

## CONTENS

### ADDRESS OF PUBLISHER & EDITOR'S OFFICE:

GDAŃSK UNIVERSITY  
OF TECHNOLOGY

Faculty of Ocean Engineering  
& Ship Technology  
G. Narutowicza 11/12  
80-233 Gdańsk, POLAND

tel.: +48 58 347 13 66  
fax: +48 58 341 13 66

### EDITORIAL STAFF:

Jacek Rudnicki  
| Editor in Chief  
Przemysław Wierzchowski  
| Scientific Editor  
Jan Michalski  
| Editor for review matters  
Aleksander Kniat  
| Editor for international  
relations  
Kazimierz Kempa  
| Technical Editor  
Piotr Bzura  
| Managing Editor

Price:  
single issue: 25 zł

Prices for abroad  
single issue:  
- in Europe EURO 15  
- overseas USD 20

### WEB:

[www.bg.pg.gda.pl/pmr/pmr.php](http://www.bg.pg.gda.pl/pmr/pmr.php)

ISSN 1233-2585

- |    |  |
|----|--|
| 3  | <b>Sławomir Świerczyński, Krzysztof Czaplewski</b><br><i>M-ESTIMATION AS A TOOL SUPPORTING A VESSEL TRAFFIC CONTROLLER<br/>IN THE VTS SYSTEM</i>   |
| 14 | <b>Stanisław Gucma, Wojciech Ślącza</b><br><i>METHODS FOR OPTIMIZATION OF SEA WATERWAY SYSTEMS<br/>AND THEIR APPLICATION</i>   |
| 20 | <b>Ming-Cheng Tsou</b><br><i>CELESTIAL NAVIGATION FIX BASED ON PARTICLE SWARM OPTIMIZATION</i>   |
| 28 | <b>Emre Peşman , Deniz Can Kolukısa , Metin Taylan</b><br><i>NUMERICAL AND ANALYTICAL APPROACHES FOR ROLL MOTION ANALYSIS<br/>IN REGULAR LONGITUDINAL WAVES</i>  |
| 36 | <b>Adam Krężel, Katarzyna Bradtke, Agnieszka Herman</b><br><i>USE OF SATELLITE DATA IN MONITORING OF HYDROPHYSICAL<br/>PARAMETERS OF THE BALTIC SEA ENVIRONMENT</i>  |
| 43 | <b>Mirosława Ostrowska, Mirosław Darecki, Adam Krężel, Dariusz<br/>Ficek, Kazimierz Furmańczyk</b><br><i>PRACTICAL APPLICABILITY AND PRELIMINARY RESULTS OF THE BALTIC<br/>ENVIRONMENTAL SATELLITE REMOTE SENSING SYSTEM (SATBAŁTYK)</i> |
| 50 | <b>Tomasz Zapadka, Adam Krężel, Marcin Paszkuta, Mirosław.<br/>Darecki</b><br><i>DAILY RADIATION BUDGET OF THE BALTIC SEA SURFACE FROM SATELLITE<br/>DATA</i>  |
| 57 | <b>Marek Jakubowski</b><br><i>INFLUENCE OF PITTING CORROSION ON FATIGUE AND CORROSION<br/>FATIGUE OF SHIP AND OFFSHORE STRUCTURES, PART II: LOAD - PIT -<br/>CRACK INTERACTION</i>   |
| 67 | <b>Przemysław Strzelecki, Janusz Sempruch, Krzysztof Nowicki</b><br><i>COMPARING GUIDELINES CONCERNING CONSTRUCTION OF THE S-N<br/>CURVE WITHIN LIMITED FATIGUE LIFE RANGE</i>   |
| 75 | <b>Kamila Kustroń</b><br><i>LASER DOPPLER VIBROMETER BASED EXAMINATION OF THE EFFICIENCY<br/>OF INTRODUCING ARTIFICIAL DELAMINATIONS INTO COMPOSITE SHELLS</i>   |
| 82 | <b>Xiao-Yan Xu, Janusz Mindykowski, Tomasz Tarasiuk, Chen Cheng</b><br><i>STUDY OF HARMONIC DETECTION METHODS UNDER NON-IDEAL<br/>CONDITIONS IN SHIP POWER NETWORK</i>   |
| 90 | <b>Jerzy Kowalski</b><br><i>AN EXPERIMENTAL STUDY OF EMISSION AND COMBUSTION<br/>CHARACTERISTICS OF MARINE DIESEL ENGINE IN CASE OF CYLINDER<br/>VALVES LEAKAGE</i>  |
| 99 | <b>Henryk Krawczyk, Michał Nykiel, Jerzy Proficz</b><br><i>TRYTON SUPERCOMPUTER CAPABILITIES FOR ANALYSIS<br/>OF MASSIVE DATA STREAMS</i>  |

## Editorial

POLISH MARITIME RESEARCH is a scientific journal of worldwide circulation. The journal appears as a quarterly four times a year. The first issue of it was published in September 1994. Its main aim is to present original, innovative scientific ideas and Research & Development achievements in the field of :

### **Engineering, Computing & Technology, Mechanical Engineering,**

which could find applications in the broad domain of maritime economy. Hence there are published papers which concern methods of the designing, manufacturing and operating processes of such technical objects and devices as : ships, port equipment, ocean engineering units, underwater vehicles and equipment as well as harbour facilities, with accounting for marine environment protection.

The Editors of POLISH MARITIME RESEARCH make also efforts to present problems dealing with education of engineers and scientific and teaching personnel. As a rule, the basic papers are supplemented by information on conferences , important scientific events as well as cooperation in carrying out international scientific research projects.

### Scientific Board

**Chairman : Prof. JERZY GIRTLER - Gdańsk University of Technology, Poland**

**Vice-chairman : Prof. ANTONI JANKOWSKI - Institute of Aeronautics, Poland**

**Vice-chairman : Prof. MIROŚLAW L. WYSZYŃSKI - University of Birmingham, United Kingdom**

<b>Dr POUL ANDERSEN</b> Technical University of Denmark Denmark	<b>Prof. WOLFGANG FRICKE</b> Technical University Hamburg- Harburg Germany	<b>Prof. YASUHIKO OHTA</b> Nagoya Institute of Technology Japan
<b>Dr MEHMET ATLAR</b> University of Newcastle United Kingdom	<b>Prof. STANISŁAW GUCMA</b> Maritime University of Szczecin Poland	<b>Dr YOSHIO SATO</b> National Traffic Safety and Environment Laboratory Japan
<b>Prof. GÖRAN BARK</b> Chalmers University of Technology Sweden	<b>Prof. ANTONI ISKRA</b> Poznań University of Technology Poland	<b>Prof. KLAUS SCHIER</b> University of Applied Sciences Germany
<b>Prof. SERGEY BARSUKOV</b> Army Institute of Odessa Ukraine	<b>Prof. JAN KICIŃSKI</b> Institute of Fluid-Flow Machinery of PASci Poland	<b>Prof. FREDERICK STERN</b> University of Iowa, IA, USA
<b>Prof. MUSTAFA BAYHAN</b> Süleyman Demirel University Turkey	<b>Prof. ZYGMUNT KITOWSKI</b> Naval University Poland	<b>Prof. JÓZEF SZALA</b> Bydgoszcz University of Technology and Agriculture Poland
<b>Prof. VINCENZO CRUPI</b> University of Messina, Italy	<b>Prof. JAN KULCZYK</b> Wrocław University of Technology Poland	<b>Prof. TADEUSZ SZELANGIEWICZ</b> Technical University of Szczecin Poland
<b>Prof. MAREK DZIDA</b> Gdańsk University of Technology Poland	<b>Prof. NICOS LADOMMATOS</b> University College London United Kingdom	<b>Prof. WITALIJ SZCZAGIN</b> State Technical University of Kaliningrad Russia
<b>Prof. ODD M. FALTINSEN</b> Norwegian University of Science and Technology Norway	<b>Prof. JÓZEF LISOWSKI</b> Gdynia Maritime University Poland	<b>Prof. BORIS TIKHOMIROV</b> State Marine University of St. Petersburg Russia
<b>Prof. PATRICK V. FARRELL</b> University of Wisconsin Madison, WI USA	<b>Prof. JERZY MATUSIAK</b> Helsinki University of Technology Finland	<b>Prof. DRACOS VASSALOS</b> University of Glasgow and Strathclyde United Kingdom
	<b>Prof. EUGEN NEGRUS</b> University of Bucharest Romania	

# M-ESTIMATION AS A TOOL SUPPORTING A VESSEL TRAFFIC CONTROLLER IN THE VTS SYSTEM

Sławomir Świerczyński<sup>1</sup>

Krzysztof Czaplewski<sup>2</sup>

1) Polish Naval Academy, Poland

2) Gdynia Maritime University, Poland

## ABSTRACT

*In order to improve maritime safety and the efficiency of vessel traffic, systems supervising vessel traffic, i.e. VTS (Vessel Traffic Service), started to be created. These systems are aimed to control vessel traffic in waters where traffic congestion, a large concentration of vessels or the presence of navigational hazards creates a risk of collision or stranding. VTS systems constitute maritime safety centres and they must be equipped with appropriate devices in order to be fully functional. Among devices that provide information about vessels are coastal radar stations which are located around a monitored sea area. This kind of spatial arrangement of these stations can be used to simultaneously obtain information about every vessel, but such observations may be fraught with serious errors. Therefore, the estimation methods that are employed and developed in geodesy can be used to improve the accuracy with which a vessel's position is determined. The Interactive Navigational Structure, i.e. IANS, is an example of how these methods can be applied in navigation; this term has already been introduced into the literature (Czaplewski, 2004). The text below presents the theoretical assumptions underlying the use of IANS as a tool supporting a vessel traffic controller using the VTS system in his/her work. This presentation is supported by a numerical test that was performed in the waters of the Bay of Gdańsk which are covered by the VTS system.*

**Keywords:** navigation, radar navigation; Vessel Traffic Services; M-estimation; Interactive Navigational Structures

## INTRODUCTION

Since the volume of vessel traffic is constantly growing, increasing emphasis should be placed on maritime safety. Marine traffic safety systems function in many sea areas; these systems are mainly aimed to improve the safety of ship transport, protect the marine environment and respond to emergencies. Sea areas are monitored by a network of coastal radar stations which constitute one of the main sources of information about vessels. Radars are to detect not only large vessels entering zones that are monitored by the VTS system, but also small ones, such as sports and fishing vessels. If the radar cross-section of particular vessels is low, radar reflectors are placed on them. International standards specify under which conditions and at what sea state radar devices should be able to detect vessels.

Based on signals that are reflected from objects, a radar image is created and it represents the situation within radar coverage. As a result, one can track vessels, detect obstructions to navigation and see an image of the shoreline, etc. both under difficult weather conditions and when visibility is poor.

The radar subsystem that monitors vessel traffic in the Bay of Gdańsk, which is referred to as the VTS system for the Bay of Gdańsk, is an example of coastal radar stations functioning within VTS systems (Bole, Wall, Norris 2014). Figure 1 presents the distribution of coastal radar stations.

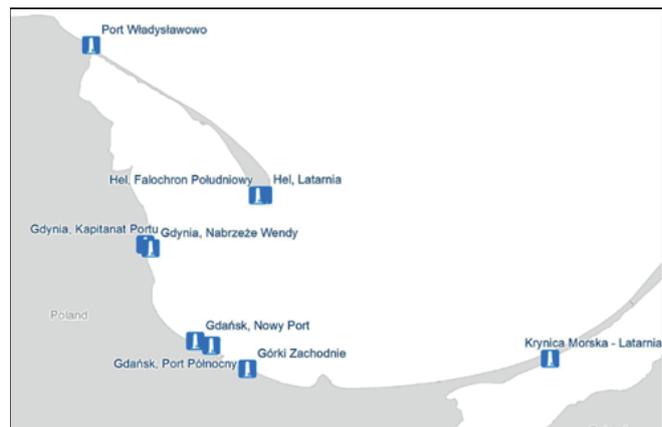


Fig. 1. Location of coastal radar stations functioning as part of the VTS system for the Bay of Gdańsk (source: Maritime Office in Gdynia)

Depending on the needs, information about a particular situation in a given sea area can be obtained from either one or several radars. The number and spatial arrangement of radars in VTS systems depends on the size and geospatial location of a given sea area. Radars are located around specific waters and they cover the whole area that is supervised by the VTS system. This is very important when return echoes from targets are

weak or there has been a failure in one of the stations. In such cases one receives a large number of navigational data which can be used to improve the parameters (position) of a vessel by employing modern methods of adjustment calculus that are used in geodesy. If navigational information is received from many independent radar stations, this allows one to more accurately determine the positions of vessels in systems supervising vessel traffic (Monahan, 2008).

If a vessel's position is determined accurately, this also means that maritime safety is improved, which is particularly important in sea areas with a large volume of traffic where the risk of collision is relatively high. Since vessels move when sailing it is impossible to repeat measurements at the same position. Therefore, both classical estimation methods and modern

M-estimation methods that are used in geodesy can be utilised to improve the accuracy with which vessels' positions are determined (Czaplewski, 2004; Wiśniewski, 2009).

As for robust estimation, a key issue is to select the appropriate attenuation function. Several

M-estimation methods are used in adjustment calculus; these methods utilise different attenuation functions or weight functions. Among the most well-known methods of this kind are: Huber's method, Hampel's method and the Danish attenuation function (Borre, Jorgensen, Kubik, 1983; Hampel, Ronchetti, Rousseeuw, Stahel, 2005). They were applied in the process of determining the position fix of a vessel by using radar navigation methods.

## OBTAINING NAVIGATIONAL INFORMATION FROM RADAR OBSERVATIONS

### USING THE DISTANCE

Observations from coastal radar stations, where a vessel's position at sea is determined based on the measurement of distance, can be utilised to measure navigational parameters which are used to establish a vessel's position (Bole, Wall, Norris 2014). The measurement structure will depend on the number of radars that were used to make measurements. Figure 2 presents an example radar network. One can measure the distance to vessels from five coastal radar stations that are distributed around an area covered by the VTS system.

