Reasoning from non-stationarity

Zbigniew R. Struzik

Centrum voor Wiskunde en Informatica (CWI)
Postbus 94079, NL-1090 GB, Amsterdam, The Netherlands
email: `Zbigniew.Struzik@cwi.nl`

Willem J. van Wijngaarden

Afdeling Verloskunde en Gynaecologie, H4 - 205, Academisch Medisch Centrum,
Postbus 22660, 1100 DD Amsterdam, The Netherlands
email: `W.J.vanWijngaarden@amc.uva.nl`

Robert Castelo

Institute of Information and Computing Sciences, Utrecht University,
Postbus 80.089, NL-3508 TB Utrecht, The Netherlands
email: `roberto@cs.uu.nl`

Abstract

Complex real-world (biological) systems often exhibit intrinsically non-stationary behaviour of their temporal characteristics. We discuss local measures of scaling which can capture and reveal changes in a system's behaviour. Such measures offer increased insight into a system's behaviour and are superior to global, spectral characteristics like the multifractal spectrum. They are, however, often inadequate for fully understanding and modeling the phenomenon. We illustrate an attempt to capture complex model characteristics by analysing (multiple order) correlations in a high dimensional space of parameters of the (biological) system being studied. Both temporal information, among others local scaling information, and external descriptors/parameters, possibly influencing system's state, are used to span the search space investigated for the presence of a (sub-)optimal model. As an example, we use fetal heartbeat monitored during labour.

---

1 Introduction

Universal descriptions of complex phenomena, like that using multifractal cascades, may prove to be inadequate for a complete description, even if valid on a restricted temporal range or in isolated or free-running conditions. The

inadequacy of such models is often demonstrated by non-stationarity of their characteristics. This aspect is usually neglected (through the selection of examples supporting the theory) or filtered out.

However, failure of the model to explain the phenomenon fully may well provide significant insight into the dependence of the system on the external conditions. Structures emerging from the non-stationarity of the (most) sophisticated description available very likely indicate that more dimensional embedding may be needed for a proper description of the system under study. Such embedding and the right parameter choice for the more adequate model may not be easy to identify. In science, this search trajectory for better models has been routinely carried out at the cost of repeated experiments under varied input or environment conditions (or modes of interaction with the environment) [1]. Such a search for a better model in, for example, biomedical science may be shortened now, since data for many experiments is digitally/automatically stored and available for widespread analysis.

Of course, computerised analysis of data and model fitting is not new. The computer has been used to do analysis or simulation in the natural sciences for some forty years. But until very recently, in the natural sciences the computer has only been used to provide answers to hypotheses involving relatively few degrees of freedom. In any instance of a more dimensional problem, a human expert had to investigate each and every output of the run of the program. Conclusions from the series of analyses were then drawn and theories posed, subject to testing. This process, usually tedious, is near to or just impossible when the volume of data and the dimensionality of the problem increase. It is not always possible to think of heuristics which would shorten the search path for interesting relationships or take a cross-section of the data along the most promising plane of dimensions.

However, due to the modern technology of data analysis and reasoning, such 'automated research' may be possible.[1] By incorporating the most advanced models available with explicitly identified instances where they fail to describe the model (as is manifested in non-stationarity), new theories can be generated and (dis-)proven by the system. The scientist, of course, remains in the centre of such an automated research trajectory - in fact the automated reasoning is merely meant to add to the expert's guided process, as a tool capturing, (dis-)proving and visualising the non-evident dependences.

In this paper, we first briefly present the characterisation of time series by a local version of the Hurst exponent in section 2. Next, using the example of fetal heartbeat during labour, we suggest capturing the non-stationary behaviour of the local h with a simple heuristic - linear integral (cumulative indicator).

---

[1] See [2] for introduction to Bayesian techniques.

Finally, in section 4, we suggest the recovery of multiple order dependencies between the cumulative product (integral) of the non-stationarity of local h and other characteristics of the fetal heartbeat, with external information on the fetal outcome and about the pregnancy. The paper closes with conclusions and suggestions for future developments.

2 Local Hölder exponent h makes possible the detection of non-stationarities in multifractal characteristics of complex systems

For the stationary fractional Brownian noise, we would expect that any local estimate of the Hurst exponent h is equal to the mean or global Hurst exponent H. Of course, for finite length samples and single realisations, we will have fluctuations in the local h exponent but they should prove to be marginal and diminish with increasing statistics. This will not be the case with a multifractal. The local h will show a wide range of exponents regardless of the resolution and sample size [3–5]. What we would expect to remain unchanged (or stationary) for the multifractal (cascade) is the multifractal spectrum of h, i.e. $D(h)$.

