

PROBABILITY MEASURES OF LIKELIHOOD OF DIAGNOSIS OF THE TECHNICAL STATE OF MAIN COMBUSTION ENGINES OF SEA-GOING SHIPS

Jerzy Girtler

Gdansk University of Technology
Faculty of Ocean Engineering & Ship Technology
Department of Ship Power Plants
phone: +48 58 347-24-30, fax : +48 58 347-19-81
e-mail: jgirtl@pg.gda.pl

Abstract

The paper presents reasons for the need of discerning the notions: likelihood of diagnosis and rightness of diagnosis while taking operating decisions. A probability formula for establishing the right diagnosis has been derived as a likelihood measure of diagnosis.

For deriving the formula, the theory of semi-Markov processes and Bayes' formula of the conditional probability have been applied. Other probability measures of likelihood of diagnosis have been also presented. These measures concern the technical state of such important systems as combustion engines of sea-going ships. However, they can be of use also for other ship systems.

The paper provides description of a stochastic decision situation following from operation of combustion engines in any operating system and on this background there is presented some proposals of applying the technical diagnostics for controlling the engine operating process.

Formulating a problem for the combustion engines, solving the problem, diagnostic inference and likelihood of diagnosis as well a simplified model of combustion engine operation taking into account diagnosis and control in time of operation are presented in the paper.

Keywords: *combustion engine, diagnostic system, likelihood of diagnosis, theory of semi-Markov processes*

1. Introduction

Diagnostic inference [1, 3, 8] in a stochastic decision situation enables establishing a diagnosis at determined likelihood. Knowledge of the likelihood of diagnosis is indispensable to take rational operating decision – which is an operating decision worked out in result of applying the optimum calculus [7, 10]. It is also known that any arbitrary decision (decision on initiation of a determined action and holding it up to the time of achieving the wanted result which can be, but does not have to be the same as the assumed aim of the action) should be taken only after analysis of effects of the action. Deciding resulting in taking a decision should be understood as making a non-random selection, when up to the moment the decision is taken the probabilistic and stochastic measures of phenomena, events and processes occurring in the phase of using the system, are applied as well as the likelihood measures of diagnosis [4, 7-8]. These measures are indispensable to elaborate decision information (J_D) which makes taking the decision possible, e.g. the decision whether the defined ship can be used to perform the given task or needs renovation of its state first to be able to perform the task. The complexity of the situation in which the mentioned information is being elaborated, has been presented in Fig. 1 [3].

The likelihood of diagnosis is defined in different ways [8]. It can be accepted that the likelihood of diagnosis is [3, 5, 8-9]:

- *in descriptive meaning*, a characteristic of diagnosis, describing the degree of identification of the real (past, actual, future) state (technical and energetic, which together form the operating

state) of a diagnosed system (SDN), e.g. main combustion engine of a sea-going ship, by a diagnosing system (SDG),

- *in valuing meaning*, a characteristic of diagnosis, being expressed with values of essential, in particular cases, indexes that describe the degree of identification of the state of SDN by SDG. The identification consists in classifying the real state of SDN by SDG to the known class of diagnostic model states. Such classification equals taking a diagnostic decision (DG) and in consequence – an operating decision.

The indexes that describe the degree of identification of the state of SDN, are following:

- probability of taking the right diagnostic decision (of elaborating or establishing the right DG),
- ratio of the expected number of identified (in fixed time) states of SDN by SDG to the expected number of really occurred states of the same kind (in the same time interval) of SDN,
- expected value of relative frequencies of elaboration of the right diagnosis.

Accepting that the essential index describing (defining) the degree of identification of the real state of SDN by SDG, is the mentioned probability of elaborating the right DG, the likelihood of diagnosis can be defined (in valuing meaning) as follows: reliability of diagnosis is a probability of taking the right diagnostic decision (the probability of establishing the right diagnosis), so the probability of classifying the supposed real state (the state being identified by SDG) of SDN to such a class of diagnostic model states, to which this real state belongs and should be classified by SDG.