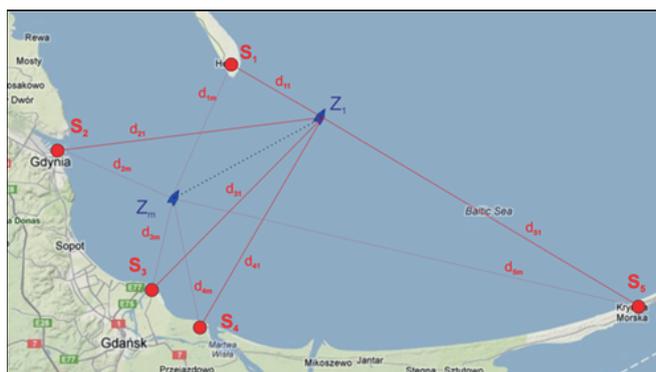


Fig. 2. Distance measurement structure (source: own work)

The distances that are measured from coastal stations with known geographic coordinates will help in determining a vessel's estimated position. This method, which is used based on observations made by coastal radars, is to allow one to correct the coordinates of a vessel's position that have been determined by an officer in charge of a navigational watch.

In this case, the adjustment problem entails determining the position fix of a vessel at sea (in a sea area covered by a VTS system) based on radar observations and the vessel's dead reckoning position that has been established by the officer in charge of a navigational watch on the ship.

The distances from coastal radar stations that have been measured allow one to define the system of observation equations for five coastal stations. The expected values can be determined by using the following formula (Wiśniewski, 2009):

$$d_i^o = \sqrt{(X_{S_i} - X_{Z_j})^2 + (Y_{S_i} - Y_{Z_j})^2} \Big|_{\substack{i=1,\dots,5 \\ j=1,\dots,m}} \quad (1)$$

where:

$(X_{S_i}, Y_{S_i})$  - coordinates of radar stations' positions,

$(X_{Z_j}, Y_{Z_j})$  - coordinates of the vessel's positions.

The system of linear adjustment equations takes the following form (Wiśniewski, 2009):

$$v_{ij} = \frac{\partial d_{ij}}{\partial X_{Z_j}} \hat{d}_{X_{Z_j}} + \frac{\partial d_{ij}}{\partial Y_{Z_j}} \hat{d}_{Y_{Z_j}} + d_{ij}^o - d_{ij}^{obs} \Big|_{\substack{i=1,\dots,5 \\ j=1,\dots,m}} \quad (2)$$

where:

$v_{ij}$  - adjustment of the bearing measured from the  $i$ -th radar station ( $i = 1, \dots, 5$ ) to the  $j$ -th position of the vessel,

$d_{ij}^o$  - the expected value of the distance from a radar to the vessel for  $Z_j^o(X_{Z_j}^o, Y_{Z_j}^o)$ ,

$d_{ij}^{obs}$  - the observed value for  $Z_j^o(X_{Z_j}^o, Y_{Z_j}^o)$ ,

$Z_j$  -  $j$ -th position of the vessel ( $j = 1, \dots, m$ ).

### USING THE BEARINGS

In order to determine a vessel's position at sea one can also use other observations, for example, bearings from coastal radar stations (Bole, Wall, Norris 2014). Figure 3 presents an example measurement structure for the case that is discussed here. For each of the measurements made by several coastal radar stations we obtain bearings having true value  $NR_{ij}(i = 1, \dots, 5, j=1, \dots, m)$  for the  $i$ -th observation made relative to the vessel's  $j$ -th position. The results of observations (for example,  $NR_{ijobs}$ ) having true values are fraught with measurement errors, and it is often assumed that a measurement error is a random variable which has certain assumed descriptive parameters. Since there are errors, especially random ones, it may be difficult to determine the value of a measurement. The value of a measurement is actually a theoretical argument and it can only be estimated by using certain legitimate methods. The methods of estimating the descriptive parameters of random variables that have been developed in mathematical statistics provide a theoretical basis for the methods of adjusting the results of measurements (Wiśniewski, 2004; Jianjun, 1996).

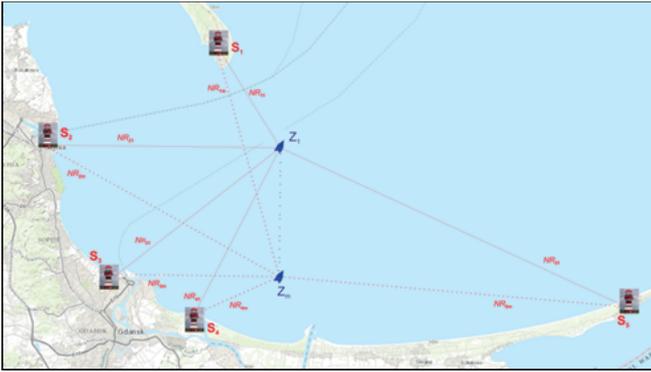


Fig. 3. Bearing measurement structure (source: own work)

Let us assume that we can use bearings that have been measured from coastal radar stations to a vessel at position  $Z_j$ . Then, the system of observation equations is as follows (Wiśniewski, 2009):

$$NR_{ij}^Z = \arctg \frac{Y_{S_i} - Y_{Z_j}}{X_{S_i} - X_{Z_j}} \Bigg|_{\substack{i=1,\dots,5 \\ j=1,\dots,m}} \quad (3)$$

where:

$(X_{S_i}, Y_{S_i})$  – coordinates of radar stations' positions,  
 $(X_{Z_j}, Y_{Z_j})$  – coordinates of the vessel's positions.

For the measurement structure which is used for the purpose of this discussion and which is presented in Figure 3, the functional model of the adjustment problem takes the following form:

$$v_{ij} = \frac{\partial NR_{ij}}{\partial X_{Z_j}} \hat{d}_{X_{Z_j}} + \frac{\partial NR_{ij}}{\partial Y_{Z_j}} \hat{d}_{Y_{Z_j}} + NR_{ij}^0 - NR_{ij}^{obs} \Bigg|_{\substack{i=1,\dots,5 \\ j=1,\dots,m}} \quad (4)$$

where:

$NR_{ij}^0$  – the expected value of a radar bearing for  $Z_j^0(X_{Z_j}^0, Y_{Z_j}^0)$  which was taken from the  $i$ -th radar station ( $i = 1, \dots, 5$ ) to the  $j$ -th position of the vessel,

$v_{ij}$  – adjustment to the observation that has been made.

In the adjustment process it is assumed that the position fix, which has been provided by an officer in charge of a navigational watch to a vessel traffic controller using the VTS system, is the expected position with the coordinates:

$$\mathbf{Z}_j^0 = \begin{bmatrix} X_{Z_j}^0 \\ Y_{Z_j}^0 \end{bmatrix} \quad (5)$$

Given that:

$$\mathbf{V} = \begin{bmatrix} v_{11} \\ \vdots \\ v_{ij} \end{bmatrix}_{\substack{i=1,\dots,5 \\ j=1,\dots,m}} \quad - \text{vector of adjustments,}$$

$$\hat{\mathbf{d}}_X = \begin{bmatrix} \hat{d}_{X_{Z_j}} \\ \hat{d}_{Y_{Z_j}} \end{bmatrix} \quad - \text{the unknown vector of increments to the expected coordinates.}$$

The matrix of coefficients of the unknowns and the matrix of constants assume different values, depending on the observations that are used. When it is distances that are used then we will obtain:

$$\mathbf{A} = \begin{bmatrix} \frac{\partial d_{11}}{\partial X_{Z_j}} & \frac{\partial d_{11}}{\partial Y_{Z_j}} \\ \vdots & \vdots \\ \frac{\partial d_{ij}}{\partial X_{Z_j}} & \frac{\partial d_{ij}}{\partial Y_{Z_j}} \end{bmatrix} \quad - \text{matrix of coefficients of the unknowns,}$$

$$\mathbf{L} = \begin{bmatrix} d_{11}^0 - d_{11}^{obs} \\ \vdots \\ d_{ij}^0 - d_{ij}^{obs} \end{bmatrix}_{\substack{i=1,\dots,5 \\ j=1,\dots,m}} \quad - \text{matrix of constants.}$$

However, if it is bearings that are used then:

$$\mathbf{A} = \begin{bmatrix} \frac{\partial NR_{11}}{\partial X_{Z_j}} & \frac{\partial NR_{11}}{\partial Y_{Z_j}} \\ \vdots & \vdots \\ \frac{\partial NR_{ij}}{\partial X_{Z_j}} & \frac{\partial NR_{ij}}{\partial Y_{Z_j}} \end{bmatrix}_{\substack{i=1,\dots,5 \\ j=1,\dots,m}} \quad - \text{matrix of coefficients of the unknowns,}$$

$$\mathbf{L} = \begin{bmatrix} NR_{11}^0 - NR_{11}^{obs} \\ \vdots \\ NR_{ij}^0 - NR_{ij}^{obs} \end{bmatrix}_{\substack{i=1,\dots,5 \\ j=1,\dots,m}} \quad - \text{matrix of constants.}$$

If we adopt the matrix structures that are presented above, i.e.  $\mathbf{A}$ ,  $\hat{\mathbf{d}}_X$ ,  $\mathbf{L}$ , the matrix system of adjustment equations takes the following form:

$$\mathbf{V} = \mathbf{A} \cdot \hat{\mathbf{d}}_X + \mathbf{L} \quad (6)$$

If the indefinite method is used, the solution to the system of equations will be as follows:

$$\hat{\mathbf{d}}_X = -(\mathbf{A}^T \mathbf{P} \mathbf{A})^{-1} \mathbf{A}^T \mathbf{P} \mathbf{L} \quad (7)$$

where:

$$\mathbf{P} = \begin{bmatrix} \frac{1}{m_{11}^2} & & \\ & \ddots & \\ & & \frac{1}{m_{ij}^2} \end{bmatrix}_{\substack{i=1,\dots,5 \\ j=1,\dots,m}} \quad - \text{is a weight matrix of the observations made.}$$

Where:

$m_{ij}$  – average error of the  $i$ -th observation made relative to the vessel's  $j$ -th position.

Therefore, the estimators of the adjusted coordinates of a vessel at sea are as follows:

$$\hat{\mathbf{Z}}_j = \mathbf{Z}_j^0 + \hat{\mathbf{d}}_{XY} = \begin{bmatrix} X_{Z_j}^0 \\ Y_{Z_j}^0 \end{bmatrix} + \begin{bmatrix} \hat{d}_{X_{Z_j}} \\ \hat{d}_{Y_{Z_j}} \end{bmatrix} = \begin{bmatrix} \hat{X}_{Z_j} \\ \hat{Y}_{Z_j} \end{bmatrix} \quad (8)$$

Where:

$\begin{bmatrix} \hat{X}_{Z_j} \\ \hat{Y}_{Z_j} \end{bmatrix}$  – estimated coordinates of a vessel at sea in the coordinate system X,Y (PL-UTM).

In order to determine the quality of the observations that have been made, a statistical model of the adjustment problem should be defined:

$$\hat{\mathbf{C}}_{\hat{\mathbf{x}}} = m_0^2 (\mathbf{A}^T \mathbf{P} \mathbf{A})^{-1} = \begin{bmatrix} m_{\hat{X}_{Z_j}}^2 & \hat{\text{cov}}(\hat{X}_j, \hat{Y}_j) \\ \hat{\text{cov}}(\hat{X}_j, \hat{Y}_j) & m_{\hat{Y}_{Z_j}}^2 \end{bmatrix} \quad (9)$$

where:

$$m_0^2 = \frac{V^T P V}{n - r} \quad \text{– estimator of the coefficient of variance,}$$

$n$  – number of observations made,

$r$  – number of the unknowns.

Then the average error of the position fix will be determined based on the formula:

$$M_{po} = \sqrt{m_{\hat{X}_{Z_j}}^2 + m_{\hat{Y}_{Z_j}}^2} \quad (10)$$

If we assume that the goal function for the present research problem is compatible with the goal function in the method of least squares, the adjustment problem can be finally written as follows:

$$\begin{cases} \mathbf{V} = \mathbf{A} \cdot \hat{\mathbf{d}}_X + \mathbf{L} \\ \hat{\mathbf{C}}_{\hat{\mathbf{x}}} = m_0^2 (\mathbf{A}^T \mathbf{P} \mathbf{A})^{-1} \\ \psi(\mathbf{V}^T \mathbf{P} \mathbf{V}) = \min \end{cases} \quad (11)$$

## SELECTED ATTENUATION FUNCTIONS

In marine navigation, the measurement structure can reflect the work of real devices that measure directions; then an angle is calculated between the direction vectors. Among such devices are the above-mentioned coastal radar stations, which are used to collect navigational data (for example, a bearing) that are needed to determine the positions of vessels in sea areas. Figure 3 presents an example structure of this kind.

The results of measurements can sometimes be fraught with serious errors. Such errors may occur, for example, because measuring instruments have been poorly fitted or

calibrated or the indications have been read incorrectly, or due to temporary changes in the parameters of the measurement environment, etc. Errors of this kind are referred to as gross errors. In order to eliminate them, one can use robust estimation methods with properly selected attenuation functions.

Many M-estimation methods are used in adjustment calculus; these methods utilise different attenuation functions or different weight functions. Among the most well-known methods of this kind are: the Danish, Huber's, and Hampel's methods. For the purpose of this study, the three functions were analysed with regard to their influence on the accuracy with which positions are determined for measurements that are fraught with gross errors; these errors are described generally in the text above based on a paper written by Borre, Jorgensen, Kubik (1983), Teunissen P.J.G. (2004) and Hampel, Ronchetti, Rousseeuw, Stahel, (2005).

The characteristics of the Danish attenuation function are that it decreases exponentially outside acceptable range  $\Delta \bar{v} \in \langle -k; k \rangle$  and takes the following form:

$$t(\bar{v}) = \begin{cases} 1 & \text{for } \bar{v} \in \langle -k; k \rangle \\ \exp[-l(|\bar{v}| - k)^g] & \text{for } |\bar{v}| > k \end{cases} \quad (12)$$

Therefore, equivalent weight values will be determined according to this formula:

$$\hat{p} = t(\bar{v})p = \begin{cases} p & \text{for } \bar{v} \in \langle -k; k \rangle \\ \exp[-l(|\bar{v}| - k)^g] p & \text{for } |\bar{v}| > k \end{cases} \quad (13)$$

The form of the Danish attenuation function is presented in Figure 4.