In figure 1, an example time series with the local Hurst exponent indicated in colour are shown. We have chosen the record of healthy (adult) heartbeat intervals and white noise for comparison. The background colour indicates the Hölder exponent h - the local counterpart to the Hurst exponent H. It is centred at the mean value corresponding with the Hurst exponent at green. The colour goes towards blue for higher h and towards red for lower h. In the same figure 1, we show corresponding log-histograms of the local Hölder exponent.[2] Each h measures a so-called singularity strength and thus a histogram provides a way to evaluate the 'singularity spectrum'. In other words, the local h measures local contribution to the multifractal spectra [5].

Stationarity in system/model characteristics like local h or spectral characteristics is, however, rarely encountered in real life experiments. Real life time series samples may 'behave multifractally' on a fixed time scale and a biological multifractal system may undergo dynamic changes in its characteristics affecting the spectrum of h. Such non-stationarity of a systems' universal characteristics may, in fact, prove important for the more complete description of the system. The system may behave quasi-multifractally in free running con-

_____

[2] They are made by taking the logarithm of the measure in each histogram bin. This conserves the monotonicity of the original histogram, but allows us to compare the log-histograms with the spectrum of singularities $D(h)$. By following the evolution of the log-histograms along scale, one can extract the spectrum of the singularities $D(h)$ (multifractal spectrum).
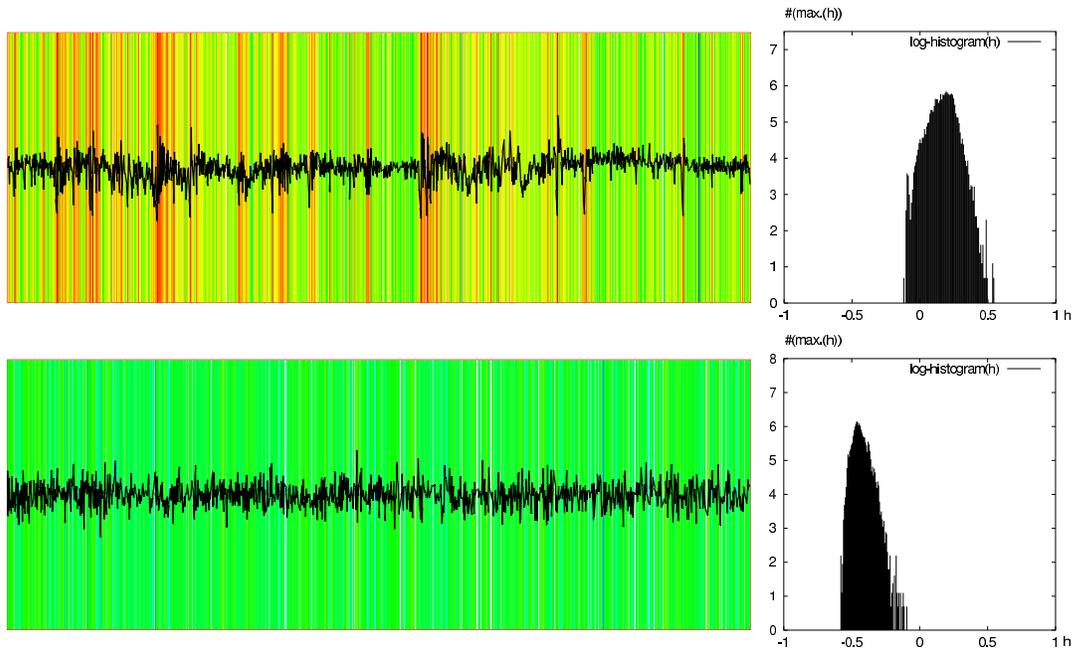
Figure 1. Left: Example time series with local Hurst exponent indicated in colour: the record of healthy heartbeat intervals and white noise. The background colour indicates the Hölder exponent locally, centred at the Hurst exponent at green; the colour goes towards blue for higher h and towards red for lower h. Right: The corresponding log-histograms of the local Hölder exponent.

ditions, but the non-linear interactions within the system may be affected by changing environmental influence. As a result the main characteristics, like mean h or spectrum width, may change with time. Often such changes may in fact show a direct correspondence with the environmental influence [6].

