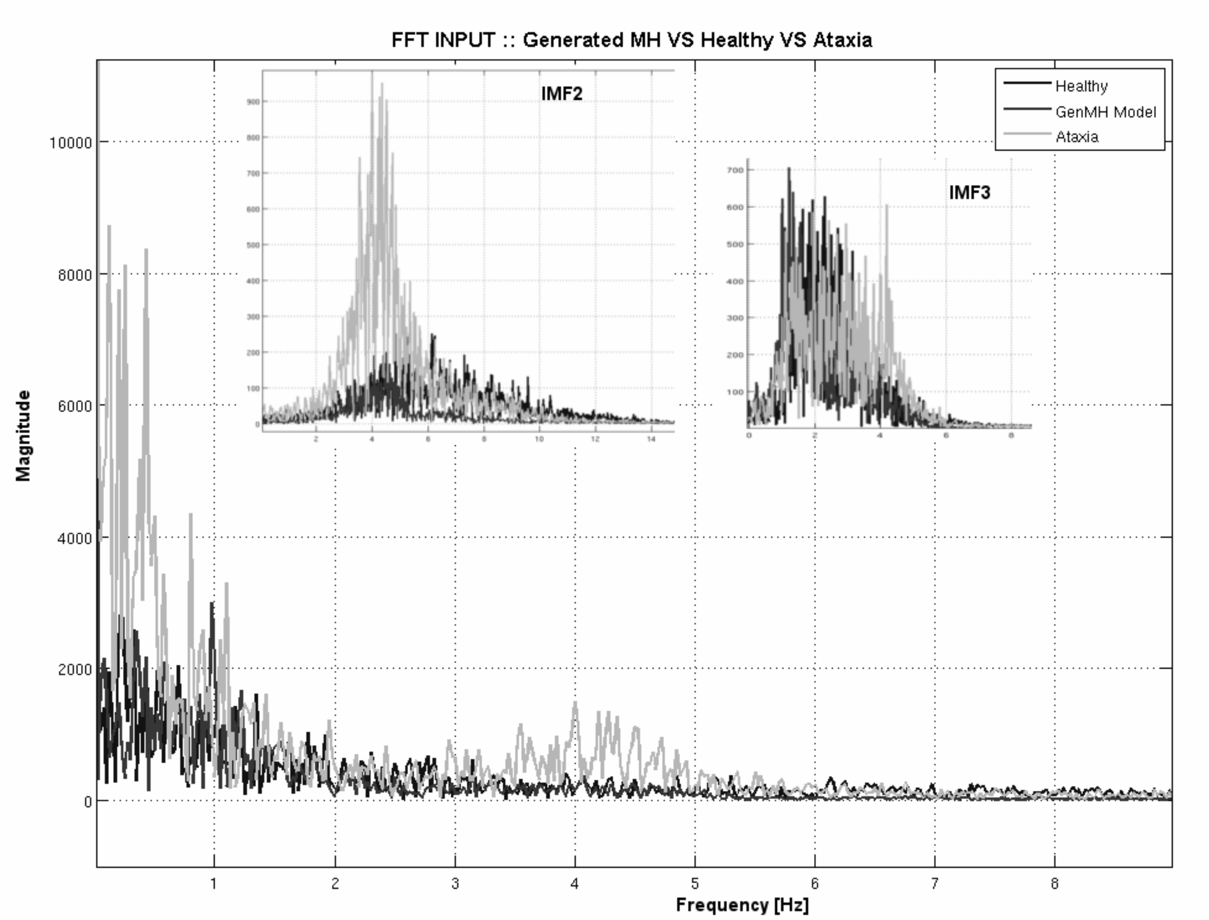

Fig. 8. *Comparison in energy-frequency domain of healthy generated, real healthy and real affected subjects; Visible is strong similarity between generated and healthy subjects compare to big signal deviances in case of affected one.*

5 Discussion and final overview

In the article has been shown possible way of stochastic data sets analyses which were received from real posturography checkup. The individual checkup is possible to consider as individual realization of Markov Chain in meaning of previously described way. Subject's CoP movement is involved only by his own abilities because of quiet sway over S.O.T test. Specific disorders or affection is then sharply projected. Also has been shown that the utilization of data filtering seems to be useful before any type of analysis employment. Affected subject in this article has strong damage of his CNS which disallows him regulate his stability. This case is specific and helps us in comparison. In the real live such signals could be passed up which is damaging next subject's live.

