# Implementing Multidimensional Inference Capability in Vibration Condition Monitoring[1]

**CEMPEL[2] Czeslaw**

**Abstract**

Contemporary measuring technology in condition monitoring of critical systems allow us to form diagnostic symptom observation vector, with components different physically, and to extract fault information from such created symptom observation matrix. This is possible by using singular value decomposition algorithm and specially written program, which enable also to optimize the dimensionality of symptom observation vector, and to extract needed diagnostic information. We can use as the next, the concept of symptom reliability and symptom hazard rate to calculate the symptom limit value, for system maintenance planning and execution. It seems to be possible to perform these task in an autonomous way, and adding also the knowledge base and learning loop, creating in this way some first approach to Condition Inference Agent (CIA).

## 1. Introduction

Condition monitoring of critical machinery depends on observation of some **symptoms**[3], (like amplitudes of vibration, the temperature, etc), and comparing them with their limit values, usually determined by some long term experience. In most cases, even for sophisticated machinery like turbo set, every measurable symptom $\mathbf{S_i}$ is monitored and assessed separately, by its specific symptom **limit value** $\mathbf{S_{il}}$. But contemporary advances of measuring technology, connected with intelligent sensors allow us to measure and process several symptoms at the same time. Moreover, we can have also as measured some parameters of system operation, like mechanical or electrical load, the temperature, etc, or at least the system **lifetime** counter $\theta$, as the first assessment of just enumerated components of so called logistic vector **L**, (see for example [CempelNatke93]).

In this way we can form symptom observation vector with many components, and measure it sequentially over the span of system life; $0 \le \theta \le \theta_b$, with each row as separate observation of symptom vector. This gives us so called Symptom Observation Matrix (**SOM**), with columns being the component of observed symptom vector **S,** and rows as successive observation; $\mathbf{S}(\theta_1)$, $\mathbf{S}(\theta_2)$, … . In other words, we have **multidimensional** symptom space for system condition monitoring, and in the theory it is possible to extract from this symptom observation space, the full description of system degradation taking place during its life. As was shown in [Cempel02] using singular value decomposition (**SVD**), and lately also principal component analysis (**PCA**) [CempelTabaszewski03], it is possible to decompose information contained in SOM into information descending independent components called **generalized symptoms**, which seems to describe independent **faults** evolving in an operating system.

As one can suspect some symptoms can be more diagnostic oriented in a given case, so there is optimization task and challenge to provide along with the condition assessment. This challenge concerns also the determination of symptom limit values $\mathbf{S_l}$, as after the decomposition of SOM we have no longer originally measured symptoms, but some generalized ones. However basing on **symptom distribution** theory elaborated initially in [Cempel91,ch4] [NatkeCempel97,ch2.3], and later **symptom reliability** and hazard in [CempelNatke00], we can solve this problem basing only on currently assessed SOM. Finally it seems to be possible to implement a learning loop into just described methodology and to try to develop some Condition Inference Agent (CIA) for diagnostics of **unitary** critical systems. These problems are described, some of them solved and synthesized in this paper, along with some computational program, which makes much easier this enormous task.

---

[1] Paper intended to Surveillance Conference, Senlis October 04

[2] Poznań University of Technology, ul Piotrowo 3, 60-965 Poznań, Poland, email:czeslaw.cemepel@put.poznan.pl

[3] Symptom is measurable quantity covariable with system condition

## 2. Multidimensionality in Condition Monitoring and Symptom Observation Redundancy

Let us take into consideration a critical machine in operation. During its life $0 < \theta < \theta_b$ , ($\theta_b$ – anticipated breakdown time), several independent **faults**; $F_t(\theta)$, t = 1,2,..u, are growing. We would like to identify and assess these faults by forming and measuring the symptom observation vector; $[\mathbf{S_m}] = [\mathbf{S_1,...,S_r}]$, which may have components different physically, like vibration amplitudes, temperature, machine load, etc. In order to track machine condition evolution (faults), we are making equidistant reading of above symptom vector in the lifetime moments; $\theta_n$, n = 1, ... p, $\theta_p \le \theta_b$ , forming in this way the rows of symptom observation matrix (SOM). From the previous research and papers [Cempel99], [Cempel03] we know that the best way of SOM pre processing is to center it (remove), and normalize (divide it) to symptom initial value; $\mathbf{S_m(0) = S_{0m}}$, of each given symptom (column of SOM). From these research it is also known that amount of diagnostic information in SOM increases if we append the lifetime $\theta$ column, as the first approximation of system logistic vector **L**. This gives us dimensionless symptom observation matrix in the form

$$O_{pr} = [S_{nm}], \qquad S_{nm} = \frac{\mathbf{S}_{nm}}{\mathbf{S}_{0m}} - 1, \tag{1}$$

where bold letters indicate primary dimensional symptoms, as taken directly from measurements.