Of course direct influence is not the only relationship possible and in some cases a long-term influence on system's characteristics is possible. A sad example of such long term influence are the many malfunctions of heart functioning. In the following, we will address in greater detail the issue of the malfunctioning of the fetal heart during labour and relate it to the non-stationarity of, among others, the local h (roughness exponent) characteristic of the heartbeat.

Methods of wavelet transform modulus maxima (WTMM) based multifractal analysis (MF) and detrended fluctuation analysis (DFA) have been demonstrated to be suitable for capturing scaling and correlation characteristics of the fluctuations in human heartbeat intervals [3,7–9]. These characteristics, obtained under a variety of conditions, have also been shown to reflect deviations in the heartbeat due to a variety of malfunctions [3,4,10,11] and physiological behaviour [6,12,13]. Obviously, these results can be considered for clinical applications.

Unfortunately, such standard methods of statistical analysis of heartbeat signals are not directly applicable to the practical problem of evaluating (the

4

characteristics of) the fetal heartbeat in real time during labour. One obvious reason for this is that these statistical techniques use long stretches of data to provide estimates of global measures like correlation exponents or multifractal spectra. A typical measurement requires over $30,000$ samples (2^{15}) to provide reliable estimates of scaling for the extraction of exponents and reliable transformation from scaling exponents to the domain of multifractal spectra. The typical heart rate of a fetus is about 130 beats per minute. The required time stretch for acquiring a sufficiently long data set (2^{15}) is thus about 250 minutes. By this time, the baby is often already born. Decisions about an intervention (such as a Caesarean section) have to be taken on the basis of $10 - 60$ minutes' observations of the heartbeat.[3]

Another, related, reason is that fetal heartbeat during labour is a highly dynamic process which needs to be monitored in real time. It is evident that in the case of the process evolving in time, we cannot expect that a global multifractal description is going to be sufficient. Such a global multifractal description may well be valid in isolated conditions and may well characterise universality class of the phenomenon [3], but for the tasks of differentiating between cases and temporal description, more sensitive measures are needed.

Such measures may well be derived from the proven description of the system in isolated conditions. The local Hölder exponent is such a potentially useful characteristic. In an isolated stationary case, it measures the local contribution to the multifractal spectra. But in dynamic conditions when the system is subjected to changes the temporal fluctuations of this characteristic can provide invaluable insight into changing conditions of the system.

3 Exploiting non-stationarity of the local Hölder exponent h in a heuristic model

In the case of the fetal heartbeat during labour, there is no reason why the local Hölder exponent $h(V_i)$ of the variability V_i component of the fetal heartbeat intervals should be stationary. It reflects dynamic changes in the condition of the fetus and the degree of stress to which it is subjected. Despite the fact that stress has a rapid effect on the heartbeat, the effects on the state of the fetus are not always immediate. This is why short dynamic changes in heartbeat characteristics (which determine the multifractal picture) may not be relevant and not representative of the state of the fetus. Rather than expanding