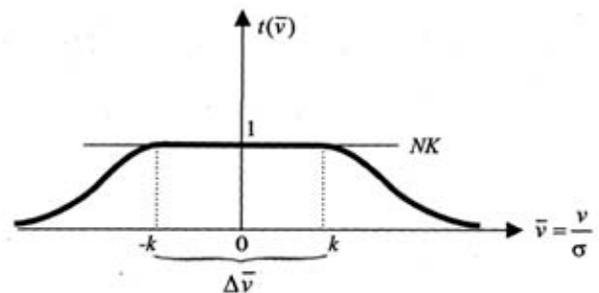


Fig. 4. Danish attenuation function (source: Wiśniewski, 2009)

Hampel's attenuation function has two additional intermediate ranges (to the left and to the right of acceptable range  $\Delta \bar{v} \in \langle -k; k \rangle$ ) in which attenuation function  $t(\bar{v})$  decreases linearly (Hampel, Ronchetti, Rousseeuw, Stahel, 2005). This function and the resulting weight function take the following forms:

$$t(\bar{v}) = \begin{cases} 1 & \text{for } \bar{v} \in \langle -k; k \rangle \\ \frac{|\bar{v}| - k_b}{k - k_b} & \text{for } |\bar{v}| \in \langle k; k_b \rangle \\ 0 & \text{for } |\bar{v}| > k_b \end{cases} \quad (14)$$

$$\hat{p} = t(\bar{v})p = \begin{cases} p & \text{for } \bar{v} \in \langle -k; k \rangle \\ \left( \frac{|\bar{v}| - k_b}{k - k_b} \right) p & \text{for } |\bar{v}| \in (k; k_b) \\ 0 & \text{for } |\bar{v}| > k_b \end{cases} \quad (15)$$

where  $k_b$  is a number that determines the limits of the additional ranges. It is usually assumed that  $k_b = 4, \dots, 6$ . The form of Hampel's attenuation function is presented in Figure 5.

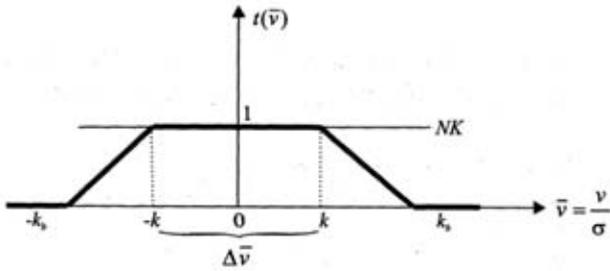


Fig. 5. Hampel's attenuation function (source: Wiśniewski, 2009)

Huber's attenuation function was the last of the analysed functions; this function assigns zero weight values to all observations with adjustments that do not belong to the acceptable range and it takes this form:

$$t(\bar{v}) = \begin{cases} 1 & \text{for } \bar{v} \in \langle -k; k \rangle \\ 0 & \text{for } \bar{v} \notin \langle -k; k \rangle \end{cases} \quad (16)$$

Thus, the weight function is as follows:

$$\hat{p} = t(\bar{v})p = \begin{cases} p & \text{for } \bar{v} \in \langle -k; k \rangle \\ 0 & \text{for } \bar{v} \notin \langle -k; k \rangle \end{cases} \quad (17)$$

The form of Huber's attenuation function is presented in Figure 6.

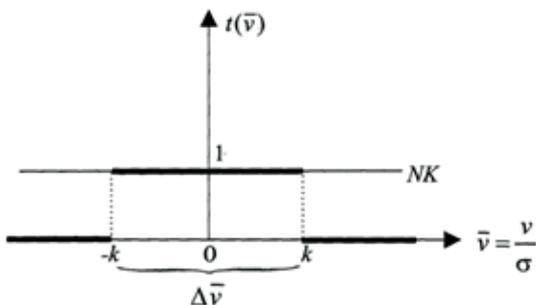


Fig. 6. Huber's attenuation function (source: Wiśniewski, 2009)

Finally, the Danish attenuation function was selected because its form was most suited to dynamic models that perform typical navigation tasks.

If one uses a method that is robust to gross errors then the adjustment problem can be expressed in the following form:

$$\left. \begin{aligned} \mathbf{V} &= \mathbf{A}\hat{\mathbf{d}}_X + \mathbf{L} && \text{functional model} \\ \mathbf{C}_X &= \sigma_0^2 \mathbf{Q}_X = \sigma_0^2 \mathbf{P}^{-1} && \text{original statistical model} \\ \hat{\mathbf{C}}_X &= \sigma_0^2 \hat{\mathbf{Q}}_X = \sigma_0^2 \hat{\mathbf{P}}^{-1} && \text{equivalent statistical model} \\ \hat{\mathbf{P}} &= \mathbf{T}(\bar{\mathbf{V}})\mathbf{P} && \text{equivalent weights} \\ \Psi(\hat{\mathbf{d}}_X) &= \mathbf{V}^T \hat{\mathbf{P}}\mathbf{V} = \mathbf{V}^T \mathbf{T}(\bar{\mathbf{V}})\mathbf{P}\mathbf{V} = \min && \text{adjustment criterion} \\ \mathbf{V}^T \mathbf{T}(\bar{\mathbf{V}})\mathbf{P}\mathbf{V} &= \hat{\mathbf{V}}^T \mathbf{T}(\hat{\bar{\mathbf{V}}})\mathbf{P}\hat{\mathbf{V}} && \text{adjustment criterion} \end{aligned} \right\} \quad (18)$$

Where:

$\hat{\mathbf{P}} = \mathbf{T}(\bar{\mathbf{V}})\mathbf{P}$  – equivalent weight matrix,

$\hat{\mathbf{C}}_X$  – equivalent covariance matrix,

$\hat{\mathbf{Q}}_X$  – equivalent cofactor matrix,

$\mathbf{T}(\bar{\mathbf{V}})$  – diagonal attenuation matrix.

For such assumptions, the solution to the adjustment problem is iterative, i.e. the algorithm that was proposed in papers written by Wiśniewski (2004, 2009) can be used to solve this problem – at the first stage of calculations adjustment is made by using the classical method of least squares. An M-estimation method with the Danish attenuation function is presented in this paper.

## M-ESTIMATION WITH THE DANISH ATTENUATION FUNCTION

As for the measurements of navigational parameters, the obtained results can be different than expected. If such errors, which are called gross errors, occur it is recommended that the measurement and results be corrected. It is, however, impossible to repeat a measurement for a moving vessel. At each time step, a vessel covers a certain distance and it “does not come back” to the point at which the last measurement was made. Serious errors can lead to using nautical charts for navigation incorrectly, and consequently pose a threat to maritime safety. It is possible to limit the influence of such errors on how a course line is drawn on a nautical chart by using modern M-estimation methods.

In a classical adjustment which is made by using the method of least squares, the results of measurements are random variables with the same standard deviation and they are treated equally; therefore, gross errors may significantly affect the measurement results. This method is not robust to serious errors. Thus, one can use robust estimation methods with properly selected attenuation functions in order to eliminate such errors. Later in the study the Danish attenuation function is used in the M-estimation procedure.

For this method, one can use the measurement structure which is shown in Figure 3 and which will be defined in the coordinate system (X, Y), where a bearing represents an observation. Radar observations can be made from coastal stations with regard to a vessel at position for a given measurement network. Bearings were measured at various time intervals.

If we assume that a given observation that has been made is fraught with gross error, for example, because the radar echo of a vessel was identified incorrectly, then a so-called

equivalent weight  $\hat{\mathbf{P}}$  will be assigned to this observation; this weight will result from attenuating the original weight  $\mathbf{P}$  (resulting from the assumed average measurement error). The attenuation process will occur in accordance with the following formula (Wiśniewski, 2009):

$$\hat{\mathbf{P}} = \mathbf{T}(\bar{\mathbf{V}})\mathbf{P} \quad (19)$$

where:  $\mathbf{T}(\bar{\mathbf{V}})$  is an attenuation function having the following basic properties:

for  $v_i \in \Delta v_i, v_j \in \Delta v_j$ :  $t(v_i) = t(v_j)$

for  $(v_i \notin \Delta v_i) \vee (v_j \notin \Delta v_j), |v_i| < |v_j|$ :  $t(v_i) > t(v_j)$

Ranges  $\Delta v_i = \langle -k\sigma_{v_i}; k\sigma_{v_i} \rangle$  are acceptable ranges for adjustments  $v_i, i = 1, \dots, n$ , and they are established with the assumed confidence level  $\gamma$ . If we assume that  $v_i$  are random variables with a normal distribution, we can write the following:

$$\gamma = P(-k\sigma_{v_i} < v_i < k\sigma_{v_i}) = P(-k < \bar{v}_i < k) = \int_{-k}^{+k} \frac{1}{\sqrt{2\pi}} \exp\left[-\frac{\bar{v}_i^2}{2}\right] d\bar{v}_i \quad (20)$$

where:

$\sigma_{v_i} = \sqrt{[\mathbf{C}_V]_{ii}}$  – standard deviation for the  $i$ -th adjustment,

$\bar{v}_i = \frac{v_i}{\sigma_{v_i}}$  – standardised adjustments with the same acceptable range  $\Delta \bar{v}_i = \langle -k; k \rangle$ . For example, for  $\gamma = 0.95$  coefficient  $k = 2$ .

Equivalent weight matrix  $\hat{\mathbf{P}} = \mathbf{T}(\bar{\mathbf{V}})\mathbf{P}$  depends on the vector of standardised adjustments  $\bar{\mathbf{V}}$ . If a method that is robust to gross errors is used, the adjustment problem can be expressed as formula (18) (Wiśniewski, 2009; Yang, Song, Xu, 2002).

For such assumptions, the solution to the adjustment problem, which involves determining such  $\hat{\mathbf{d}}_X$  that  $\xi(\hat{\mathbf{d}}_X) = \mathbf{v}^T \hat{\mathbf{P}} \mathbf{v} = \min$ , is iterative. In the first steps in the iterative process that is aimed to solve the adjustment problem, it is assumed that  $l = 0,01 \div 0,1, g = 2$ . If parameters are chosen incorrectly, this unnecessarily increases the number of steps in the iterative process which is aimed to solve a robust adjustment problem. A well-known algorithm was used to solve this problem; at the first stage bearings were adopted as measurement observations and adjustment was made by using the classical method of least squares (3) ÷ (10). Further calculations allow us to establish which of the standardised adjustments can represent gross errors (they do not belong to  $\Delta \bar{\mathbf{V}}$ ) and determine a covariance matrix of the vector of adjustments for  $m_0 = 1$ .

$$\hat{\mathbf{C}}_{\hat{\mathbf{V}}(m_0=1)} = \mathbf{P}^{-1} - \mathbf{A}(\mathbf{A}^T \mathbf{P} \mathbf{A})^{-1} \mathbf{A}^T \quad (21)$$

Then we carry out classification in order to find out if for each:

$$i = \bar{v}_i \in \Delta \bar{\mathbf{V}}$$

If all  $\bar{v}_i \in \Delta \bar{\mathbf{V}}$  then we do not perform further calculations because the values of standardised adjustments that have

been obtained fall within the acceptable range, and a new attenuation matrix will not reduce the values of weight matrices and then also the values of adjustments (within the precision range that has been assumed for the calculations).

Otherwise, if any adjustment  $\bar{v}_i \notin \Delta \bar{\mathbf{V}}$ , the next steps in the iterative process are performed (Wiśniewski, 2009):

We assume that  $j = 0$  and then

$\mathbf{V}^{(j)} = \mathbf{V}, \mathbf{P}^{(j)} = \mathbf{P}, \mathbf{C}_{\mathbf{V}}^{(j)}(m_0=1) = \hat{\mathbf{C}}_{\hat{\mathbf{V}}(m_0=1)}^{(j)}$  and the control parameters in the Danish attenuation function  $l = 0.01, g = 2$ , and the following is calculated:

by remembering that:

$$t(\bar{v}) = \begin{cases} 1 & \text{for } \bar{v} \in \langle -k; k \rangle \\ \exp[-l(|\bar{v}| - k)^g] & \text{for } |\bar{v}| > k \end{cases} \quad (22)$$

we calculate the values of the attenuation function:

$$\begin{aligned} \bar{v}_i^{(j)} \in \Delta \bar{\mathbf{V}} &\rightarrow t(\bar{v}_i^{(j)}) = 1 \\ \bar{v}_i^{(j)} \notin \Delta \bar{\mathbf{V}} &\rightarrow t(\bar{v}_i^{(j)}) = \exp[-l(|\bar{v}_i^{(j)}| - k)^g] \end{aligned} \quad (23)$$

and the attenuation matrix

$$\mathbf{T}(\bar{\mathbf{V}}^{(j)}) = \begin{bmatrix} t(\bar{v}_1^{(j)}) & & \\ & \ddots & \\ & & t(\bar{v}_n^{(j)}) \end{bmatrix}_{\text{dla } i=1, \dots, n} \quad (24)$$

Then we perform iteration and increase  $j$  by 1, i.e.  $j = j + 1$ , and we calculate weight matrices as well as matrix increases and adjustments:

$$\left. \begin{aligned} \mathbf{P}^{(j)} &= \mathbf{T}(\bar{\mathbf{V}}^{(j-1)})\mathbf{P}^{(j-1)} \\ \mathbf{d}_X^{(j)} &= -(\mathbf{A}^T \mathbf{P}^{(j)} \mathbf{A})^{-1} \mathbf{A}^T \mathbf{P}^{(j)} \mathbf{L} \\ \mathbf{V}^{(j)} &= \mathbf{A} \cdot \mathbf{d}_X^{(j)} + \mathbf{L} \end{aligned} \right\} \quad (25)$$

For the precision range that has been assumed for the calculations, the differences between elements of vectors of adjustments are checked:

$$\mathbf{V}^{(j)} \quad \text{ i } \quad \mathbf{V}^{(j-1)}$$

If these differences are larger than assumed, we calculate a covariance matrix of the vector of adjustments for  $m_0 = 1$

$$\begin{aligned} \mathbf{C}_{\mathbf{V}}^{(j)}(m_0=1) &= (\mathbf{P}^{(j)})^{-1} - \mathbf{A}(\mathbf{A}^T \mathbf{P}^{(j)} \mathbf{A})^{-1} \mathbf{A}^T \\ m_{v_i}^{(j)} &= \sqrt{[\mathbf{C}_{\mathbf{V}}^{(j)}(m_0=1)]_{ii}} \quad \bar{v}_i^{(j)} = \frac{v_i^{(j)}}{m_{v_i}^{(j)}} \end{aligned} \quad (26)$$

Then we calculate the value of the attenuation function and attenuation matrix. We increase  $j$  by 1, i.e.  $j = j + 1$ , and proceed to the next iteration step. The iteration process ends with an adjustment whereby the obtained values of standardised adjustments fall within the acceptable range, and the resulting new attenuation matrix will not reduce the values of weight matrices and then also the values of adjustments (within the precision range that has been assumed for the calculations). The final weight matrix is an equivalent matrix, and the solution that was arrived at based on this matrix is the final solution. In the equivalent weight matrix, weights that correspond to observations which are fraught with gross errors are no longer original weights (with values resulting from average measurement errors), but weights with reduced values, which are sometimes equal to zero.