As it was already said in the introduction, we apply now to the dimensionless SOM (1), the Singular Value Decomposition (SVD), and principal component analysis (PCA), in the form

$$O_{pr} = U_{pp} * \Sigma_{pr} * V_{rr}^{T}, \quad \text{(T- matrix transposition )}, \tag{2}$$

where $U_{pp}$ is p dimensional orthogonal matrix of left hand side **singular vectors**, $V_{rr}$ is r dimensional orthogonal matrix of right hand side singular vectors, and the diagonal matrix of **singular values** $\Sigma_{pr}$ is as below

$$\Sigma_{pr} = \text{diag}(\sigma_1, \dots, \sigma_l), \text{ and } \sigma_1 > \sigma_2 > \dots > \sigma_u > 0, \tag{3}$$
$$\sigma_{u+1} = \dots \sigma_l = 0, \quad l = \max(p, r), \quad u = \min(p, r).$$

The above means, that from the **r** measured symptoms we can extract only $\mathbf{u \le r}$ independent sources of diagnostic information describing evolving generalized faults $F_t$ , (see Fig. 1). As it is seen from Fig. 1 upper left picture, only a few developing faults are making essential contribution to total fault information, the rest of generalized faults are below the level of 10% noise. What is important here, that such SVD decomposition can be made currently, after each new observation of the symptom vector; n = 1, … p, and in this way we can trace the faults evolution in a system (see Fig. 5). From the current research of this idea [CempelTabaszewski03], we can say that the most fault oriented indices obtained from SVD/PCA is the first pair: ($SD_t$ , $\sigma_t$ ) , and the sum of all indices; $\mathbf{SumSD_i}$ , $\mathbf{Sum\sigma_i}$. The first fault indices $\mathbf{SD_t}$ can be named as discriminant of the fault **t**, and one can get it as the SOM product and singular vector $v_t$ , as below

$$\mathbf{SD_t = O_{pr} * v_t = \sigma_t \cdot u_t} \, . \tag{4}$$

We know from SVD theory, that all singular vectors $v_t$ are normalized to one, so the energy norm of new discriminant is simply

$$\text{Norm}(SD_t) \equiv \| SD_t \| = \sigma_t. , \; t = 1, ...,u. \tag{5}$$

The above discriminant $SD_t(\theta)$ can be also named as **fault profile**, and in turn singular value $\sigma_t(\theta)$ seems to be its advancement (energy norm).

The similar inference can be postulated to the meaning, and the evolution, of summation quantities, what can mean the total damage profile $\boldsymbol{Sum}\mathbf{SD_i(\theta)}$, and total damage advancement $\boldsymbol{Sum}\mathbf{\sigma_i(\theta)}$,

$$SumSD_i(\theta) = \sum_{i=1}^{z} SD_i(\theta) = \sum_{i=1}^{z} \sigma_i(\theta) \cdot u_i(\theta) = P(\theta),$$

$$Sum\sigma_i(\theta) = \sum_{i=1}^{z} \sigma_i(\theta) \sim \sum_{i=1}^{z} F(\theta)_i = F(\theta) \, . \tag{7}$$

But the last relation seems to be not fully validated as yet, and it seems to also, that the condition inference based on the above summation measures; ***Sum(*)*** may stand for the first approach to multidimensional condition inference, as it is clearly seen from the Fig 1. Here railroad Diesel engine named **sil54d2** was diagnostically monitored by vibration measurements[4] performed on the top of one of cylinders each ten thousand kilometer of mileage, up to the breakdown. Altogether 12 vibrational parameters were gathered in the symptom observation vector, beginning from the three acceleration amplitude measures; avg, rms, peak, three velocity, displacement, and Rice frequency measures, with the first component in the vector being always the engine lifetime $\theta$.

In this way **SOM** of the engine has 13 columns, and as it is seen from the upper top right picture it contains information concerned with several faults $F_t$ , but only two of them are prevailing the **10%** level of noise. The top right picture presents the life behavior of symptoms in SOM; from 0 km mileage up to the engine breakdown at 250.000km, together with the straight line being the course of the engine life $\theta$ . The bottom left picture shows the course of summation generalized fault discriminant ***SumSD$_i$*** , and **SD$_1$** below it, and again the rest of generalized fault discriminant is on the level of noise, near the zero line. The last picture, the bottom right shows the course of singular values $\sigma_i$ , here the prevailing information is contained again in the summation discriminant, and the first one $\sigma_1$, but the second singular value $\sigma_2$ grows substantially only after 100 thousand km mileage.