---

[3] It has to be noted that fetal heartbeat is the only indicator of the state of the fetus. Therefore the characteristics determined from it are crucial in taking the decision to carry out an operative delivery or perform a fetal scalp blood sample to assess the metabolic status in case of suspected hypoxia.
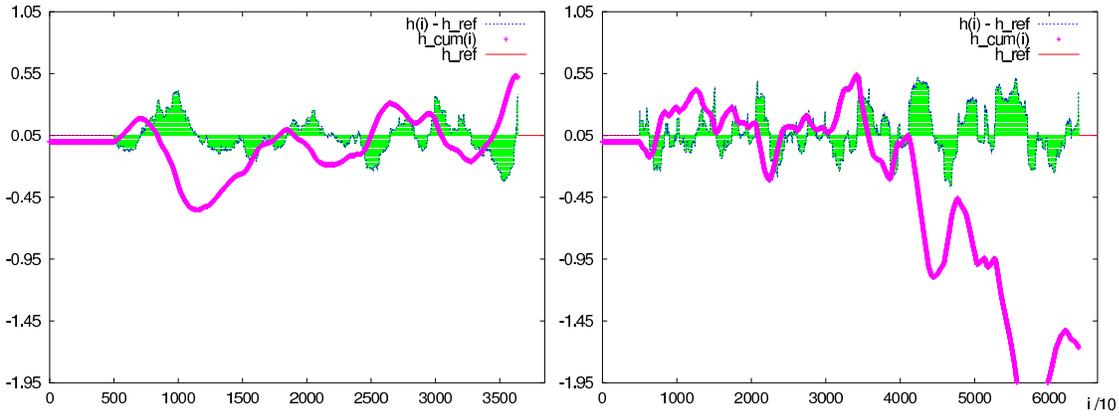
Figure 2. Two different time series of fetal heartbeat intervals were analysed using the cumulative Hölder exponent based real time indicator. The time series correspond with one good outcome (healthy) and one bad outcome (hypoxia). The good outcome can be identified by the indicator oscillating near zero or steadily increasing. In the case of hypoxia, the indicator plunges (down towards negative values). Problems can occur at any moment during labour, even after a stable condition, as is visible in the bad outcome plot. The first part of the plot is flat since this period is required for initialisation - acquiring a reference for the upper value for the Hölder exponent evaluation. Both the cumulative Hölder exponent $h_{cum}(V_i)$ (red line) and the deviation of the Hölder from the reference value $h_{ref} = 0.05$ (blue filled curve) are plotted. The $h_{cum}(V_i)$ has been rescaled by a factor 0.01.

the observation window, we suggested [14,15] using a cumulative indicator, designed to capture the non-stationarity of the local h of the variability component of the fetal heartbeat (in fact the non-stationarity of the short-time (1000-samples) mean of the local h). [4]

The cumulative h is defined from the beginning of the observation and with respect to some normal reference level h_{ref}:

$$h_{cum}(V_i) = -\sum_{l=1}^{i}(h_{eff}(V_l) - h_{ref}) \,. \tag{1}$$

The minus sign is introduced to give the h_{cum} indicator increasing direction when the level of local correlations is lower than h_{ref}. This corresponds with a healthy condition. The case of higher correlations is associated with problems and, therefore, the accumulation of a positive difference $(h_{eff}(V_l) - h_{ref})$ will lead to decreasing cumulative h.

---

[4] To use the linear integral of the non-stationarity component of the variability of the fetal heartbeat is simply a heuristic. Of course, other functional dependence than linear integral is possible but the discovery of a more suitable model of such possible functional dependence is the subject of our future research, involving the techniques described in section 4.

6

We have tested several examples of fetal heartbeats and found a good correlation with the fetal outcome, as determined by the blood tests [14,15]. In figure 2, we plot two cases, one representing good outcomes and one bad outcomes. The cumulative indicator steadily increasing or remaining within some margin of fluctuations indicates no problems and a good prediction. When the indicator plunges down, it calls for intervention. This can, of course, happen at any moment during labour. The nature of this process is dramatically non-stationary, and a period of positive evaluation can be interrupted at any stage (for example by the occlusion of the umbilical cord due to movement). One of the examples shown (figure 2 right) shows the cumulative indicator plunging after a prolonged homeostasis.

4 Reasoning with non-stationarities - data mining or looking for the unknown model

The idea of an automatic system that provides some kind of support to an expert in reasoning about the (experimental) evidence had its origins some thirty years ago (in what initially was called 'expert systems'). However, it was not until the last decade, that the seminal paper by Lauritzen and Spiegelhalter [16] provided the necessary machinery to achieve such an ambitious goal. It made it feasible, in the so-called 'probabilistic expert systems', for the probabilities to be propagated through the variables involved in the reasoning process. Such propagation makes it possible to obtain probabilities assigned to the values of the variables involved in the reasoning problem, given evidence provided about the status of another set of variables (usually exogenous ones).