### NUMERICAL TEST

For a measurement network consisting of five coastal radar stations observations were made relative to a vessel floating in the Bay of Gdańsk. Bearings were the navigational observations that were made. Since the measuring series was long, the paper presents measurements for ten positions of the vessel. Table 1 presents the values of these observations. For the purposes of this study, one of the measurements is “fraught” with gross error, which in practice may occur if a radar operator has taken an incorrect bearing or if the targeted

object has been incorrectly identified. For all positions, a gross error has occurred in the bearing taken from the radar station in the port on the Hel Peninsula.

Table 2 presents the coordinates of the coastal radar stations from which bearings were taken.

Table 2. Coordinates of the coastal radar stations (Source: own work).

Item no.	Coastal station	Rectangular coordinates (X,Y)
1	Hel Lighthouse	X = 6052476.63 Y = 357945.55
2	Gdynia_KP Harbour Master	X = 6045669.81 Y = 341309.47
3	Gdynia_S Harbour Master	X = 6045119.44 Y = 342083.22
4	Gdańsk North Port Harbour Master	X = 6031298.79 Y = 348189.74
5	Radar Tower Górk Zachodnie	X = 6027017.31 Y = 355714.79

Due to the large volume of research material, this paper only presents the results of the calculations for position  $Z_2$ . The results of all the adjustments that have been made are shown in Figure 7. In order to simplify the computational process, the authors decided that further calculations would

Table 1. Bearings, distances and coordinates for a vessel floating in the Bay of Gdańsk. (Source: own work).

Measurement point	The coastal radar stations					$Z_j^0$
	Hel Lighthouse	Gdynia_KP Harbour Master	Gdynia_S Harbour Master	Gdańsk North Port Harbour Master	Radar Tower Górk Zachodnie	
$Z_1$	213.9°	114.5°	113.1°	0.8°	334.4°	X = 6042478.33 Y = 348331.82
$Z_2$	213.7°	114.6°	113.3°	0.7°	334.3°	X = 6042455.38 Y = 348340.88
$Z_3$	213.7°	114.7°	113.3°	0.8°	334.3°	X = 6042436.72 Y = 348349.55
$Z_4$	213.6°	114.8°	113.4°	0.8°	334.3°	X = 6042423.17 Y = 348355.15
$Z_5$	213.5°	114.9°	113.6°	0.9°	334.3°	X = 6042389.72 Y = 348367.64
$Z_6$	213.3°	115.2°	113.9°	1.0°	334.3°	X = 6042336.15 Y = 348387.79
$Z_7$	213.2°	115.4°	114.2°	1.0°	334.3°	X = 6042313.62 Y = 348395.46
$Z_8$	212.8°	116.0°	114.9°	1.3°	334.3°	X = 6042194.18 Y = 348437.39
$Z_9$	212.3°	116.4°	115.3°	1.4°	334.2°	X = 6042121.38 Y = 348460.89
$Z_{10}$	211.8°	117.3°	116.3°	1.8°	334.2°	X = 6041954.26 Y = 348519.17

be performed in a rectangular coordinate system and not in a geographic coordinate system.

The coordinates of a vessel's position that were determined by an officer in charge of a navigational watch are treated as the expected coordinates of a vessel's position at sea. For this observational arrangement, a system of adjustment equations is assumed, like in formula (4); after the above-mentioned data are substituted for the variables in this formula, the following elements of the matrix system of adjustment equations are obtained – formula (6).

When the observations that have been made are adjusted as part of the first step which is aimed at identifying gross errors by using the classical method, the following results are obtained:

The matrix of coefficients of the unknowns:

$$\mathbf{A} = \begin{bmatrix} 0.00286 & -0.00298 \\ -0.00674 & -0.00308 \\ -0.00775 & -0.00330 \\ -0.00007 & 0.00513 \\ 0.00144 & 0.00302 \end{bmatrix}$$

The matrix of constants:

$$\mathbf{L} = \begin{bmatrix} 223.78 - 213.75 \\ 114.57 - 114.66 \\ 113.06 - 113.29 \\ 0.77 - 0.73 \\ 334.47 - 334.33 \end{bmatrix} = \begin{bmatrix} 10.03 \\ -0.09 \\ -0.23 \\ 0.04 \\ 0.014 \end{bmatrix}$$

If we assume that the average observation error is  $m_j = 0.5^\circ$  for  $i = 1, \dots, 5$  and that the observations that have been made are independent of one another, the weight matrix of the observations takes the following form:

$$\mathbf{P} = \begin{bmatrix} 4 & 0 & 0 & 0 & 0 \\ 0 & 4 & 0 & 0 & 0 \\ 0 & 0 & 4 & 0 & 0 \\ 0 & 0 & 0 & 4 & 0 \\ 0 & 0 & 0 & 0 & 4 \end{bmatrix}$$

The vector of increments to the expected coordinates is:

$$\hat{\mathbf{d}}_X = -(\mathbf{A}^T \mathbf{P} \mathbf{A})^{-1} \mathbf{A}^T \mathbf{P} \mathbf{L} = \begin{bmatrix} \hat{d}_{X_{z_2}} \\ \hat{d}_{Y_{z_2}} \end{bmatrix} = \begin{bmatrix} -557.78 \\ 796.05 \end{bmatrix}_{[m]}$$

Therefore, the vector of adjustments will take the value:

$$\mathbf{V} = \begin{bmatrix} v_1 \\ v_2 \\ v_3 \\ v_4 \\ v_5 \end{bmatrix} = \begin{bmatrix} 6.07 \\ 1.21 \\ 1.47 \\ 4.17 \\ 1.74 \end{bmatrix}$$

Thus, the estimator of the position fix of a vessel at sea, which has been determined by using observations made from coastal

radar stations, is:

$$\hat{\mathbf{Z}}_1 = \begin{bmatrix} \hat{X}_{z_1} \\ \hat{Y}_{z_1} \end{bmatrix} = \begin{bmatrix} X_{z_1}^0 \\ Y_{z_1}^0 \end{bmatrix} + \begin{bmatrix} \hat{d}_{X_{z_1}} \\ \hat{d}_{Y_{z_1}} \end{bmatrix} = \begin{bmatrix} 6042455.38 \\ 348340.88 \end{bmatrix} + \begin{bmatrix} -557.78 \\ 796.05 \end{bmatrix} = \begin{bmatrix} 66041897.60 \\ 349136.93 \end{bmatrix}$$

The estimators for the observations are, respectively:

$$\begin{bmatrix} \hat{NR}_1 \\ \hat{NR}_2 \\ \hat{NR}_3 \\ \hat{NR}_4 \\ \hat{NR}_5 \end{bmatrix} = \begin{bmatrix} 219.8 \\ 115.9 \\ 114.8 \\ 4.9 \\ 336.1 \end{bmatrix}_{[^\circ]}$$

The average error of the estimated coordinates of the vessel is, respectively:

$$m_{\hat{X}_{z_2}} = \sqrt{1988} = 478.1 \text{ [m]}$$

$$m_{\hat{Y}_{z_2}} = \sqrt{4554} = 639.3 \text{ [m]}$$

Finally, the average error of the position is:

$$M_{po} = \sqrt{m_{\hat{X}_{z_2}}^2 + m_{\hat{Y}_{z_2}}^2} = 798.3 \text{ [m]}$$

Based on the obtained results, we must determine which of the standardised adjustments can represent gross errors. Let us assume, for the purpose of the calculations, that  $\gamma = 0.95$ , and thus  $k = 2$ . The acceptable range  $\Delta \bar{v}$  is as follows:  $\Delta \bar{v} \in \langle -k; k \rangle = \langle -2; 2 \rangle$ . In order to establish which of the standardised estimators of adjustments can represent gross errors (they do not belong to  $\Delta \bar{v}$ ), we determine a covariance matrix of the estimator for the vector of adjustments  $\hat{\mathbf{V}}$  for  $m_0 = 1$ .

Then we carry out the classification:

$$\hat{\mathbf{C}}_{\hat{\mathbf{V}}(m_0=1)} = \begin{bmatrix} 0.15130 & 0.02854 & 0.03774 & 0.10470 & 0.04162 \\ 0.02854 & 0.14979 & -0.11445 & 0.01575 & 0.02913 \\ 0.03774 & -0.11445 & 0.11899 & 0.01178 & 0.03044 \\ 0.10470 & 0.01575 & 0.01178 & 0.11590 & -0.06476 \\ 0.04162 & 0.02913 & 0.03044 & -0.06476 & 0.21402 \end{bmatrix}$$

$$\bar{v}_1 = \frac{(\hat{v}_1)_{(m)}}{(\sqrt{0.15130})_{(m)}} = 15.6 \notin \Delta\bar{v}$$

$$\bar{v}_2 = \frac{(\hat{v}_2)_{(m)}}{(\sqrt{0.14979})_{(m)}} = 3.1 \notin \Delta\bar{v}$$

$$\bar{v}_3 = \frac{(\hat{v}_3)_{(m)}}{(\sqrt{0.11899})_{(m)}} = 4.2 \notin \Delta\bar{v}$$

$$\bar{v}_4 = \frac{(\hat{v}_4)_{(m)}}{(\sqrt{0.11590})_{(m)}} = 12.2 \notin \Delta\bar{v}$$

$$\bar{v}_5 = \frac{(\hat{v}_5)_{(m)}}{(\sqrt{0.21402})_{(m)}} = 3.8 \notin \Delta\bar{v}$$

$$\hat{\mathbf{P}}^{(1)} = \mathbf{T}(\bar{\mathbf{V}}^{(0)})\mathbf{P}^{(0)} = \begin{bmatrix} 0.09897 & 0 & 0 & 0 & 0 \\ 0 & 3.89754 & 0 & 0 & 0 \\ 0 & 0 & 3.61422 & 0 & 0 \\ 0 & 0 & 0 & 0.49091 & 0 \\ 0 & 0 & 0 & 0 & 3.75768 \end{bmatrix}$$

coordinates' increments:

$$\hat{\mathbf{d}}_X^{(1)} = -(\mathbf{A}^T \hat{\mathbf{P}}^{(1)} \mathbf{A})^{-1} \mathbf{A}^T \hat{\mathbf{P}}^{(1)} \mathbf{L} = \begin{bmatrix} -72.92 \\ 89.62 \end{bmatrix}_{(m)}$$

the vector of adjustments;

$$\hat{\mathbf{V}}^{(1)} = \mathbf{A} \cdot \hat{\mathbf{d}}_X^{(1)} + \mathbf{L} = \begin{bmatrix} 9.56 \\ 0.12 \\ 0.04 \\ 0.51 \\ 0.31 \end{bmatrix}_{[^\circ]}$$

The obtained results show that none of the standardised estimators of adjustments belongs to the acceptable range. If we employed methods that are used in geodesy, we would have to reject measurements that are fraught with such errors and then repeat the measurements. However, it would be difficult to cause a vessel which is going to a port to return and then make the measurements for the same positions again. In order to neutralise these errors, we can adjust the results of measurements again, thus making the results of observations robust to gross errors, by using the Danish attenuation function. If we assume that  $\mathbf{V}^{(0)} = \hat{\mathbf{V}}$ ,  $\mathbf{P}^{(0)} = \hat{\mathbf{P}}$  and  $l = 0.02$ ,  $g = 2$  (control parameters in the Danish attenuation function), we can calculate the following: the value of the attenuation function (where  $\Delta\bar{v} \in \langle -k; k \rangle = \langle -2; 2 \rangle$ );

$$\bar{v}_1^{(0)} = \bar{v}_1 \notin \Delta\bar{v} \rightarrow t(\bar{v}_1^{(0)}) = \exp[-l(15.6 - k)^g] = 0.025$$

$$\bar{v}_2^{(0)} = \bar{v}_2 \notin \Delta\bar{v} \rightarrow t(\bar{v}_2^{(0)}) = \exp[-l(3.1 - k)^g] = 0.974$$

$$\bar{v}_3^{(0)} = \bar{v}_3 \notin \Delta\bar{v} \rightarrow t(\bar{v}_3^{(0)}) = \exp[-l(4.2 - k)^g] = 0.904$$

$$\bar{v}_4^{(0)} = \bar{v}_4 \notin \Delta\bar{v} \rightarrow t(\bar{v}_4^{(0)}) = \exp[-l(12.2 - k)^g] = 0.123$$

$$\bar{v}_5^{(0)} = \bar{v}_5 \notin \Delta\bar{v} \rightarrow t(\bar{v}_5^{(0)}) = \exp[-l(3.8 - k)^g] = 0.939$$

and the attenuation matrix:

$$\mathbf{T}(\bar{\mathbf{V}}^{(0)}) = \begin{bmatrix} 0.025 & 0 & 0 & 0 & 0 \\ 0 & 0.974 & 0 & 0 & 0 \\ 0 & 0 & 0.904 & 0 & 0 \\ 0 & 0 & 0 & 0.123 & 0 \\ 0 & 0 & 0 & 0 & 0.939 \end{bmatrix}$$