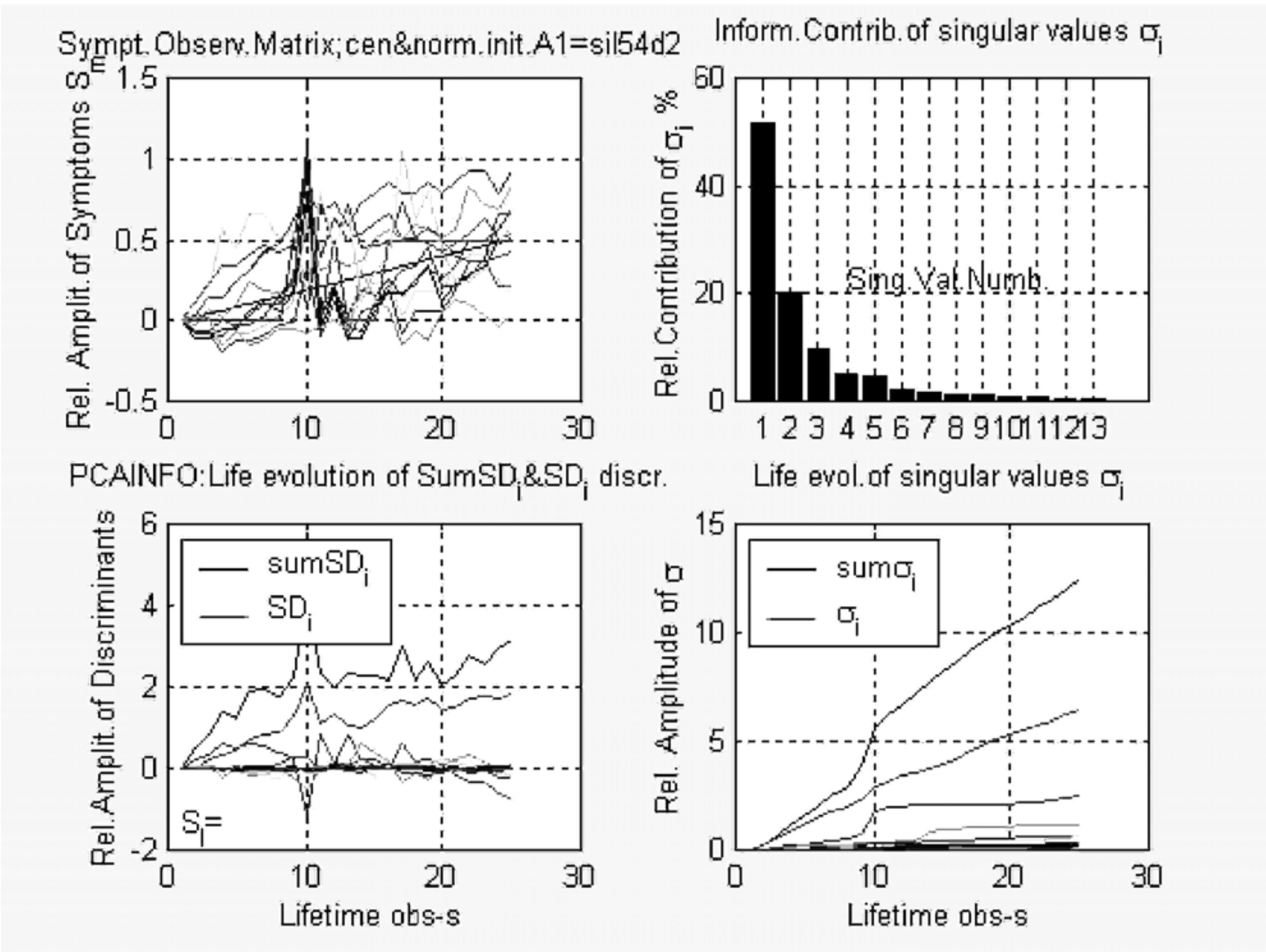


Fig. 1. The information contents of symptom observation matrix for a Diesel engine **sil54d2**, and independent fault indices **SD$_i$** , $\sigma_i$ as discovered by SVD/PCA computation program.

[4] Author is grateful to Prof. F. Tomaszewski for borrowing the data.

Now we can ask the question, the 13 component of symptom observation vector and only one significant generalized fault can be observed ! Hence, it has to be great redundancy in our symptom observation space and some of measured symptoms can be omitted, can one say which one, and how many of them? Next Fig. 2 gives some answer to this problem presenting two pictures, upper one with assessment of information contribution given by each symptom to the overall information resource in the SOM. One can notice clearly here, that last three symptoms (11 – 13), being the Rice frequencies of engine vibration can be really omitted. More detailed information on the contribution of each symptom to $SD_1$ discriminant is shown on the bottom picture of Fig. 2. We can notice here, that again symptoms 11 – 13 are fully redundant, and the most informative symptom in our symptom observation vector is No 3. That means the root mean square acceleration amplitude $A_{rms}$ , and the next one no7 the peak vibration velocity $V_{peak}$ are essential, and the same is true for overall information contribution. Also the life symptom θ (no 1) gives quite substantial amount of information to the overall resource in **SOM**, as well as to $\mathbf{SD_1}$ .