2. Formulating a problem for the combustion engines

In operating practice, decisions concerning operation of combustion engines on sea-going ships are taken by the users of diagnosis (decision-makers) at different stages of a diagnosing process. That results from work of SDG which in case of the same type of combustion engines (Fig. 1), can be differently fitted to diagnostic inference. Among operated diagnosing systems (SDG) that are fitted to diagnose a main combustion engine of sea-going ships, as SDN (diagnosed system), there are such ones which can be named: complex and local [8]. In the both cases, diagnostic inference consists of: measuring, symptomatic, structural and operating inferences [1]. In each kind of the inference there are made mistakes which accumulate and influence the likelihood of diagnosis negatively. From SDG's work results that first of all the likelihood of diagnosis depends on (Fig. 1):

- technical state of SDG and especially, possibilities of failures of this system,
- possibilities of occurring SDN's states which have not been considered in the diagnostic task (Z_D),
- possibilities of changes (random changes above all) of the admission vector (Z_D) and the control vector (Y), and changes of the characteristics of random perturbations (K_G), what in consequence means that the mentioned Z_D, Y, K_G can be known only in approximation at the time of diagnosing SDN's state,
- SDG sensitiveness to changes of Z_D, Z_G and Y and existence of K_G .

When SDG is in full (total) ability during diagnosing SDN's state and none of SDN's states other than considered in Z_D appears at this time, accuracy of the diagnosis can be assumed. This results from the fact that the accuracy of diagnosis can be understood (in descriptive meaning) as a characteristic of the diagnosis, describing the degree of identification of SDN's real state by SDG on the condition that SDG is in full ability and none of SDN's states other than considered in Z_D appears. Such understood accuracy of diagnosis enables interpreting it (in valuing meaning) as a characteristic of diagnosis expressed with values of, essential for particular cases, indexes which describe the degree of identification of SDN's state by SDG on the condition that the SDG is in full ability and none of SDN's states other than considered in Z_D appears. Thus, elaboration of the likelihood measure of diagnosis should take the mentioned condition into account where the measure of accuracy of diagnosis can be obtained.