Then we perform calculations by using the iteration method; as part of step no. 1 we calculate the following: the attenuation matrix;

diagonal elements of matrix  $\mathbf{C}_{\mathbf{V}(m_0=1)}^{(1)}$

$$\mathbf{C}_{\mathbf{V}(m_0=1)}^{(1)} = \begin{bmatrix} 9.566 & 0.057 & 0.095 & 0.619 & 0.271 \\ 0.057 & 0.144 & -0.130 & -0.012 & 0.019 \\ 0.095 & -0.130 & 0.124 & -0.049 & 6.165 \cdot 10^{-3} \\ 0.619 & -0.012 & -0.049 & 1.296 & -0.336 \\ 0.271 & 0.019 & 6.165 \cdot 10^{-3} & -0.336 & 0.108 \end{bmatrix}$$

the value of the attenuation function;

$$v_1^{(1)} = 3.09 \notin \Delta v \rightarrow t(v_1^{(1)}) = \exp[-l(3.09 - k)^g] = 0.997$$

$$\bar{v}_2^{(1)} = 0.33 \in \Delta\bar{v} \rightarrow t(\bar{v}_2^{(1)}) = 1$$

$$\bar{v}_3^{(1)} = 0.11 \in \Delta\bar{v} \rightarrow t(\bar{v}_3^{(1)}) = 1$$

$$\bar{v}_4^{(1)} = 0.44 \in \Delta\bar{v} \rightarrow t(\bar{v}_4^{(1)}) = 1$$

$$\bar{v}_5^{(1)} = 0.94 \in \Delta\bar{v} \rightarrow t(\bar{v}_5^{(1)}) = 1$$

and the attenuation matrix

$$\mathbf{T}(\bar{\mathbf{V}}^{(1)}) = \begin{bmatrix} 0.977 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

Since only one standardised adjustment falls within the acceptable range, the iterative process of solving the adjustment problem by using a method that is robust to gross errors can be accelerated by adopting increasingly higher values of parameter  $l$ . Finally, the following solution is obtained:

$$\hat{\mathbf{d}}_X^{(1)} = -(\mathbf{A}^T \hat{\mathbf{P}}^{(1)} \mathbf{A})^{-1} \mathbf{A}^T \hat{\mathbf{P}}^{(1)} \mathbf{L} = \begin{bmatrix} -36.07 \\ 19.87 \end{bmatrix}_{(m)}$$

$$\hat{\mathbf{V}}^{(1)} = \mathbf{A} \cdot \hat{\mathbf{d}}_X^{(1)} + \mathbf{L} = \begin{bmatrix} 9.87 \\ 0.09 \\ -0.02 \\ 0.15 \\ 0.15 \end{bmatrix}_{[^\circ]}$$

$$m_{\hat{x}_{Z_2}} = 108.68 [m]$$

$$m_{\hat{y}_{Z_2}} = 195.80 [m]$$

$$M_{po} = \sqrt{m_{\hat{x}_{Z_2}}^2 + m_{\hat{y}_{Z_2}}^2} = 223.94 [m]$$

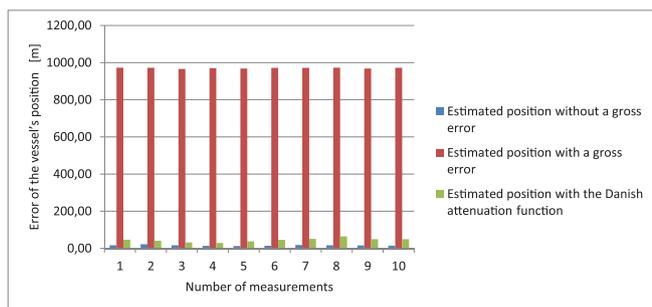


Figure 7. Error of the vessel's position which was determined during the study (source: own work).

## CONCLUSIONS

Radar is one of the basic tools which are used in marine and land navigation and which makes it possible to visualise the surrounding area. This tool is used as one of the main sources of the information about vessel traffic at sea that is passed on to vessel traffic monitoring centres. If information about moving objects is obtained from several radar stations, this may be a step forward in assessing maritime safety since more complete and more accurate data will be available about vessel traffic in waters that are monitored by radar stations. Once it is possible to determine vessels' exact positions, maritime safety in coastal waters will improve because information that is necessary for safe navigation will be sent to the bridge.

When measuring the parameters that are used to determine a vessel's position, we can obtain results which will be fraught with serious errors if the measurement process is disturbed, and in extreme cases such errors can largely deviate from the expected results.

Consequently, if a vessel's position is erroneously determined, nautical charts may be incorrectly used for navigation and this can pose a threat to maritime safety.

The authors are not familiar with any other solutions which would involve simultaneously using information obtained from all coastal radar stations operating within the VTS system as well estimation methods that are employed in geodesy. At the same time, the use of the whole body of information that is provided by VTS systems as well as of modern M-estimation methods opens up new possibilities for the VTS system operators as regards traffic control and updating the positions of vessels in waters that are monitored by systems supervising vessel traffic.

In order to reduce the impact of gross errors on the process of establishing a vessel's position, one can use robust estimation methods with properly selected attenuation functions; this will determine the accuracy with which the position will be established. M-estimation methods can significantly correct such errors and limit their influence on how a course line is drawn on a nautical chart. This paper presents basic robust adjustment methods. The analysis has shown that a method utilising the Danish attenuation function is best suited to the case that is presented here. If the other, i.e. Hampel's and Hubert's, attenuation functions are used when serious errors occur, this leads to an unexpected shift in a vessel's position.

## REFERENCES

1. Bole A.G., Wall A.W., Norris A., (2014). Radar and ARPA Manual, Third Edition: Radar, AIS and Target Tracking for Marine Radar Users 3rd Edition, Elsevier, ISBN 978-0-08-097752-2, Waltham, MA.
2. Borre K.J, Jorgensen P.C., Kubik K., (1983). Robust Adjustment of the Danish Fundamental Triangulation Network. Zeszyty Naukowe Akademii Górniczo - Hutniczej im. Stanisława Staszica, nr 79, s. 57 - 69, Kraków.
3. Czaplewski K. (2004). Positioning With Interactive Navigational Structures Implementation. Annual of Navigation no 7/2004, ISSN 1640-8632, Gdynia.
4. Hampel F.R, Ronchetti E.M, Rousseeuw P.J, Stahel W.A, (2005). Robust Statistics. The Approach Based on Influence Functions. John Wiley, ISBN 978-0-471-73577-9, New York, NJ
5. Jianjun Z. (1996). Robustness and the Robust Estimate. Journal of Geodesy, nr 70, s. 586 - 590, Heidelberg.
6. Monahan K. (2008) The Radar Book: Effective Navigation and Collision Avoidance, Fine Edge Productions, ISBN 978-1-932310-36-8, Anacortes, WA.
7. Teunissen P.J.G. (2004). Adjustment theory: an introduction (mathematical geodesy and positioning), Delft University Press, ISBN 978-90-407-1974-5, Delft.

8. Wiśniewski Z. (2004). Metody opracowania wyników pomiarów w nawigacji i hydrografii. AMW, ISBN 83-87280-65-8, Gdynia.
9. Wiśniewski Z. (2009). Rachunek wyrównawczy w geodezji (z przykładami). UWM, ISBN 83-7299-399-8, Olsztyn.
10. Yang Y., Song L., Xu T., (2002). Robust Estimator for Correlated Observations Based on Bifactor Equivalent Weights. Journal of Geodesy, nr 76, s. 353 - 358, Heidelberg.

#### **CONTACT WITH THE AUTHORS**

Sławomir Świerczyński

Polish Naval Academy  
s.swierczynski@amw.gdynia.pl

Krzysztof Czaplewski

Gdynia Maritime University  
krzysztof@czaplewski.pl  
**POLAND**

# METHODS FOR OPTIMIZATION OF SEA WATERWAY SYSTEMS AND THEIR APPLICATION

Stanisław Gućma, Prof.

Wojciech Ślącza, Assoc. Prof.

Maritime University of Szczecin, Poland

## ABSTRACT

*This paper presents developed optimization methods for sea waterway systems. Objective function of optimization of sea waterway systems have been formulated and its detail form for fairways has been determined. Three probabilistic methods for determining safe bed-breadth of fairway are described. Limitations of the presented methods are also discussed. The developed optimization methods for sea waterway systems have been applied to determine parameters of a modernized fairway between Świnoujście and Szczecin as well as its navigation systems.*

**Keywords:** Optimization of sea waterway systems; sea traffic engineering; probabilistic methods of determination safe breadth of fairways.

## INTRODUCTION

System of sea waterways formulated in terms of sea traffic engineering consists of several separated segments (n). Each of them consists of three basic elements [1]:

- waterway subsystem,
- ship position determination subsystem ( navigation subsystem),
- ship traffic control subsystem.

The elements mutually interact and significantly influence features of the whole system.

In practice all the sea waterway segments are often subjected to the same traffic control subsystem. On this assumption a sea waterway system can be reduced to two basic elements only :

- waterway subsystem,
- navigation subsystem.

System of sea waterways is defined by parameters of its elements ( subsystems). The two elements of sea waterway system at its particular segments are function of conditions for safe operation of ships. Therefore a system of i-th segment of sea waterway can be described by the following matrix expression:

$$\begin{bmatrix} \mathbf{A}_i \\ \mathbf{N}_{in} \end{bmatrix} = \mathbf{f}(\mathbf{W}_i)$$

In the above given formula the following matrix description was applied :

- for waterway subsystem :

$$\mathbf{A}_i = \begin{bmatrix} t_i \\ l_i \\ D_i \\ h_i \end{bmatrix}$$

where:

$t_i$  – type of i-th segment of waterway;

$l_i$  – length of i-th segment of waterway;

$D_i$  – breadth of accessible navigation area of i-th segment of waterway;

$h_i$  – minimum depth of i-th segment of waterway;

- for navigation subsystem:

$$\mathbf{N}_{in} = \begin{bmatrix} d_{in} \\ m_{in} \\ n_{in} \end{bmatrix}$$

where:

$d_{in}$  – accuracy of n-th navigation system at i-th segment of waterway (standard deviation);

$m_{in}$  – accessibility of n-th navigation system at i-th segment of waterway (depending on day-time and visibility);

$n_{in}$  – reliability of n-th navigation system at i-th segment of waterway ( technical reliability);

A state vector of conditions for safe operation of ships at i-th segment of a considered waterway can be presented as follows [1]:

$$\mathbf{W}_i = [t_{yp}, L_c, B, T, V_i, C_i, \mathbf{H}_i]$$

where:

$t_{yp}$  – type of „ a maximum ship”;

$L_c$  – length of „ a maximum ship”;

$B$  – breadth of „ a maximum ship”;

$T$  – draught of „ a maximum ship”;

$V_i$  – permissible speed of „ a maximum ship” over i-th segment of waterway;

$C_i$  – tug assistance at i-th segment of waterway (required number of tugs and their towing powers);  
 $H_i$  – vector of hydrometeorological conditions permissible for „a maximum ship” over i-th segment of waterway.

## OPTIMIZATION OF SEA WATERWAY SYSTEM

The optimization of sea waterway system is carried out at determination of parameters of basic elements of a system under construction or re-construction (waterway subsystem, navigation subsystem). Parameters of the sea waterway system elements are function of designed (assumed). conditions of ship’s safe operation (vector of state of ship’s safe operation).

In issues of optimizing the parameters of sea waterway systems, an objective function is cost of construction (modernization) of waterway system elements and operational cost of its subsystems. On the assumptions, the objective function can be described as follows [5]:

$$Z = (A1 + A2 + N1 + N2 + S) \rightarrow \min$$

where:

$$\begin{aligned} A1 &= f_1(\mathbf{D}_i, h_{xy}), \text{ wherein } (x, y) \in \mathbf{X1} \times \mathbf{Y1}, \\ A2 &= f_2(\mathbf{D}_i, h_{xy}), \\ N1 &= f_3(\mathbf{D}_i, h_{xy}), \\ N2 &= f_4(\mathbf{D}_i), \\ S &= f_5(\mathbf{D}_i), \end{aligned}$$

i.e.:

$$Z = F(\mathbf{D}_i, h_{xy}) \rightarrow \min$$

under constraints:

$$\left. \begin{aligned} \mathbf{d}_i(1 - \alpha) \subset \mathbf{D}_i \\ \bigwedge_{p(x, y) \in \mathbf{D}_i} h_{xy}(t) \geq T_{xy}(t) + \Delta_{xy}(t) \end{aligned} \right\}$$

where:

$\mathbf{D}_i$  – accessible navigation water area complying with safe depth condition at i-th segment of waterway;  
 $d_i(1-\alpha)$  – safe manoeuvring area at i-th segment of waterway for a „maximum ship” manoeuvring in set navigation conditions, determined on the confidence level of  $1-\alpha$ ;  
 $Z$  – cost of construction and operation of a waterway system;  
 $A1$  – cost of construction ( re-construction ) of a waterway;  
 $A2$  – cost of operation of a waterway;  
 $N1$  – cost of construction of a ship position determination subsystem (navigation systems);  
 $N2$  – cost of operation of navigation systems;  
 $S$  – ship operation cost connected with passing a considered waterway ( pilotage, tug assistance etc );  
 $h_{xy}$  – water area depth in the point  $x, y$ ;

$T_{xy}$  – ship draught in the point  $x, y$ ;  
 $\Delta_{xy}$  – underkeel clearance in the point  $x, y$ .

In the case of sea waterways of constant depth ( $h_{xy} = \text{const}$ ) the objective function can be expressed as follows:

$$Z = F(\mathbf{D}_i)$$

under constraints:

$$D_j \geq d_j(1 - \alpha)$$

where:

$D_j$  – accessible bed-breadth of a fairway in j-th point of fairway axis;  
 $d_j(1-\alpha)$  – safe breadth of a fairway on bottom level of a „maximum ship” manoeuvring in set navigation conditions in j-th point of fairway axis, determined on the confidence level of  $1-\alpha$ .