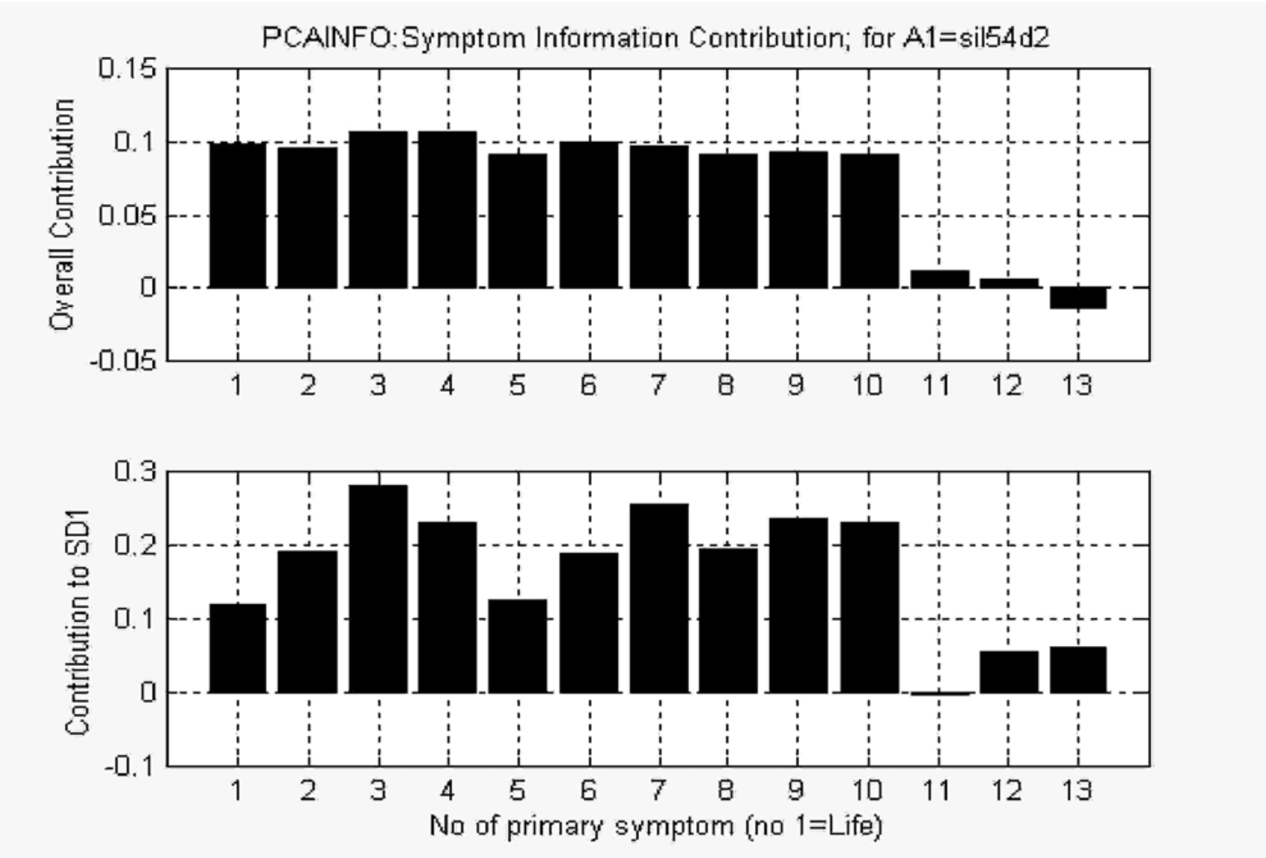


Fig. 2. The redundancy assessment of symptoms in the symptom observation space for the engine **sil54d2** (no 11 – 13 redundant).

Another question can be posed with respect o observation vector, namely what kind of symptom we should chose to minimize observation redundancy ? May be to change the sensor location is enough, and we can use $A_{rms}$ or similar symptom with large life dynamics in a different places of our object ? This problem addresses the next figure 3, where similar diesel engine were monitored vibrationally, by measuring $A_{rms}$ at the top of each cylinder. It is seen from all pictures of the Fig. 3, that by measuring only one symptom, even with sensor separation over half a meter, gives us the information on the same fault only, and there is no gain in multiplying another sensor location. New information can be brought only by the new symptom which is different physically or has quite different frequency spectrum (acoustic emission, ultrasound, etc). We can use the same vibration symptom only if the damping of vibration in the structure is substantial, giving **no leakage** of information among sensors.