This framework generalises the logistic model, used very often for medical problems, in two ways. First, because there is no limitation of binary values to variables, but one can arbitrarily assess the values of variables involved when the values of the others changes. Second, because this framework makes it possible to express the complex relationships between the variables in terms of conditional independencies, which goes a step further than the usual hierarchical modelling of the factors in a logistic model [17]. The process of propagating probabilities is fast and makes it possible to obtain a response in real time as new evidence (in medical terms 'symptoms') is entered to the system.[5]

Another use of a probabilistic expert system is the ability to provide insight into the data used.This data may contain clues that can uncover facts relevant to the domain where we are using it (in our case obstetrics). In particular, re-

---

[5] There is already commercial software for this purpose being successfully used (http://www.hugin.com, http://www.norsys.com).

cent work in Bayesian networks and simulation methods, allows the extraction of such valuable information from data.[6]

Thus, the approach we use combines expert knowledge of the subject matter with an automatic recovery of the relationships in the system from the current data available [21,22]. This, in turn, may lead to the discovery of unknown relationships that are relevant to the domain expert.
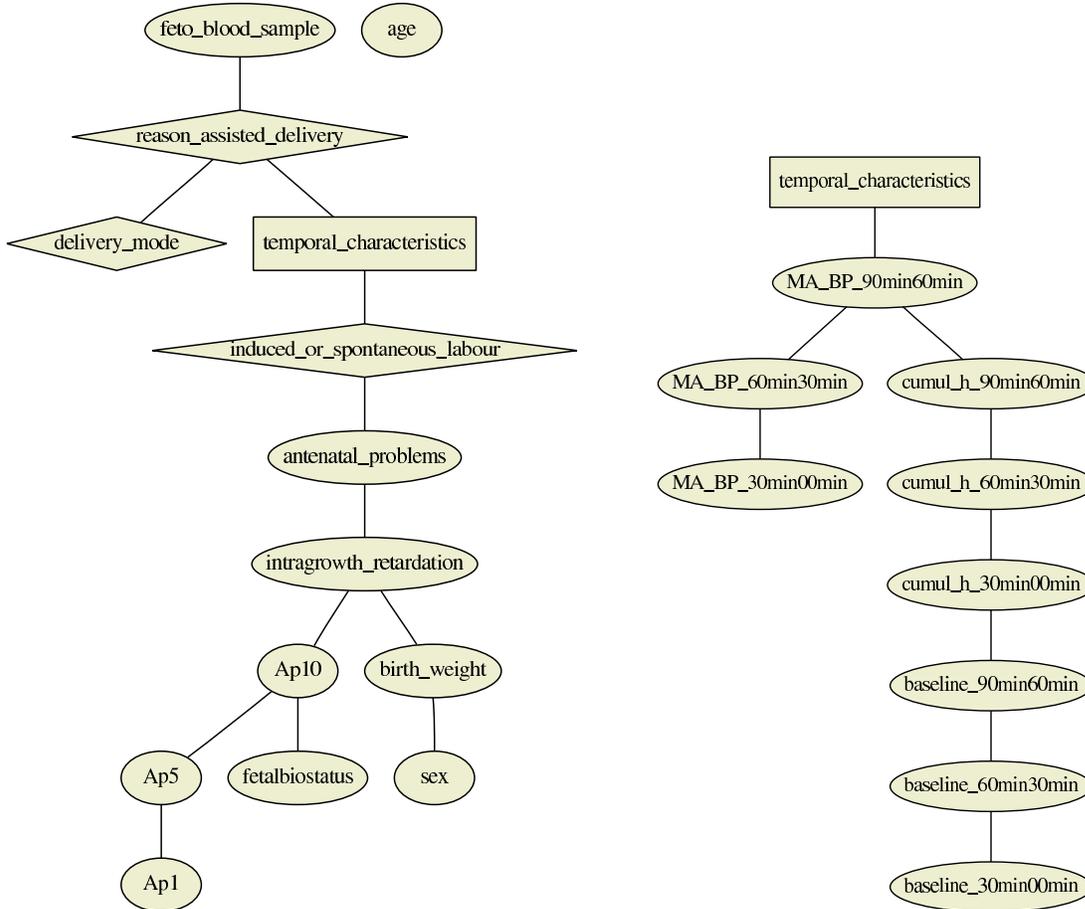

Figure 3. Simplified version of the best dependence network obtained in the experiment. We split the time series sub-network (denoted: 'temporal characteristics', in rectangle) from the main one. Note that heartbeat temporal characteristics are only connected to biochemical variables (fetoblood sample - fbs, fetalbiostatus) and fetal outcome AP* by decisions of the obstretician (marked by the diamond shapes).

It is important to realise that in our approach we are looking for an unknown model, therefore not just model-fitting. This is because we do not know 'a

---

[6] For example, software BUGS [18] (http://www.mrc-bsu.cam.ac.uk/bugs) has proven useful in the Bayesian modeling of complex biomedical systems [19,20].

piori' the dimensionality of the system and do not know the functional dependence between the parameters (experimentally observed, 'input' and 'output' values). In particular we also do not know which parameters are relevant. This approach, especially when 'automated' and applied to large volumes of experimental data, is sometimes referred to as data mining.