## PROBABILISTIC METHODS FOR DETERMINING THE SAFE BED- BREADTH OF FAIRWAY

Safe breadth of fairway bed is a random variable of normal distribution [2; 6]. In optimizing sea waterway systems, probabilistic methods have found their application for determining the safe breadth of fairway bed. These are simulation methods and CIRM method (CIRM stands for Sea Traffic Engineering Centre). In the probabilistic model safe breadth of manoeuvring water area is a random variable. For the parameter normal distribution is usually assumed, which shows a good conformity with real life (Fig.1). Distribution of distance between extreme points of manoeuvring water area and fairway midline can be expressed in the following form [4]:

$$\begin{aligned} d_{lj}(y) &= \frac{1}{\delta_{lj} \sqrt{2\Pi}} e^{-\frac{(y-m_{lj})^2}{2\delta_{lj}^2}} \\ d_{pj}(y) &= \frac{1}{\delta_{pj} \sqrt{2\Pi}} e^{-\frac{(y-m_{pj})^2}{2\delta_{pj}^2}} \end{aligned}$$

$d_{lj}(y); d_{pj}(y)$  – densities of distribution of distances between fairway midline and left-hand and right-hand side boundaries of manoeuvring water area;

$m_{lj}; \delta_{lj}$  – mean value and standard deviation of distances between fairway midline and left-hand side boundary of manoeuvring water area;

$m_{pj}; \delta_{pj}$  – mean value and standard deviation of distances between fairway midline and right-hand side boundary of manoeuvring water area;

In simulation methods a simplified method for determining the safe breadth of fairway consists in determining it in function of longitudinal coordinate of fairway axis which,

in this case, is equivalent to water area axis. Breadth of safe manoeuvring water area in  $i$ -th point of fairway or port entrance can be determined for a given series of simulation manoeuvres (series of trials) as follows [4]:

$$d_j(1-\alpha) = (y_{lj} + cs_{lj}) + (y_{pj} + cs_{pj})$$

where:

$d_j(1-\alpha)$  – safe breadth of a fairway in  $j$ -th point of fairway on the confidence level of  $1-\alpha$ ;

$Y_{lj}; Y_{pj}$  – mean arithmetic value of maximum distances from ship's point to the left and right to  $j$ -th point of fairway axis (estimators:  $m_l$  and  $m_p$ );

$s_{lj}; s_{pj}$  – standard deviation of trial series for maximum distances from ship's point to the left and right to  $j$ -th point of fairway axis (estimators:  $\sigma_l$  and  $\sigma_p$ );

$c$  – a coefficient depending on an assumed confidence level (e.g. taken as  $c=1,96$  for the confidence level:  $1-\alpha = 0,95$ ).

Arrangement of a safe manoeuvring water area on a fairway is presented in Fig.1.

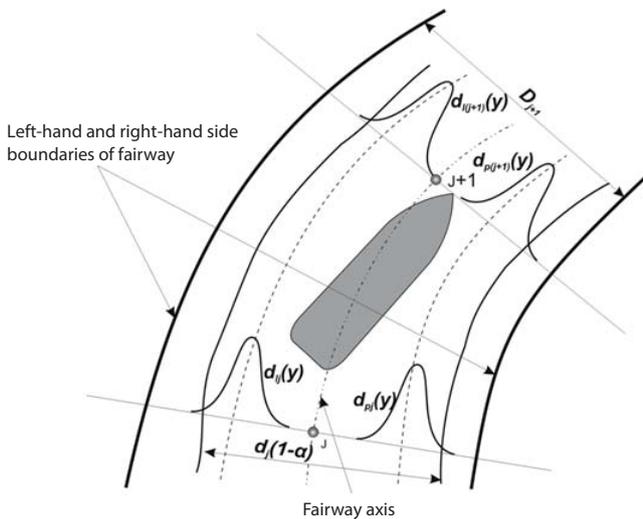


Fig. 1. Arrangement of safe manoeuvring water area on fairway

In the CIRM deterministic-probabilistic method for determining the safe breadth of fairway it was assumed that the standard deviation of maximum distances from ship's point to the left and right to  $j$ -th point of fairway axis is approximated by standard deviation of error in directional location of ship's side to the left and right, determined for an appropriate confidence level. The error in directional location of ship's side is determined for a navigation system used by the ship over a given segment of fairway [3]. Values of the error are properly described by normal distribution. It is called a navigational component of safe manoeuvring water area, i.e.:

$$d_{npj}(1-\alpha) = c\sigma_{jp}$$

$$d_{nlj}(1-\alpha) = c\sigma_{jl}$$

Under the above mentioned assumptions the safe breadth of fairway can be determined as follows (Fig. 2) [9]:

$$d_j(1-\alpha) = d_{mj} + d_{npj}(1-\alpha) + d_{nlj}(1-\alpha) \text{ [m]}$$

where:

$d_j(1-\alpha)$  – safe breadth of manoeuvring water area on the confidence level of  $(1-\alpha)$ , [m];

$d_{mj}$  – manoeuvring component of breadth of safe manoeuvring water area determined by using the deterministic method, [m];

$d_{npj}(1-\alpha); d_{nlj}(1-\alpha)$  – navigational (left and right) components of breadth of safe manoeuvring water area on the confidence level of  $(1-\alpha)$ , [m].

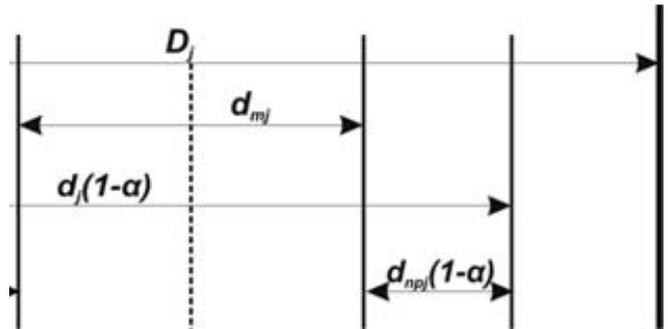


Fig. 2. Safe breadth of fairway on a given confidence level

In both the cases, after taking into account a breadth margin associated with bank effect, the following safety condition for navigation over a given fairway is to be fulfilled:

$$D_j \geq d_j(1-\alpha) + d_{rp} + d_{rl}$$

where:

$d_{rp}; d_{rl}$  – breadth margin for right and left side of fairway,

In the detail design stage of sea waterways including determination of safe bed-breadth values of fairways the three following methods may be used:

1. The simulation method for determining ship motion in real time by making use of non-autonomous models including navigator as an element [7]. The limitations of the method are:

- possible application to all kinds of waterways except of long (> 3 NM) straight segments of fairways;
- required application of appropriate manoeuvring simulators and software;
- proper experience required from persons making use of the method for simulation research and mathematical modeling;
- a high accuracy of the method which depends on applied simulators, ship models, design of an experimental system and run of processing the obtained results;
- the method is labour- and cost-consuming and requires to carry out long-lasting investigations engaging a large group of navigators (experts).

2. The simulation method for determining ship motion in an accelerated time run by using autonomous models including an element being a model of decision-making navigator [7].

The limitations of the method are:

- possible application, presently, to only straight fairways and in some cases to fairway bends (appropriate accuracy);
- proper experience required from persons making use of the method in the area of design of human decision-making process and ship motion simulation in accelerated time run;
- the method is labour-consuming in the stage of formulating the simulation model.

Run of such experiment is fast and not labour-consuming.

3. The CIRM deterministic-probabilistic method developed by Sea Traffic Engineering Centre, Maritime University of Szczecin [7].

The limitations of the method are the following:

- possible application to all kinds of waterways;
- a lower accuracy compared with other simulation methods;
- the method, being empirical, does not require to carry out any experimental investigations.

## APPLICATION OF THE DEVELOPED METHODS TO MODERNIZATION OF ŚWINOUJŚCIE-SZCZECIN FAIRWAY

The modernization of Świnoujście-Szczecin fairway consists in dredging the fairway up to 12,5 m depth, its appropriate broadening as well as fitting it with navigation systems making safe navigation of to-be-operated ships, possible.

As assumed, the basic navigation systems for Świnoujście-Szczecin fairway will be the following:

- a terrestrial system (optical one) accessible at the visibility up to 2 NM;
- a pilotage navigation system (PNS);

whereas an additional system will be:

- a radar system.

The „maximum ships” assumed for the calculations are the following [6]:

- a cruising ship ( $L_c = 260$  m;  $T = 9,0$  m);
- a container carrier ( $L_c = 210$  m;  $T = 11,0$  m);
- a bulk carrier ( $L_c = 195$  m;  $T = 11,0$  m);

and, taking into account that the ships do not carry dangerous goods, the confidence level equal to

$$1 - \alpha = 0,95$$

was assumed.

Under such assumptions the objective functions was expressed in the following form:

$$Z = \min D_j$$

with the constraints:

$$d_{ijk}(0,95) \leq D_j$$

where:

$Z$  – cost of modernization of Świnoujście-Szczecin fairway (its dredging, broadening, and modernizing the navigation system);

$D_j$  – accessible bed-breadth of the fairway at 12,5 m depth contour in  $j$ -th point of fairway axis;

$d_{ijk}$  – safe bed-breadth of the fairway in  $j$ -th point of fairway axis for  $i$ -th type of „maximum ship” passing through the fairway in  $k$ -th navigation conditions.

Determination of a minimum value of the safe bed-breadth,  $d_{ijk}(0,95)$ , of the fairway was performed in two phases.

In the 1<sup>st</sup> phase a minimum safe bed-breadth of the fairway on the confidence level of 0,95 was determined for particular segments of Świnoujście-Szczecin fairway between 16<sup>th</sup> and 67<sup>th</sup> km of its length by using the CIRM method. Fig. 3 shows calculation results for the cruiser going from Świnoujście to Szczecin.

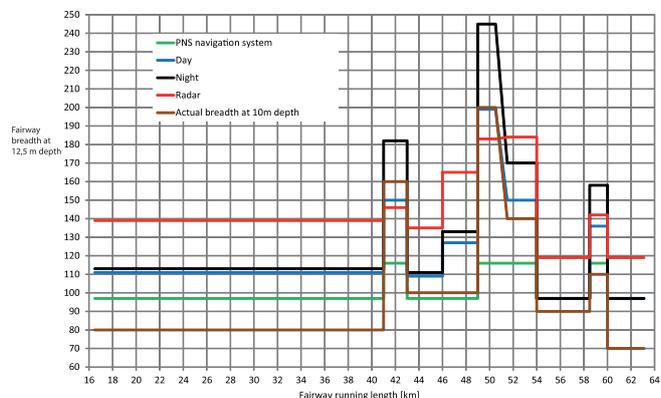


Fig. 3. Safe breadth of Świnoujście-Szczecin fairway for the cruiser going from Świnoujście to Szczecin, calculated for the existing navigation marking on the confidence level of 0,95

On the basis of calculation results reached with the use of CIRM method, final values of bed-breadth of the fairway were obtained for its straight segments as well as parameters of their navigation systems.

In the 2<sup>nd</sup> phase, the simulation method of ship motion in real time was applied to calculate safe breadth values of fairway bends. The research was performed by using a Konsberg Maritime AS Polaris multi-bridge manoeuvring simulator fitted with 3D-projection visualization. By means of the simulator, 3D geometrical models of the considered fairway bends as well as simulation models of the examined „maximum ships” were formed, assuming safe isobates for 12,5 m depth determined by using CIRM method.

The investigations were carried out in series for two directions of wind of 10 m/s speed and sea current of 0,7 kn speed. Particular simulation passages in series, of 12÷15 in number, were executed by different pilots. On the basis of the values of safe bed-breadth of the fairway, calculated on the 0,95 confidence level, accessible navigation water areas for Świnoujście-Szczecin fairway of 12,5 m in depth and parameters of navigation systems for these areas were determined.

Exemplary results of the 2<sup>nd</sup> phase investigations for the Święta bend (58,5 km ÷ 60,0 km) are presented in Fig. 4 and 5.

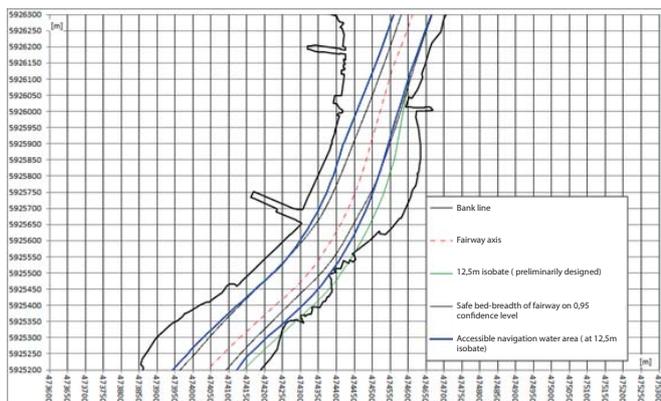


Fig. 4. Święta bend. Safe bed-breadth values of the fairway, determined on 0,95 confidence level, and the accessible navigation water area ( north part)

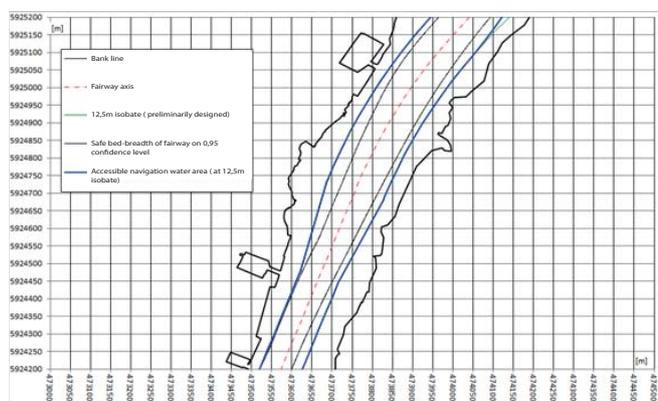


Fig. 5. Święta bend. Safe bed-breadth values of the fairway, determined on 0,95 confidence level, and the accessible navigation water area ( south part)

## CONCLUSIONS

In this paper has been described the optimization method for sea waterway systems, based on three probabilistic methods for determining safe values of fairway bed-breadth. These are the following:

- the simulation method for determining ship motion in real time by making use of non-autonomous models;
- the simulation method for determining ship motion in accelerated run of time;
- the CIRM deterministic probabilistic method.