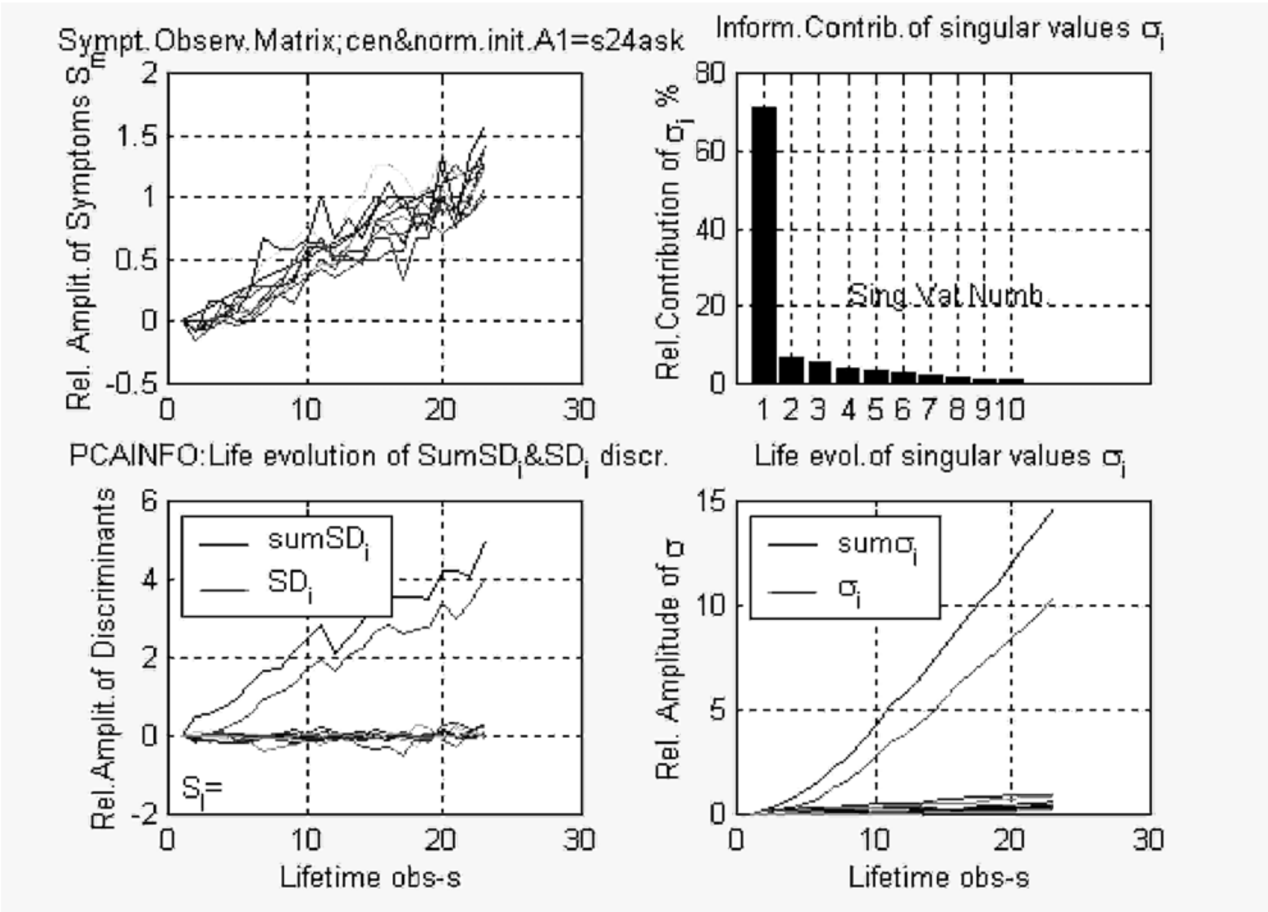


Fig. 3. Engine fault description and differentiation by 9 sensors measuring the same symptom $A_{rms}$, but located on different cylinders of another railroad diesel engine **S24** of the same type.

### 3. Generalized Symptom Reliability for Condition Inference

Looking at Fig 1 and / or Fig 3 bottom left pictures we may know now the course of generalized symptoms in each particular case of multidimensional observation of critical mechanical system. We can also exclude redundant symptoms not carrying useful diagnostic information (see Fig. 2). But how to proceed with diagnostic inference and elaboration of "**go/no go**" maintenance decisions? On what basis we can determine the limit values $\mathbf{S_l}$ for generalized symptoms $SD_1$, or ***SumSD$_i$*** or both, shown on bottom left pictures of Fig.1 and 2 ?
But we can make the statistics of observations from the calculated generalized symptoms. Being more specific the cumulative distribution of generalized symptom of machine being in **good condition**. It was shown by the present author in several papers that such cumulative symptom distribution is equivalent to **symptom reliability R(S)** [Cempel91], [CempelNatke00], and we can get from this also the new quantity called symptom **hazard rate**. Not going into the theory presented elsewhere, the symptom reliability can be used for determination of symptom limit value $\mathbf{S_l}$ by using Neyman-Pearson rule of statistical decision theory [Cempel91,ch4.3]. If we determine, or assume, the allowed probability of unneeded (erroneous) repair of machines being in good condition, say **A**, knowing also the needed availability of the machine set, say **G** , so formula leading to determination of the symptom limit value $\mathbf{S_l}$ is simply

$$\mathbf{G \bullet R(S_l) = A\ , \quad \Rightarrow \quad S_l = f\,[\,R(S), A, G,\,]\ .} \tag{8}$$

It seems to be simple to carry such calculation by some statistical program, Matlab® for example, moreover it was show also [Cempel91] that symptom reliability can be transformed to average symptom life curve $\mathbf{S(\theta/\theta_b)}$ defined in the dimensionless system lifetime. The result of such calculations is shown on pictures of Fig. 4, where scale of symptom value was enlarged by **+1** due to calculation convenience.