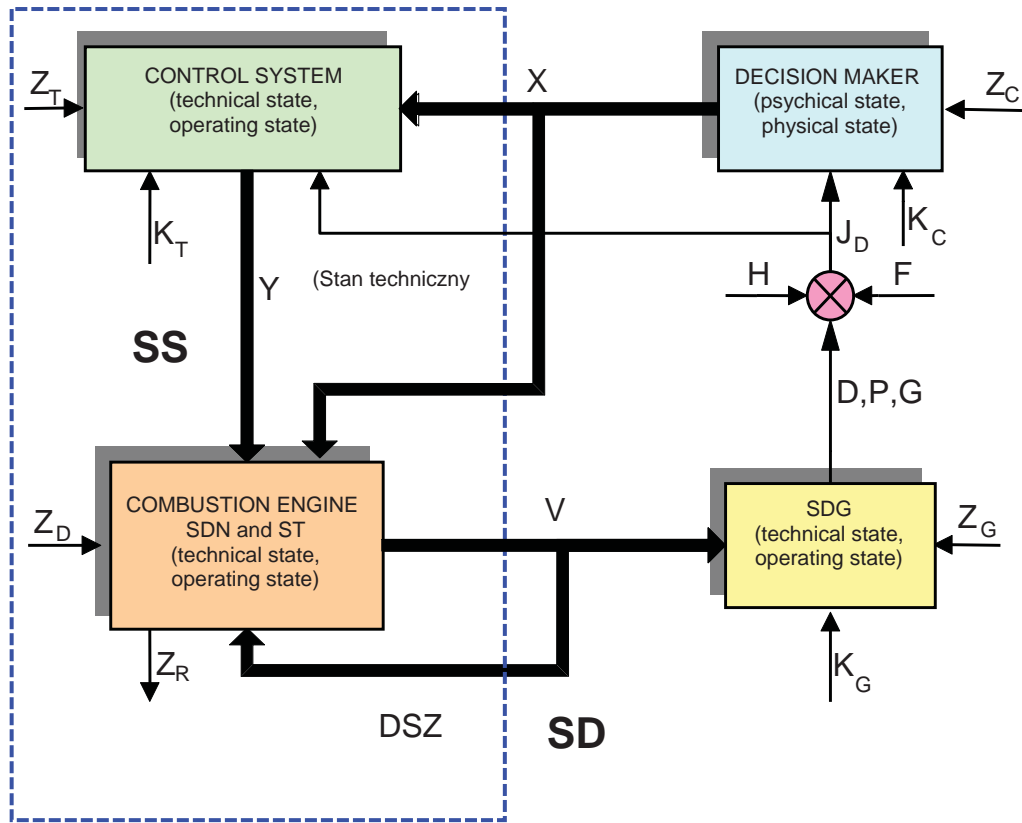


Fig. 1. A simplified model of combustion engine operation taking into account diagnosis and control in time of operation: *D* - diagnosis, *DSZ* – destructive feedback, *F* – criterion function, *G* – genesis, *H* – operating constraints, *I_D* – decision information, *K_C* – disturbances in taking operating decisions, *K_D* - disturbances to diagnosed system (SDN) and control system (ST), *K_G* – disturbances to diagnosing system (SDG), *K_T* – disturbances to control system, *P* – prognosis, *SD* – diagnostic system (SDN and SDG), *V* – vector of initial process, *X* – vector of decisions, *Y* – vector of control, *Z_C* – supply for decision maker, *Z_D* – delivered power supply, *Z_G* – power supply for diagnosing system, *Z_R* – lost power supply, *Z_T* – power supply for control system

From the mentioned reasons follows that after performance of *n*-diagnostic tests and inferences with formulated *n*-diagnoses of SDN's state (e.g. technical state) by using the proper SDG, the right diagnoses can be established (thus, the expected SDN's state can be rightly said that is the same as real or belongs to the given class of the model states) in the quantity: $m < n$. The other quantity: $k = n - m$ indicates that the diagnoses have not been right. This means that the following quantity can be accepted by intuition as the measure for establishing the right diagnosis (measure of likelihood):

$$h = \frac{m}{n} = 1 - \frac{k}{n}. \quad (1)$$

Considering that at the known number *n* of expected (planned) diagnostic tests and inferences we do not know what value *m* takes, therefore for determining likelihood randomness of obtaining the right diagnosis on SDN's state should be taken into account. Hence, it can be admitted that establishing the right diagnosis is a random event because during performance of test and diagnostic inference in determined conditions it can appear but does not have to. In case when the event of formulating the right diagnosis was often observed in the past, a great chance can be assumed that it will also occur in future (in the same conditions). This chance can be defined with probability of taking the right diagnostic decision (establishing the right diagnosis) which therefore can be accepted as a likelihood measure of diagnosis.

Considering the possibility of formulating the right diagnosis of SDN's state in the time interval $[0, t]$ two random variables can be tested: $N(t)$ which determines the number of possible-to-perform tests and diagnostic inferences, and $M(t)$ which determines the number of possible-to-formulate right diagnoses. During such tests the expected values of $E\{N(t)\}$ and $E\{M(t)\}$ can be defined, too. Thus, the likelihood measure of diagnosis can also be a quantity which can be called a likelihood index and defined with the following formula:

$$w = \frac{E\{M(t)\}}{E\{N(t)\}}. \quad (2)$$

The presented likelihood measures of diagnosis do not reflect clearly the fact that identification of the state S of SDN by SDG proceeds in consequence of observing the adequate vector K of values of diagnostic parameters – being generated by SDN, e.g. a ship combustion engine (Fig. 1). They do not reflect also the fact that SDG can, just like SDN, be in different states which the likelihood of diagnosis depends on. Thus, such formula determining the probability $P(S/K)$ should be derived so that would reflect these mentioned facts.

3. Solving the problem

The process of applying SDG is the process $\{W(t): t \geq 0\}$ of which values can be elements of the set:

$$D_d = \{d_1, d_2, d_3\}, \quad (3)$$

with the following interpretation:

- d_1 - state of active use (u) of SDG (operation state of this system), which appears when SDG is in the state of full ability (s_1), thus d_1 means diagnosing of SDN's state by SDG when SDG is in the state s_1 , so $d_1 = (u, s_1)$,
- d_2 - the state of active use (u) of SDG, which appears when SDG is not in the state s_1 , but in the state $\sim s_1$, so in the state of partial ability (s_2) or in the state of disability (s_3) being the state which makes formulation of a right diagnosis impossible, so the state $d_2 = (u, \sim s_1)$,
- d_3 - state of active use (u) of SDG, which appears when SDG is in the state s_1 and at the same time SDN finds in the state s_0 that has not been taken into consideration in ZD, so it is the state which cannot be identified by SDG, so $d_3 = (u, s_1, s_0)$.