The developed optimization method may be applied in the detail design stage of waterway system; however, taking into account accuracy and cost consumption of the particular methods for determining safe values of fairway bed-breadth, the following sequence of their application has been proposed:

- Phase 1. Optimization of parameters of waterway system, carried out by means of CIRM method to reach the following:
- determination of optimum values of parameters of straight segments of fairways and their navigation systems;
  - preliminary determination of parameters of complex manoeuvring water areas of fairways, to be used in Phase 2.

Phase 2. Optimization of parameters of complex manoeuvring water areas of waterway system, carried out by means of simulation methods of ship motion in real time to reach the following:

- determination of optimum values of parameters of complex manoeuvring water areas of fairways and their navigation systems.

The above described two-phase optimization method of sea waterways was applied to the modernization project of Świnoujście-Szczecin fairway. The modernization consisted in dredging the fairway from 10,5 m up to 12,5 m depth, that forced:

- to change conditions of safe operation of ships sailing along the fairway;
- to determine optimum parameters of the fairway and its navigation systems.

## BIBLIOGRAPHY

1. Gucma S. (Editor): „Navigational analysis of modernization project of Świnoujście-Szczecin fairway (dredging up to 12,5m depth)” (in Polish). Scientific research ordered by Europrojekt Gdańsk S.A.; project investor: Szczecin Maritime Administration (Urząd Morski w Szczecinie). Maritime University of Szczecin, Szczecin 2014
2. Gucma L.: „Modeling of factors for risk of collision between floating units and port facility structures” (in Polish). (Studies 44), Maritime University of Szczecin, Szczecin 2005
3. Gucma S.: „Pilotage navigation” (in Polish). Publisher: Foundation for Promotion of Shipbuilding Industry and Maritime Economy (Fundacja Promocji Przemysłu Okrętowego i Gospodarki Morskiej), ISBN 83-919488-5-4. Gdańsk 2004.
4. Gucma S.: „The width of a fairway - comparison of models”. Journal of Konbin, No 1(13)2010, ISSN 1895-8281. Warszawa 2010.
5. Gucma S.: “Optimization of sea waterway system parameters in marine traffic engineering”. Journal of Konbin, No 2(26)2013, ISSN 1895-8281. Warszawa 2013.
6. Gucma S.: “Systems approach to sea waterways”. Scientific Bulletins of Maritime University of Szczecin, No 38 (110), 2014, ISSN 1733-8670. Szczecin 2014.

7. Gucma S., Gucma L., Zalewski P. : „Simulation research methods in sea traffic engineering” (in Polish). A monograph edited by Gucma S. ; ; Publishing House of Maritime University of Szczecin (Wydawnictwo Naukowe Akademii Morskiej w Szczecinie), ISBN 978-83-89901-29-3, Szczecin 2008.
8. Gucma S., Ślęczka W., Zalewski P. : „Parameters of fairways and navigation systems determined with the use of safe navigation criteria” (in Polish). A monograph edited by Gucma S. ; Publishing House of Maritime University of Szczecin (Wydawnictwo Naukowe Akademii Morskiej w Szczecinie), ISBN 978-83-89901-82-8. Szczecin 2013.
9. Gucma S., Zalewski P.: “Deterministic-probabilistic method of waterway design parameters determination”. Annual of Navigation, No 19/2012, part 2., ISSN 1640-8632. Gdynia 2012.

## CONTACT WITH AUTHOR

Stanisław Gucma  
s.gucma@am.szczecin.pl

Wojciech Ślęczka  
w.slaczka@am.szczecin.pl

Maritime University of Szczecin  
Wały Chrobrego 1-2 Str.  
70-500 Szczecin  
POLAND

# CELESTIAL NAVIGATION FIX BASED ON PARTICLE SWARM OPTIMIZATION

Ming-Cheng Tsou, Assoc. Prof.

National Kaohsiung Marine University, Taiwan

## ABSTRACT

*A technique for solving celestial fix problems is proposed in this study. This method is based on Particle Swarm Optimization from the field of swarm intelligence, utilizing its superior optimization and searching abilities to obtain the most probable astronomical vessel position. In addition to being applicable to two-body fix, multi-body fix, and high-altitude observation problems, it is also less reliant on the initial dead reckoning position. Moreover, by introducing spatial data processing and display functions in a Geographical Information System, calculation results and chart work used in Circle of Position graphical positioning can both be integrated. As a result, in addition to avoiding tedious and complicated computational and graphical procedures, this work has more flexibility and is more robust when compared to other analytical approaches.*

**Keywords:** Particle Swarm Optimization; Celestial navigation; Intercept method

## INTRODUCTION

Even today, when navigation is dominated by GPS, a traditional celestial fix still serves as a valuable backup measure. Nevertheless, as we enter the 21st century, traditional methods for computing a fix using celestial navigation can no longer meet the requirements of modern vessels in terms of calculation speed and precision. The need arises for further improvements that utilize information technology. The 2010 amendment to the STCW Code placed a continuing emphasis on celestial navigation-related education and training. It also encouraged the usage of an electronic nautical almanac and celestial navigation calculation software. In response, many researchers have resorted to computer programs to deal with celestial navigation positioning. Problems that could not be solved previously using an inspection table can now be solved with the Spherical Triangle Method or the Vector-Matrix method, which give vessel positions directly. With these efforts, great advancements have been made in celestial navigation technology.

Because the independence of celestial navigation can complement other navigation methods, research into how to apply information technology in a celestial navigation approach proves especially relevant. This study uses particle swarm optimization (PSO) from the field of swarm intelligence, which mimics natural swarm optimization behaviours, due to its superior search ability. This technique is combined with a geographical information system (GIS) and the principle of using celestial circles of equal altitude for a fix to give a fast and accurate calculation of the Most Probable Position (MPP). The proposed method and framework can potentially be integrated into an Electronic Chart Display and Information System (ECDIS).

## PRINCIPLES OF A CELESTIAL NAVIGATION FIX

The purpose of celestial navigation, as traditionally practiced, is to determine the latitude and longitude of a vessel at a specific time, through observations of the altitudes of celestial bodies, which are used to determine the observed circle of position (COP). When more than two sets of data are obtained, the vessel's position can be calculated through graphical, combined graphical and computational, or direct computational procedures using sight reduction methods such as High-Altitude Observation and the Intercept Method (IM), or computational methods. The basic principles of a celestial fix remain unchanged today, and are the basis for a number of methods. The following is a review of several celestial navigation fix methods with their respective advantages and shortcomings.

### CIRCLE OF POSITION FIX PRINCIPLE

According to the relationship between celestial and Earth coordinates, in which they are each other's projections, an observer's COP is the projection of the circle of zenith distance onto the surface of the Earth. The centre of the COP is the Geographic Position (GP) of the celestial body. The radius of the circle is the zenith distance (co-altitude) of the celestial body (Figure 1). In order to estimate the vessel's position, one must observe at least two celestial bodies from the same location, thus producing two celestial COPs and two points of intersection. The point of intersection closest to the estimated vessel position is the observed vessel position. The principle of the COP fix is quite simple and can theoretically be carried out as long as one can plot the COP directly onto a chart.

However, it is not feasible in practice for the following reasons. Firstly, the radii of most circles of equal altitude are too large to be plotted on a chart. Secondly, graphical distortion at high latitudes is apparent on the commonly used Mercator chart, and the distortion increases with the latitude of the GP. Therefore, a graphical fix by directly plotting the COP is limited only to high-altitude observations. However, a high-altitude observation is only suitable for an observed altitude of greater than  $87^\circ$  (Chen et al., 2003). It is difficult to use a sextant for high-altitude sighting, and the probability of a bright star being near the zenith at any given time and place is small. Thus, there is little desire for high-altitude observations in ordinary navigation.

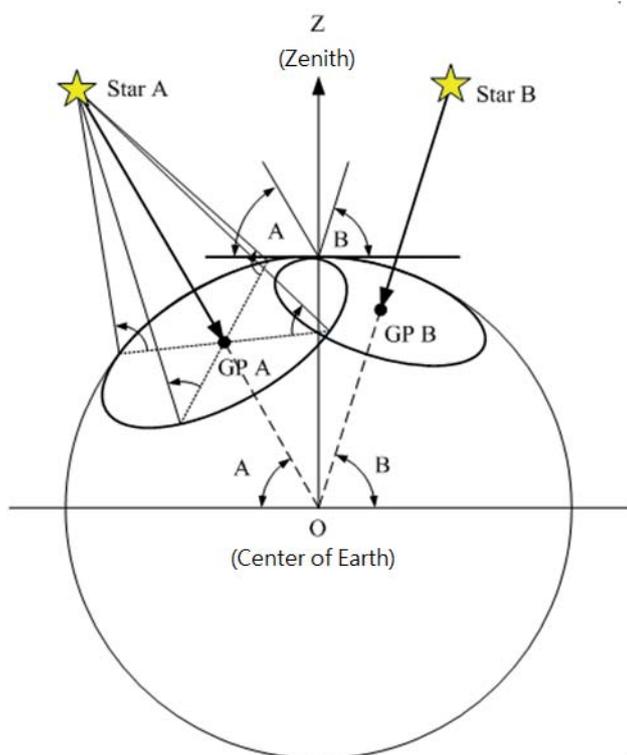


Fig. 1. The COP fix principle

### LINE OF POSITION FIX PRINCIPLE

Because the altitudes of most celestial bodies are less than  $87^\circ$ , Marcq de St Hilaire first introduced the Assumed Position (AP) to form the altitude difference, or the Intercept Method (IM), to overcome the limitation of high-altitude observations (Peacock, 2011), which has become the basis of virtually all present-day celestial navigation methods. The basic concept of the IM is to choose an AP at the nearest probable position and take it as the reference position to calculate the altitude and azimuth. By comparing the computed altitude ( $H_c$ ) and observed altitude ( $H_o$ ), the difference between the two altitudes (called the intercept or altitude difference) can be obtained. Therefore, once the AP, the computed azimuth of the body, and the intercept are all determined, the Line of Position (LOP) can be plotted. The COP can then be converted to the line of position (LOP).

However, the entire process consists of observing, calculating, and plotting. Even when performed by professional seafarers, one astronomical positioning will take around 20 minutes. Moreover, the accuracy of finding astronomical vessel positions by the IM is subject to the following two restrictions:

1. The distance between the AP and vessel's actual position largely affects the accuracy of the result. Therefore, this distance should not exceed 30 nautical miles (NM).
2. When the altitude of the observed celestial body exceeds  $70^\circ$  or the vessel is navigating in waters at high latitude, the resulting error of curvature will increase when using the LOP in lieu of the COP on the Mercator Chart.

Because of these restrictions limiting the accuracy, the IM has inherent drawbacks. Thus, a direct computation method is required, and the concept of the circle of equal altitude is reconsidered.

### DIRECT COMPUTATION METHOD FOR CELESTIAL NAVIGATION FIX

With recent advances in information technology, the problems of the celestial navigation fix can be solved using the Spherical Triangle Method, the Vector-Matrix Method, or other computational methods with the aid of computer programs. There are three classes of solutions:

1. Exact solutions to a two-body fix, such as those by Chiesa and Chiesa (1990), Spencer (1990), Gibson (1994), Chen et al. (2003), Hsu et al. (2005) and González (2008). These methods are based on full 3-D geometry, vector solution, or spherical trigonometry, and are not dependent on approximations or an assumed position.
2. Methods which are based on straight lines of position (each of which is a small arc of a circle of position) on a flat Earth near an estimated or assumed position. These methods can be applied to any number of observations  $\geq 2$ . Traditional hand-calculation and chart-based methods fall into this category, as does the least-squares method by DeWitt (1974), which was independently derived by Yallop and Hohenkerk (1990) and described in the Nautical Almanac and used in HMNAO's NavPac software.
3. Other least-squares techniques, including those of Watkins and Janiczek (1978), Severance (1989), Metcalf and Metcalf (1991), Kaplan (1995) and Wu (1991), are not based on circles or lines of position at all. Although they minimize the sum of the observational variances (the variance being the square of the altitude intercept) they do not rely on any geometric approximations. Some of them require an estimated (a priori) position, and in the case of Kaplan's algorithm, an estimated course and speed. Tsou (2012) employs genetic algorithm, similar to this study, to solve celestial navigation fix problems. This method can prevent from converging toward a local optima, but require a longer calculation time.

If direct computation is adopted, the choice of initial reference positions, such as the DR position or the EP, can be unconstrained. Furthermore, the previously described limitations are eliminated. In a computational procedure, the precision of a solution is no longer a major issue, as no graphical procedures and tables are required. In these circumstances, upon obtaining an accurate celestial Greenwich Hour Angle (GHA) and Declination (Dec), the precision of a celestial fix primarily depends on the accuracy of the observed altitude of the celestial body, provided that calculations are carried out correctly. This study is founded on these principles and improves upon the theory of the COP fix.

## CELESTIAL NAVIGATION FIX BASED ON PARTICLE SWARM OPTIMIZATION

This study utilizes the PSO technique from the field of swarm intelligence to solve celestial fix problems. Modifications are made specifically for its integration and implementation into celestial navigation.

### DIRECT FIX USING THE EQUATION OF COP.

In this study, the central idea behind the celestial circle of equal altitude fixing is to find the best fit to the altitude of a celestial body observed as a function of time. The  $H_c$  of a celestial body is given as a function of the Dec and GHA of the celestial body, and the observer's probable longitude ( $\lambda$ ), and latitude ( $L$ ) by:

$$\sin H_c = \sin L \cdot \sin Dec_i + \cos L \cdot \cos Dec_i \cdot \cos(GHA_i - \lambda) \quad (1)$$

where  $i = 1, \dots, n$  stands for the  $i$ th celestial body and its observation data.

The GHA and the Dec can be found according to time of observation in the Nautical Almanac or its electronic edition. Equation 1 is the equation of the celestial COP. Since the actual vessel position can be seen as a function of the altitude of the celestial body, which is a nonlinear function of  $L$  and  $\lambda$ , we cannot solve Equation 1 directly. Therefore, the latitude and longitude of the actual vessel position may be obtained by comparing the difference between  $H_c$  and  $H_o$  using the PSO technique and an appropriate fitting algorithm. With all the appropriate Dec and GHA values in Equation 1, any combination of altitudes of celestial bodies can be used.