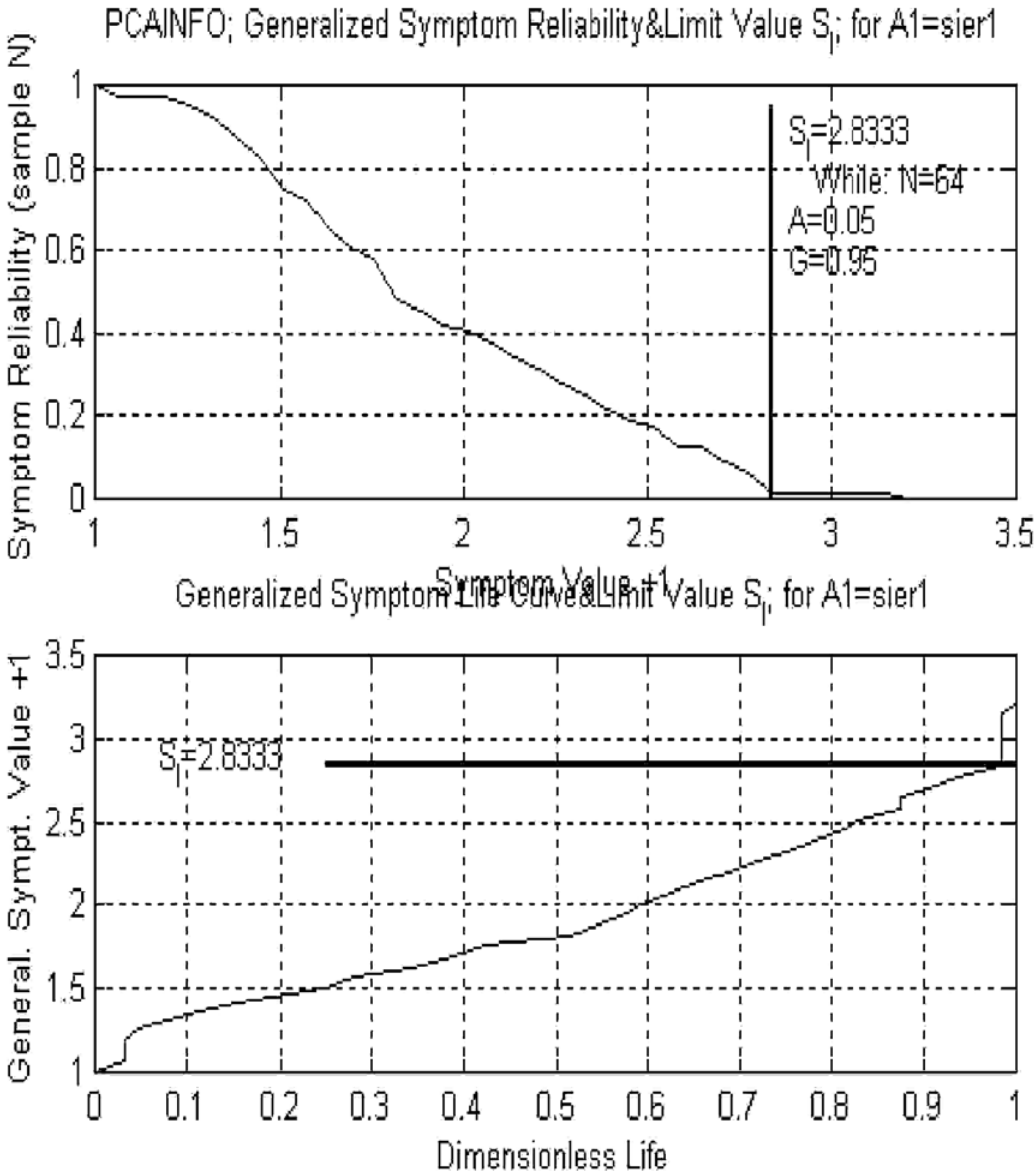


Fig. 4 Generalized symptom reliability and generalized symptom life curve of industrial fan observed multi dimensionally

From these two pictures one can conclude, that both of them can be used for effective condition inference and condition based maintenance. The generalized symptom reliability allows us to assess the symptom limit value $\mathbf{S_l}$ **,** while generalize symptom life curve enable us to trace the life evolution of our critical system, and to make right maintenance decision just on time.

**4. Real Time Inference and Optimization of Symptom Observation Matrix**

We have shown above being in the multidimensional observation space that it is possible to optimize the dimensionality of symptom observation vector, keeping its redundancy as small as possible (see Fig. 2). It was shown also, that we can calculate the symptom limit value $\mathbf{S_l}$ and average symptom life curve $\mathbf{S(\theta)}$ for multidimensional inference (see Fig. 4). And now one can ask, well it was possible to show it at the end of system life time, just near $\theta_b$ , but what about the beginning of system operation. Can we make the same after the beginning of system observation ?

It seems to be, that in general it depend on the smoothness of system operation and its loading, but just to show how it was in the elaborated cases on Diesel engines, please analyze the sequence of building the fault discriminant $\mathbf{SD_1}$ and ***SumSDi*** , when the new row of observation has been added into **SOM** of the industrial fun, with very unstable operation. Even that, we can observe that generalized discriminants are stabilizing just after tens of observation, like on Fig. 5. Much more smooth is symptom reliability curve as well as average symptom life curve, as it can be seen from Fig. 4.