In our experiment, about a thousand cases have been investigated, continuous variables have been discretised and time series characteristics derived from the heartbeat for three 1/2 hour intervals just before birth. The time series characteristics comprise [14,15]:

(1) short-time average heartrate level - 'baseline'
(2) short-time average integral of accelerations and decelerations[7]
(3) cumulative Hölder exponent.

Next, models were generated by Monte Carlo simulation and tested by maximum likelihood criterion. The best models were evaluated by human experts. Among the recovered facts, see figure 3, we obtained primary agreement of the best networks with the expert's knowledge.

In particular we discovered that heartbeat features are connected to the fetal outcome variables like AP scores and biochemical variables (fetoblood samples - fbs) via the decisions of the obstretitian. This captures the known fact that the fetal heartbeat characteristics are used by the obstretitian but are not necessarily directly related to the fetal outcome. The exact nature of the correlation between heartbeat characteristics and the fetal outcome is the subject of future work. In particular the functional temporal dependence between the heartbeat characteristics and the fetal outcome is a potentially interesting direction of study.

We have been able to recover correlations between features and their monotonic temporal ordering (dependence). Cumulative h has been demonstrated to play a substantial role in the network of dependencies, on a par with the established monitoring quantities like mean baseline heartrate and acceleration/decceleration integral. We recovered monotonic temporal relationships in the main temporal characteristics of the time series, as well as in the fetal outcome descriptors AP1, AP5, AP10 (the state of the baby 1, 5, 10 minutes after birth).

As a curiosity we have also confirmed less obvious relationships which were validated by the obstretrics expert - the age of the mother seems not to be correlated with other variables like the fetal outcome (our sample contained

_____
[7] These are sudden drops in the heartrate of the fetus, which can be caused by a number of events such as compression of the umbilical cord, increased intracranial pressure of the fetus during contractions or (temporary) hypoxia.

cases spanning 14-42 years of age and effectively covering the interval from 21-35 years - Gaussian distribution with mean 28 and stdev 7).

Also, there seems to be a gender related asymmetry in the recovered model. The gender of the baby is linked to the outcome variables like weight and growth retardation. This seems to confirm a known fact that males are weaker.

5 Conclusions and suggestions for future developments

We have demonstrated a case of 'automated' statistical experimentation motivated by large number of experiments and non-stationarity of model description.

We have elaborated on the example where non-stationarity arises in multifractal description for fetal heartbeat during labour. We have used the failure of the model to provide useful characteristics of the fetal heartbeat in real time during labour.

In particular we have discovered that heartbeat features (including non-stationarity based cumulative h measure) are connected to the fetal outcome variables like AP scores and fbs via decisions of the obstretitian.

We have shown that the heuristic of linear integral (cumulative indicator) of the non-stationarity component of the variability of the fetal heartbeat is relevant. Other than linear integral dependence is possible and may be more accurate for the functional form of dependence on the non-stationarity (measure) of local h. The discovery of such a possibly more suitable model of this functional dependence will be the subject of our future research, involving the techniques described in section 4.

As a closing remark - we have discovered that although men (the male gender) may well be blamed for a crazy life style as the primary cause of their shorter life span, the truth of the matter is, we are predetermined by nature from the very start :-).

References

[1] E. T. Jaynes, chapter 10: Physics Of 'Random Experiments', Probability Theory: The Logic Of Science, http://bayes.wustl.edu/etj/prob.html

[2] D. Heckerman. A tutorial on learning with Bayesian networks. Technical Report MSR-TR-95-06, Microsoft Research, March, 1995 (revised November, 1996). http://research.microsoft.com/scripts/pubdb/pubsasp.asp?recordID=63

[3] P.Ch. Ivanov, M.G. Rosenblum, L.A. Nunes Amaral, Z.R. Struzik, S. Havlin, A.L. Goldberger and H.E. Stanley, Multifractality in Human Heartbeat Dynamics, *Nature* **399**, 461-465, (1999).