It can be accepted that the time of SDG's work in the arbitrary mentioned state $d_i \in D_d$ ($i = 1, 2, 3$) is a random variable with: distribution $F_i(t) = P\{T < t\}$, continuous density $f_i(t)$ and positive expected value $E(T_i)$. It can be assumed that variables $T_i = (i = 1, 2, 3)$ are mutually independent [5]. In time T_1 SDG is in the state d_1 which, after ending the time T_{12} , can change into the state d_2 with probability p_{12} or after ending the time T_{13} – into the state d_3 with probability p_{13} . The state d_2 is in time T_2 and d_3 – in T_3 . SDG's state can change from d_2 into d_1 when the user notices that SDG is damaged and then makes repairs on it. This change proceeds after the end of the time T_{21} with probability p_{21} . SDG's state can change from d_3 into d_1 when the user notices identification of the state s_0 of SDN (not identified earlier by SDG) and makes repairs on it. This change proceeds after the end of the time T_{31} with probability p_{31} . In the realization time of the process $\{W(t): t \geq 0\}$ different random variables can be observed that determine the moments $\tau_0 = 0, \tau_1, \tau_2, \dots$, in which transitions of the process' states take place.

If the state of the process $\{W(t): t \geq 0\}$ is known in arbitrary moment τ_n ($n = 1, 2, \dots$), duration time of this state and the state which will occur in the moment τ_{n+1} can be considered as stochastically independent from the states which occurred previously in the moments $\tau_0, \tau_1, \tau_2, \dots, \tau_{n-1}$ and from time intervals of their duration.

Thus, it can be accepted that the process $\{W(t): t \geq 0\}$ is a semi-Markov process. Such

a process is completely defined [5, 7] when the initial distribution is known and the functional matrix is given.

From the presented considerations results that the initial distribution of the process $\{W(t): t \geq 0\}$ is as follows:

$$p_1 = P\{W(0) = d_1\} = 1, \quad p_i = P\{W(0) = d_i\} = 0 \text{ dla } i = 2, 3 \quad (4)$$

and its functional matrix is of the following form:

$$Q(t) = \begin{bmatrix} 0 & Q_{12}(t) & Q_{13}(t) \\ Q_{21}(t) & 0 & 0 \\ Q_{31}(t) & 0 & 0 \end{bmatrix}. \quad (5)$$

According to that [3]:

$$P_1 = E(T_1)M^{-1} \quad (6)$$

in which: $M = E(T_1) + p_{12}E(T_2) + p_{13}E(T_3)$

where:

p_{ij} - probability of transition of the process $\{W(t): t \geq 0\}$ from the state d_i into the state d_j ($d_i, d_j \in D; i, j = 1, 2, 3; i \neq j$).

Probability P_1 is of the following interpretation:

$$P_1 = \lim_{t \rightarrow \infty} P\{W(t) = d_1\}. \quad (7)$$

Probability P_1 can be considered as a probability of occurrence of the event A_1 , which determines the use of SDG in duration time of the state d_1 of the process $\{W(t): t \geq 0\}$, so $P_1 = P(A_1); A_1 = \{d_1\}$.

Any state of SDN (which belongs to the set of the states considered in a diagnostic task) can be identified by SDG when:

- event A_1 occurs, which defines the situation: "the state d_1 of the process $\{W(t): t \geq 0\}$ is on",
- event K occurs, which determines appearing the vector of values of diagnostic parameters,
- event K occurs in consequence of occurring the event S , which determines appearing an important (for the user) state of SDN considered in ZD, that needs to be classified to a class of model diagnostic states.

Therefore, the likelihood of diagnosis can be defined by the probability of occurring the events: A_1, S, K at the same time, basing on the following dependences [8]:

$$P(A_1 \cap S \cap K) = P(A_1)P(S|A_1)P(K|A_1 \cap S), \quad (8)$$

$$P(A_1 \cap S \cap K) = P(K)P(S|K)P(A_1|K \cap S). \quad (9)$$

From the equations: (8) and (9) follows that the probability $P(S/K)$ as the likelihood measure of diagnosis, can take the following form:

$$P(S|K) = \frac{P(A_1)P(S|A_1)P(K|A_1 \cap S)}{P(K)P(A_1|K \cap S)}. \quad (10)$$

Occurrence of the event A_1 does not influence the probability of occurring the event S , what is obvious because the events: S and A_1 are independent. That means: $P(S|A_1) = P(S)$. When SDG is reliable (if the process $\{W(t): t \geq 0\}$ is always in the state $d_1 \in D$), always occurs the event which consists in occurring the event K , providing that the event A_1 has occurred. In this situation, the dependence: $P(K/A_1 \cap S) = P(K|S)$ proceeds. Apart from that, having a reliable SDG, the event

A_1 can always be observed, assuming that the events: K and S have occurred at the same time. Therefore, it is obvious that in case of reliable SDG (such SDG for which $P(A_1) = 1$) it should be taken into account that $P(A | K \cap S) = 1$ and $P(K | A_1 \cap S) = P(K | S)$. Thus, providing that $P(A_1) = 1$, the formula (10) can be simplified to the following form:

$$P(S|K) = \frac{P(S)P(S|K)}{P(K)} \quad (11)$$

what results in obtaining the likelihood measure of diagnosis [2].