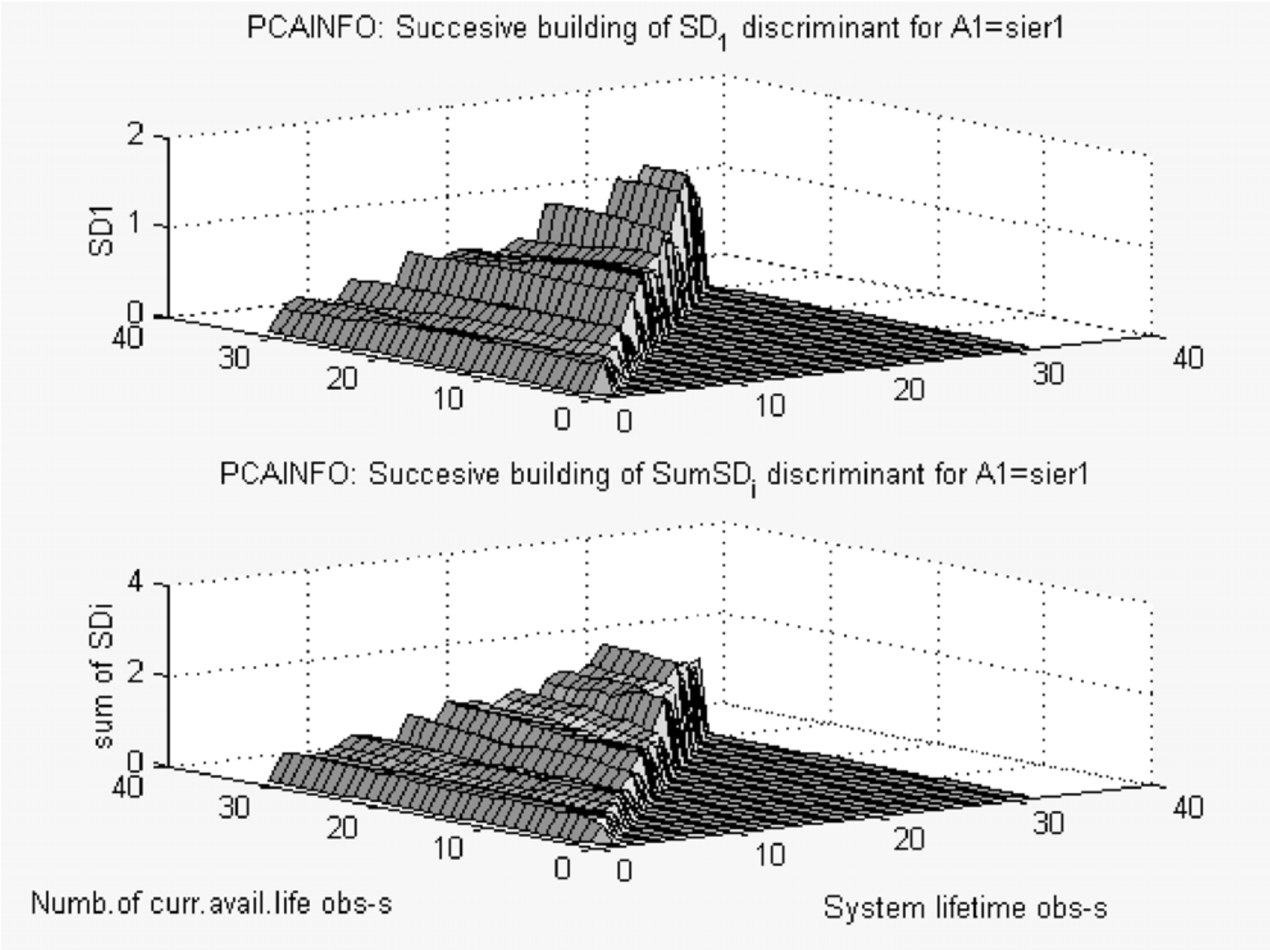


Fig 5. Successive building of fault discriminants during sequential increase of system life (observations)

Summing up, this problem of real time observation and real time inference on system condition, it seems to be workable, and we can elaborate all problems step by step as below.

1. Chose the set of condition related symptoms from the primary group of measured symptoms,
2. extract condition related information from the set of monitored symptoms,
3. to form generalized fault symptoms as the image of evolving faults in a system
4. to assess currently the limit values for each generalized fault
5. to assess the system condition and make proper maintenance decision
6. to perform condition forecasting on the basis of acquired object related specific knowledge, some general knowledge, and to communicate and implement it.

When this work will precede automatically, by means of some learning loop, one can say we have some Condition Inference Agent (**CIA**). It seems to the present author that realization of this task is not far away goal, but only next step in the intelligent multidimensional diagnostic observation of critical system. How may it proceeds in general is shown on the next Fig. 6, and one can see that we must learn how to incorporate learning loop into **CIA**, and how to build and implement diagnostic knowledge base for a specific critical system. All of this is ahead of us, and with contemporary knowledge [Cichosz00], [Agent04], we can make it workable soon.