Dyadic filtering has been depicted as usable and able to project and discover most of signal perturbations against desired state. This could be understand as the first step in next human pattern research. Searching of patterns constants has been done by using MCMC as the one part of investigation. The Metropolis-Hastings algorithm has been introduced in the text as reliable tool.

Impact of next research is also in reducing cost of health care and doctor's time which is requested for analysis.

6 References

[1] Huang, N.E., Shen, Z., Long, S. R., Wu, M.C., Shih, H. H., Zheng, Q. Yen,N.-C., Tung, C.C., and Liu H.H.: "*the empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis*", Proc. R.Soc.London, Ser. A 454, 903-995,1998

[2] Z. Wu, N.E. Huang, "Ensemble Empirical Mode Decomposition and Noise-assisted Data Analysis Method", World Scientific Publishing Company, Advances in Adaptive Data Analysis Vol. 1, No. 1, 2009, p.1–41

[3] N. E. Huang, Z. Shen, S. R. Long, M. C. Wu, E. H. Shih, Q. Zheng, C. C. Tung and H. H. Liu, "*The empirical mode decomposition method and the Hilbert spectrum for non-stationary time series analysis*", Proc. Roy. Soc. London 454A , 1998, p. 903–995

[4] M. Kopecky: *Analýza obrazu s využitím bílého šumu pro zpracování dat, ARAP 2009, 2009, pp.94-100*

[5] Diks, C.: "*Nonlinear Time Series Analysis: Methods and Applications*", Vol. 4 in series Nonlinear Time Series and Chaos, edited by H. Tong (World Scientific, Singapore). 1999, ISBN: 9810235054

[6] Kantz, H., Schreiber. T.,: "*Nonlinear Time Series Analysis*", Cambridge University Press 1999, 304 pp

[7] Wu, Z., Huang N.E.: "*A study of the characteristics of the white noise using the empirical mode decomposition method*", Proc. R. Soc. London, Ser. A, 460, 1597-1611, 2004

[8] Flandrin, P.,G.Rilling and P. Concalces: "*Empirical mode decomposition as a filter bank*", IEEE Signal Process, 2006, 112-115

[9] Brighm, E.O. : "*The Fast Fourier Transform*", Prentice Hall, 1974

[10]T.S. Kapteyn, W. Bles, Ch. J. Njiokiktjien, L. Kodde, C.H. Massen,and J.M.F. Mol: "*Standardization in Platform Stabilometry being a Part of Postulagraphy*", SPEI Editeur Paris (1983), pp 7:321-326

[11]L. Ladislao, S. Fioretti: "*Nonlinear analysis of postulography data*", Springer 2007, 45: 679-688

[12]Patel, M., Fransson, P.A., Lush, D., Gomez, S.: "*The effect of foam surface properties on postural stability assessment while standing*", ScienceDirect 2008, pp. 649-656

[13]L. Chiari, A. Cappello, D. Lenzi, U. D. Croce*:" An improved technique for the extraction of stochastic parameters from stabilogram*", Gait Posture 12,2000,, 225-234

[14]A. Shumway-Cook, FB. Horak: "*Assessing the influence of sensory interaction of balance*", Phys Ther 66:1, 1986, Suggestion from the field, pp. 1548-1550

[15] SENSORY ORGANIZATION TEST (SOT), 2010, http://resourcesonbalance.com/neurocom/protocols/sensoryImpairment/SOT.aspx

[16] Markov Chain Monte Carlo, 2010, http://en.wikipedia.org/wiki/Markov_chain_Monte_Carlo

[17] Gibbs Sampling, 2010, http://en.wikipedia.org/wiki/Gibbs_sampling

[18] Metropolis-Hasting Algorithm, 2010, http://en.wikipedia.org/wiki/Metropolis%E2%80%93Hastings_algorithm

[19] Markov Chain, 2010, http://en.wikipedia.org/wiki/Markov_chain

[20] Stochastic Process, 2010, http://en.wikipedia.org/wiki/Stochastic_process