At the assumption that SDG works without any failure when diagnosing SDN's state, the likelihood measures of diagnosis can also be presented in the form of quantities of the dependences: (1), (2) and (10).

4. Diagnostic inference and likelihood of diagnosis

In the operating practice it is necessary to establish diagnoses D (Fig.1) of SDN's (e.g. combustion engine's) state as inferences which can be logically deduced from the premises, like the values of diagnostic parameters building the vector K , recorded by SDG suggesting existence (or just occurrence) of the state S of the mentioned SDN. This kind of inference (diagnostic inference) is a non-deductive inference. Thus, answers are very important to the questions: how far can one trust the inferences resulting from a non-deductive inference?, in what degree such inferences can be accepted as reliable and used for taking operating decisions?

At formulating a diagnosis (inference) of the state S of SDN, the sentence K is accepted as a completely granted assumption. The sentence says that a defined vector (in this case, it is vector K), and no other diagnostic parameters' values, has been recorded by SDG. The sentence S says that a definite state (in this case it is state S), and not other one, of SDN is the inference established basing on the sentence K , being the result of the ended non-deductive inference. In this case the inference is a reductive one [8] proceeding in the following way: if the implication $S \Rightarrow K$ is true and its direct successor (K) is true, the direct predecessor (S) of the implication is also true. This means that the presumed state of SDN is accepted as S because the vector K of values of diagnostic parameters has been recorded by SDG.

When the sentence S is an inference established basing on the sentence K (being considered as a completely granted assumption) in the process of a non-deduction inference, it can be admitted that the sentence K gives appearance of verisimilitude to the sentence S . Considering K , the measure of appearance of verisimilitude can be a logical probability [11] of the sentence S . Taking $K^*(n)$ as the following sentence: K^* is a set of n – results of tests and diagnostic inferences on the state S of the diagnosed system, and $S^*(m)$ - as the sentence: S^* is confirmed by m - results of tests and diagnostic inferences, the logical probability of the state S , taking vector K into account, can be determined as:

$$P_L(S|K) = \frac{m}{n} = h; \quad m \leq n. \quad (12)$$

One can easily noticed that the formula (12) is the same as the formula (1) which determines the frequency of a random event. From the formula (12) follows that the degree of likelihood for which the diagnosis is taken as reliable, can be higher than the value of the logical probability h of the inference that SDN is in the state S on the ground of the assumption which is the observed vector K taken for completely granted because of occurring (or existing) the state S . When $n \rightarrow \infty$, determined by the formula (1) frequency of a random event tends to the statistical probability [8], what for this case can be determined with the formula:

$$P_S(S|K) = \lim_{n \rightarrow \infty} \frac{m_S}{n}. \quad (13)$$

Determining the both probabilities: logical (P_L) and statistical (P_S), it should be taken into consideration that they concern the repeating events: S (that the state S has appeared) and K (that the vector K has appeared), which can occur only together. It means that assumption is provided that the event S appears only when K occurs.

From the considerations above results that the frequency h , determined by the formula (1), can be considered as a logical probability (P_L) when the tests and diagnostic inferences are repeated many times, and when the number of tests and inferences is great (in theory $n \rightarrow \infty$) it can be considered as a statistical probability (P_S).

Formulating a diagnosis about the state S of SDN, the sentence S (which says that SDN is in the state S) is a hypothesis and the sentence K is a result of a diagnostic test. The sentence K^* is a set of sentences K (set of results of diagnostic tests) being the tests which confirm m – times or do not confirm (forge) $n-m$ - times the hypothesis S . The hypothesis on SDN's state can be, in this case, formulated in the following way: SDN finds in the state S because the vector K of values of diagnostic parameters is being observed.