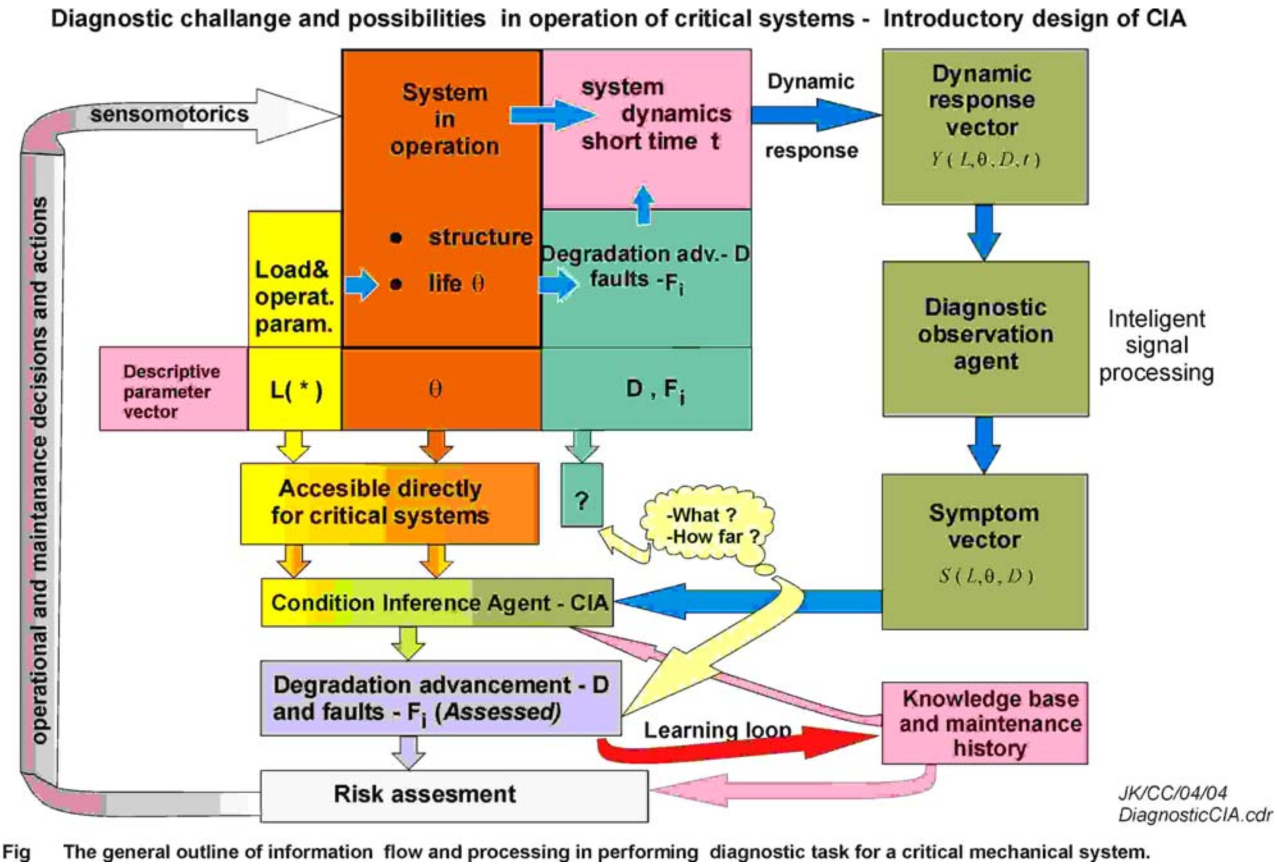


Fig. 6 Anticipated information flow in the diagnostic observation and computation in the design of future Condition Inference Agent

## 5. Conclusions

The paper is some synthesis of contemporary work on multidimensional diagnostics of critical systems. It was shown here, that we can form symptom observation vector with many components, being the basis for symptom observation matrix (**SOM**). On the basis of **SOM** and singular value decomposition (**SVD**) we can extract all condition related information, and optimize the dimensionality of symptom observation vector. Starting from generalized fault discriminants we can form (in real time) the symptom reliability **R(S)**, for estimation of symptom limit value $\mathbf{S_l}$, , which enable us to infer on system condition and make right maintenance decision. It seems to the present author, that this task can be made by some autonomous software entity called Condition Inference Agent – **CIA**, and right now this is ahead of us.

**Acknowledgment**

Author is indebted to Dr M. Tabaszewski for his current help in improving Matlab® **pcainfo.m** program.

## 6. References

1. Cempel C., Natke H. G., System life cycle, system life, the model based technical diagnostics; a view on holistic modeling, **System Research**, Vol. 10, No 1, 1993, pp 53-63.
2. Cempel C., Multidimensional Condition Monitoring of Mechanical Systems in Operation, **Mechanical Systems and Signal Processing**, Vol. 17, No 6, 2003, pp 1291-1303.

3. Cempel C., Tabaszewski M., Krakowiak M., The Extraction Methods of Multidimensional Diagnostic Information, Proceedings of **XXX Conference on Machinery Diagnostics**, Węgierska Górka, March, 2003.
4. Cempel C., **Vibroacoustic Condition Monitoring**, Ellis Horwood, London 1991, p 212.
5. Cempel C., Natke H. G., Yao J. P. T., Symptom Reliability and Hazard for Systems Condition Monitoring, **Mechanical Systems and Signal Processing**, Vol. 14, No 3, 200, pp 495-505.
6. Natke H. G., Cempel C., **Model-Aided Diagnosis of Mechanical Systems**, Springer Verlag, Berlin, 1997, p 248.
7. Cempel C., Innovative Developments in Systems Condition Monitoring, Keynote Lecture for **DAMAS '99**, Damage Assessment of Structure Conference, Dublin, June 1999.
8. Cempel C., Pareto Law of Damage Evolution in Application to Vibration Condition Monitoring, **Bulletin of the Polish Academy of Sciences**, Techn. Sci.Vol.39, No4, 1991, pp 574- 588.
9. Cichosz P., **Self Learning Systems**, WNT, Warsaw, 2000, p894.
10. Agent Builder®, Build it smarter, Internet 2004; http://www.agentbuilder.com/