[4] P.Ch. Ivanov, L.A. Nunes Amaral, A.L. Goldberger, S. Havlin, M.G. Rosenblum, H.E. Stanley and Z.R. Struzik, From 1/f Noise to Multifractal Cascades in Heartbeat Dynamics, *Chaos* **11** Special Issue on Unsolved Problems of Noise, Ed. Derek Abbott, to appear, (2001).

[5] Z. R. Struzik, Determining Local Singularity Strengths and Their Spectra with the Wavelet Transform, *Fractals*, **8**, No 2, (2000).

[6] Z. R. Struzik. Revealing Local Variablity Properties of Human Heartbeat Intervals with the Local Effective Hölder Exponent. Fractals **9**, No 1, (2001).

[7] M. Kobayashi, T. Musha, 1/f Fluctuation of Heartbeat Period, *IEEE Trans Biomed. Eng.*, **29**, 456-457 (1981).

[8] C.-K. Peng, J. Mietus, J.M. Hausdorff, S. Havlin, H.E. Stanley and A.L. Goldberger Long-Range Anticorrelations and Non-Gaussian Bahavior of the Heartbeat *Phys. Rev. Lett.*, **70**, 1343-1346 (1993).

[9] J. B. Bassingthwaighte, L. S. Liebovitch and B. J. West. *Fractal Physiology*, (Oxford University Press, 1994).

[10] R.G. Turccot, M.C. Teich, Fractal Character of the Electrocardiogram: Distinguishing Heart-failure and Normal Patients, *Ann. Biomed. Eng.* **24**, 269-293 (1996).

[11] M. Meyer, Scaling Properties of Heartbeat Interval Fluctuations in Health and Disease, in *Fractals and Beyond*, Ed. M.M. Novak, pp. 33-42, (World Scientific 1998)

[12] P. Ch. Ivanov, A. Bunde, L. A. N. Amaral, J. Fritsch-Yelle, R. M. Baevsky, S. Havlin, H. E. Stanley, and A. L. Goldberger, Sleep-wake Differences in Scaling Behavior of the Human Heartbeat: Analysis of Terrestrial and Long-term Space Flight Data, *Europhys. Lett.* **48**, 594-600 (1999).

[13] A. Bunde, S. Havlin, J.W. Kantelhardt, T. Penzel, J.-H. Peter and K. Voigt, Correlated and Uncorrelated Regions in Heart-rate Fluctuations During Sleep, *Phys. Rev. Lett.* **85**, 3736 (2000).

[14] Z. R. Struzik, W. J. van Wijngaarden, Cumulative Effective Hölder Exponent Based Indicator for Real Time Fetal Heartbeat Analysis During Labour. In M. M. Novak, Emergent Nature: Fractals 2002, (World Scientific, 2002).

[15] Z. R. Struzik, W. J. van Wijngaarden. Cumulative Effective Hölder Exponent Based Indicator for Real Time Fetal Heartbeat Analysis During Labour. Technical Report INS-R0110, CWI, Amsterdam, The Netherlands, November 2001. Available from `www.cwi.nl/~zbyszek`

[16] S.L. Lauritzen, D.J. Spiegelhalter, Local Computations with Probabilities on Graphical Structures and their Applications to Expert Systems (with discussion). *Journal of the Royal Statistical Society, Series B*, **50**, 157-224, (1988).

[17] J. Whittaker, Graphical Models in Applied Multivariate Statistics, (Wiley, New York, 1990).

[18] W.R. Gilks, A. Thomas, D.J. Spiegelhalter, A Language and a Program for Complex Bayesian Modelling. *The Statistician*, **43**, 169-178, (1994).

[19] D.J. Spiegelhalter, Bayesian Graphical Modelling: A case-study in Monitoring Health Outcomes. *Applied Statistics*, **47**, 115-133, (1998).

[20] H.H Hundborg, M. Hojbjerre, O.B. Christiansen, S.L. Lauritzen, Familial Tendency to Fetal Loss Analyzed with Bayesian Graphical Models by Gibbs Sampling. *Statistics in Medicine*, **19**, 2147-2168, (2000).

[21] T. Kocka and R. Castelo, Improved Learning of Bayesian Networks. In *Proceedings 17th Conference on Uncertainty in Artificial Intelligence*, 269-276, (Seattle, Morgan Kauffman, 2001).

[22] P. Giudici and R. Castelo, Improving Markov Chain Monte Carlo Model Search for Data Mining, to appear in *Journal of Machine Learning*, (Kluwer Academic Publishers, 2001).