The suggested likelihood measures of diagnosis are objective. In case when the measures cannot be used because of different reasons (e.g. technical, economical, organizational) the likelihood of diagnosis can be determined with help of psychological (subjective) probability [11]. This probability determines the degree of conviction (certainty) of a diagnosis' user about the chances of coming such expectations true that SDN's state according to the formula considered in the diagnosis, is the state S . Acceptance of the diagnosis as reliable or not reliable through applying this probability is subjective because depends on the knowledge of the person who formulates (elaborate) the diagnosis. It differs from the objective probability by the fact that reflects subjective estimation of SDN's state, according to the relation: this state of SDN is more probable than any each other one or otherwise – this state of SDN is the most probable, or the most probable SDN's state is S , etc. The probability is indeed comparable but deprived of numerical measures which determine a particular degree of acceptance. The diagnosis about the state S of SDN can be considered only as more or less reliable. Therefore, the psychological (subjective) probability is not a good likelihood measure of diagnosis. From this reason, application of it should be limited in the operating practice of combustion engines on sea-going ships.

5. Summary

Occurrence of the event $A_1 = \{d_1\}$ is a necessary (but not sufficient) condition to enable identification of the state S of the given SDN. During diagnosing, when the process $\{W(t): t \geq 0\}$ is in the state d_1 , the events which consist in changing Z_D , Y and K_G can occur. Occurrence of these events makes elaboration of a reliable diagnosis difficult what causes difficulties in taking the right operating decision, as well. The less complex is SDG, the more difficult is elaboration of the diagnosis. When applying a complex SDG the all possible SDN's states are considered in ZD and then $A_1 = \{d_1, d_3\}$. Thus, in practice the likelihood of diagnosis (in case of using a local SDG) can be greater than $P(A_1) = P_1 < 1$, what follows from the formula (6).

The formula (10) which enables defining the likelihood of diagnosis, has been built by applying the limiting distribution of the semi-Markov process $\{W(t): t \geq 0\}$ and Bayes' formula.

The presented semi-Markov model of the process $\{W(t): t \geq 0\}$ is of the essential practical meaning because of easiness in defining the estimators of the probability p_{ij} and empirical distribution functions, being in approximation the distribution functions $F_{ij}(t)$, needed for calculating the elements $Q_{ij}(t)$ of the functional matrix $Q(t)$, and because of simplicity in estimating the expected values $E(T_j)$ of random variables T_j , meaning the duration time of the state $d_j \in D$ ($j = 1, 2, 3$) [10].

The presented measures of the likelihood of diagnosis are readable because for the extreme

cases the following are assigned:

- value 1, when the diagnosis is completely reliable,
- value 0, when the diagnosis is completely unreliable.

For other cases, when the diagnosis can be estimated as reliable in a certain degree, this degree is needed to be defined precisely by matching the likelihood of diagnosis with a value from the non-negative real numbers interval $R_+ = (0, 1)$.

References

- [1] Będkowski, L., *Elementy diagnostyki technicznej*, Wyd. WAT, Warszawa 1992.
- [2] Cempel, C., *Diagnostyka wibroakustyczna maszyn*, Wyd. Politechnika Poznańska, 1985.
- [3] Girtler, J., *Significance of technical diagnostics in decision control of combustion engines operation process*, *Archiwum Motoryzacji* 2004 Vol. 7, No. 4, pp. 389-407.
- [4] Girtler, J., *Statistic and probabilistic measures of diagnosis likelihood on the state of self-ignition combustion engines*, *Journal of Polish CIMAC*. Vol. 2, No. 2(2007), pp. 57-63.
- [5] Girtler, J., *A probabilistic concept of load assessment of self-ignition engines*, *Polish Maritime Research*. - Vol. 15, No. 2 (56) (2008), pp. 65-70.
- [6] Girtler, J., *Wiarygodność diagnozy a podejmowanie decyzji eksploatacyjnych*, *Kongres Diagnostyki Technicznej. KDT'96, T II*, pp. 271–276, Gdańsk 1996.
- [7] Girtler, J., *Zastosowanie bayesowskiej statystycznej teorii decyzji do sterowania procesem eksploatacji urządzeń*, *XXII Zimowa Szkoła Niezawodności*. Wyd. SPE KBM PAN – ITE, s. 55-62, Szczyrk, 1994.
- [8] Girtler, J., *Diagnostyka jako warunek sterowania eksploatacją okrętowych silników spalinowych*, Wyd. WSM. *Studia* Nr 28, Szczecin 1997.
- [9] Girtler, J., Kitowski, Z., Kuriata, A., *Bezpieczeństwo okrętu na morzu*, WKiŁ s. 61-94, Warszawa 1995.
- [10] Grabski, F., *Teoria semi-markowskich procesów eksploatacji obiektów technicznych*, *ZN WSMW (AMW)*, Nr 75A, Gdynia 1982.
- [11] Pabis, S., *Metodologia i metody nauk empirycznych*, PWN, Warszawa 1985.