### PARTICLE SWARM OPTIMIZATION

PSO is a computational method that was originally proposed by Eberhart and Kennedy (1995) for simulating behaviours in a bird flock to optimize a problem by iteratively trying to improve a candidate solution with regard to a given measure of quality. When using PSO to solve optimization problems, the solution corresponds to the position of a particle in the search space. Each particle has its own position and velocity, which decides the direction and distance of movement. In the process of iteration, every particle's movement is influenced by its local best-known position ( $p_{best}$ ) and is also guided

toward the best-known position ( $g_{best}$ ) in the search-space, which are updated as better positions are found by other particles. This is expected to move the swarm towards the best solutions. PSO does not require that the optimization problem be differentiable as is required by classic optimization methods such as gradient descent and quasi-Newton methods and can search very large spaces of candidate solutions.

Particle information can be represented by an  $n$ -dimensional vector. Its position can be expressed by  $X = (x_1, x_2, \dots, x_n)$ , and its velocity by  $V = (v_1, v_2, \dots, v_n)$ . The update equations for particle velocity and position are, respectively:

$$v_j^{t+1} = wv_j^t + c_1r_1(pb_{best_j}^t - x_j^t) + c_2r_2(g_{best_j}^t - x_j^t) \quad (2)$$

and

$$x_j^{t+1} = x_j^t + v_j^t \quad (3)$$

where  $w$  is the inertial coefficient, the main purpose of which is to generate a disturbance in order to prevent the calculation from becoming trapped in a local optimum. The maximum value of  $w$  is 0.9, and it decreases linearly with the evolutionary process to 0.4. A larger  $w$  means a stronger global optimization capacity and a weaker local optimization capacity, and vice versa. The terms  $c_1$  and  $c_2$  are acceleration coefficients (or learning factors). Usually,  $c_1 = c_2$  and both are random numbers in the interval  $[0, 2]$ . The coefficient  $c_1$  adjusts the flying velocity towards the best-known solution of the particle, while  $c_2$  adjusts the flying velocity towards the best-known solution of the entire swarm. Overly small values of  $c_1$  and  $c_2$  may cause particles to move further away from the target area and, conversely, overlarge values of  $c_1$  and  $c_2$  can result in them overshooting the target area. The values  $r_1$  and  $r_2$  are random numbers in the interval  $[0, 1]$ . The velocity  $v$  is usually limited to a certain range, i.e., it lies in the interval  $[-v_{max}, v_{max}]$ .

This study also includes the concept of a constriction factor that was later proposed by Kennedy and Eberhart (1997). It further improves on Equation (3) with:

$$x_j^{t+1} = x_j^t + nv_j^t \quad (4)$$

and

$$n = \frac{2}{|2 - f - \sqrt{f^2 - 4f}|} \quad (5)$$

where  $f = c_1 + c_2$  and  $n$  is the constriction factor. Similar to  $v_{max}$ , it is used to control and constrain the particle velocity. It also enhances the local search ability of the algorithm and improves the overall convergence.

The underlying principles and mechanisms of PSO are relatively simple, and the algorithm is easy to realize. At present, multi-body fixes in celestial navigation positioning or computer programs for the IM generally use the least squares mean method. However, a least squares mean method that proceeds from some initial value by decreasing the gradient of the goodness-of-fit parameter can converge to a local

minimum that is not the best solution when high-altitude observations are used (Metcalf and Metcalf, 1991). It was found in this study that the PSO method cannot only avoid this problem, but also possesses some advantages over other computational methods. Ten particles were used in this study. Each particle X represents a possible vessel position. It has two dimensions: longitude and latitude. The MPP is searched for using the optimization mechanism of the PSO method. More observation data will produce better estimation results.

## DESIGN OF THE OBJECTIVE FUNCTION

In the past, the most optimal vessel position in the majority of multi-body celestial fixing problems was solved for by data fitting. In this study, the variance of the altitude residuals is minimized through PSO. The altitude residual is defined as the difference between  $H_o$  and the  $H_c$  computed from Equation 1:

$$F(L, \lambda) = \min \frac{\sum_{i=1}^n \sqrt{[(\text{Sin}L \cdot \text{Sin}Dec_i + \text{Cos}L \cdot \text{Cos}Dec_i \cdot \text{Cos}(GHA_i - \lambda)) - H_{o_i}]^2}}{n} \quad (6)$$

The aim of Equation 6 is to find the most appropriate location (of longitude and latitude) in the solution space from the observational data of n celestial bodies so that F is minimized. When the result converges to meet the acceptance criteria, the MPP is found.

### CONSTRAINT CONDITIONS

Although PSO is not sensitive to initial conditions, the position of each particle X can be determined according to whether a reference position exists. If there is one, the particle position is chosen randomly based on this reference position and limited to the search range. Otherwise, the particle position can be any random location on the Earth. The reference position here is essentially different from the AP or the DR position in the IM. The reference position in this study is used to limit the search range and provide the effects of a heuristic search. Therefore, although the reference position can be set to be the same as the DR position, there is no limitation imposed on the DR position in the IM that the distance to the actual vessel position cannot exceed 30 NM. It is only an approximate reference position, and it can be hundreds or even thousands of NMs away.

However, a reference position closer to the DR position will result in a faster convergence. Therefore, if a DR position is available, it should be used as the reference position to accelerate the search. A reference position should be set in two-body celestial fix problems in order to determine which one of the two points of intersection from the COPs of the two celestial bodies is the actual vessel position. Multi-body celestial fixing does not require an initial guess for the position in this study, which means that a reference position does not need to be set. However, the number of iterations can be reduced by setting one, thus speeding up the search. The velocity of each particle must be within the search range.

### CORRECTION OF THE OBSERVED ALTITUDE IN A RUNNING FIX.

When performing the multi-body celestial fix at sea, there is some time difference between observations. It is therefore necessary to correct the zenith of the COP in the last observation to the same zenith position of the

following observation. When the distance travelled between two consecutive observations does not exceed 30 NM, the following formula is used to calculate the correction of the altitude ( $\Delta h$ ) in the running fix:

$$\Delta h = \frac{V}{60} \text{Cos} \Delta A \cdot \Delta T^m \quad (7)$$

in which V is the speed (in knots),  $\Delta A$  is the angle between the observed azimuthal direction of the celestial body and the true course of the vessel, and  $\Delta T^m$  is the time interval (in minutes) between the two observations.

### PLOTTING THE COP

Due to the large radius of the COP and graphic distortion at high latitudes when conducting graphical positioning on the Mercator chart, the application of direct graphical fix method was restricted to high-altitude observations. If the functionality of the GIS can be further modified, then the previous limitations on positioning with manual chart work will not apply on an electronic chart under a GIS environment.

This study proposes to include a function that constructs a COP, based on vector analysis (González, 2011), in a celestial navigation fix module in GIS. As long as there is a GP of the celestial body and a  $H_o$ , the entire COP can be directly constructed in a GIS environment in any projection. By utilizing the spatial analysis function in a GIS to obtain intersection points of the COP, the vessel's position can be measured, which can then facilitate positioning and serve as an aid to set the initial position, to assess the likely accuracy of the fix.

## RESULTS VALIDATION

The validity of this study was verified by using observation data from both two celestial bodies and multiple celestial bodies, and also by using data at high altitude. Using several important methods of astronomical navigation fixing published in the literature in recent years, three cases were tested and compared. The adopted comparison methods are: the traditional IM; the computerized IM by DeWitt (1974) that was published in the Nautical Almanac, U.S. Naval Observatory (Kaplan, 1995); the vector-matrix method by

Metcalf and Metcalf (1991); the Simultaneous Equal-Altitude Equation Method (SEEM) by Hsu et al., (2005); and the vector analysis method by González (2008). The methods proposed by Hsu et al., and González are only applicable in a two-body celestial fix. The other methods do not have this restriction. Visual Basic.Net 2010 was used as the development tool. In order to plot the COP and to add visual effects, the COM component of the GIS was also used. This was also used for education on celestial navigation and integration with an ECDIS in the future.

## TWO-BODY FIX

The data for this case study was taken from Hsu et al., (2005). It only contains observational data of two celestial bodies. All methods were tested using two different DR positions. Table 1 contains the relevant observation data and the results are presented in Table 2.

Tab. 1. Extract of relevant information from Hsu et al. (2005) for two-body observation

Body	ZT	Ho	GHA	Dec
Capella	20-03-58	15° 19.3'	131° 24.8'	45° 58.4' N
Alkaid	20-02-56	77° 34.9'	003° 14.2'	49° 25.7' N

Tab. 2. Two-body fix positions

Method	DR 1: $L = 41^\circ \text{N}$ , $\lambda = 017^\circ \text{W}$		DR 2: $L = 44^\circ \text{N}$ , $\lambda = 019^\circ \text{W}$	
	$L_1$	$\lambda_1$	$L_2$	$\lambda_2$
IM	41° 38.6' N	017° 08.1' W	41° 26.5' N	017° 28.1' W
DeWitt (5 iterations)	41° 39.1' N	017° 07.3' W	41° 39.1' N	017° 07.3' W
Metcalf (virtual star)	41° 39.1' N	017° 07.3' W	41° 39.3' N	017° 07.5' W
Hsu	$L = 41^\circ 39.1' \text{N}$ , $\lambda = 017^\circ 07.3' \text{W}$ (DR for reference)			
González	$L = 41^\circ 39.1' \text{N}$ , $\lambda = 017^\circ 07.3' \text{W}$ (DR for reference)			
PSO	$L = 41^\circ 39.1' \text{N}$ , $\lambda = 017^\circ 07.3' \text{W}$ (DR for reference)			

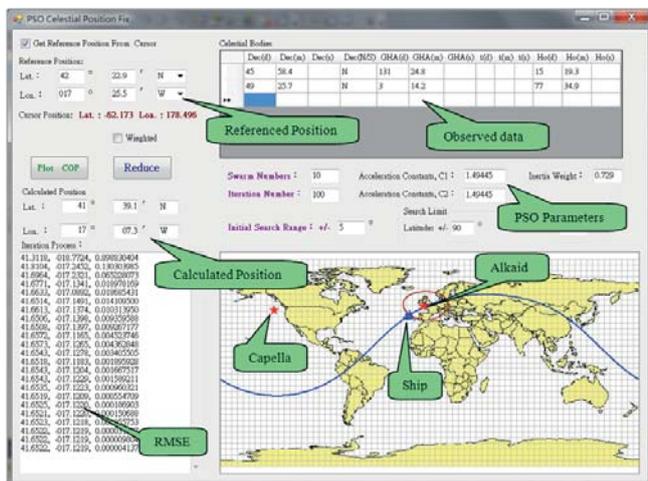


Fig. 2. Results of the two-body fix by PSO

The experimental results of the PSO are shown in Figure 2. Before execution, two COPs may first be plotted and displayed through a GIS module. The vessel position can be measured by mouse control and is used as a reference for setting a reference position. It can be seen from Table 1 that the observed altitude of Alkaid is as high as 77° 34.9', which exceeds the upper limit of 70° specified in the IM. Hsu et al., (2005) found that for high-altitude observation, an LOP drawn using the IM was

shifted due to curvature error, thus resulting in an inaccurate vessel position. It can be seen in Table 2 that all methods but the traditional IM can find the correct vessel position when the first DR position (DR 1) is used; i.e., the distance between the DR position and the actual vessel position is less than 30 NM. This implies that the traditional IM indeed has a limitation on the observed altitude.

It is worth noting that the method proposed by Metcalf was designed for a multi-body celestial fix. When used in a two-body fix, a virtual star must be assumed, and the latitude of the DR position is set to be its Dec. The longitude of the DR position needs to be converted to the GHA of the star. The computation can be performed after setting the observed altitude of the star as 90°. Therefore, the accuracy of positioning when using Metcalf's method in a two-body fix is affected by the selected DR position.

In the case of the second DR position (DR 2), a greater error occurs in the IM due to the larger distance between the DR position and the actual vessel position. DeWitt's method can reach the correct vessel position after 5 iterations. Metcalf's method, meanwhile, has produced a noticeable error. Although other methods, including PSO, can obtain the correct vessel position, some degree of prior DR knowledge is needed to help identify the answer. However, relatively speaking, the requirement on the accuracy of the initial position is lowered.

Some unlikely extreme experiments are designed in this study. One experiment is setting the DR position north of  $L = 60^\circ \text{N}$ ,  $\lambda = 17^\circ \text{W}$ , i.e., with high latitude and far away. Another is to have two COPs intercept at only one tangent point. In these two cases, the IM, DeWitt, and Metcalf methods fail while the other methods can still give the correct vessel position.

## MULTI-BODY FIX.

There are four celestial bodies involved in this case study. Table 3 presents the relevant observation data. In this multi-body fix problem, a correction on the running fix is also applied. González's and Hsu's methods are not applicable in this situation. The experimental results are tabulated in Table 4 and Figure 3. The vessel position obtained using the IM is slightly different due to errors in the graphical procedures. All other methods reach the same vessel position. Metcalf's method and the PSO technique do not require a DR position to find the correct vessel position.

Moreover, in a multi-body celestial fix, to prevent some abnormal observation data from affecting the overall accuracy, a correction on the weights is applied. Those closer to the vessel position are given a larger weight. This point is explained by comparing Figures 4 and 5. The vessel position in Figure 4 is obtained without applying weight correction, and it is at the centre of the big cocked hat. Figure 5 shows the vessel position with corrected weights, in which case it falls into the smaller cocked hat.

Results obtained using PSO show a good agreement with those obtained using other methods. The applicability of PSO on both the two-body fix and multi-body fix is thus verified.

Tab. 3. Multi-body fix data

Course: 220° Speed: 18 kts				
Body	ZT (1993/9/13)	Ho	GHA	Dec
Altair	18-00-00	37° 53.0'	325° 06.6'	08° 51.4' N
Fomalhaut	18-04-00	27° 54.0'	279° 24.2'	29° 39.1' S
Achernar	18-08-00	17° 46.5'	240° 21.7'	57° 15.8' S
Rasalhague	18-12-00	41° 35.5'	002° 04.8'	12° 34.1' N

Tab. 4. Multi-body fix positions

18-12-00 ZT DR : L = 35° S, λ = 005° E			
Method	L	λ	
IM	35° 19.0' S	005° 26.5' E	
DeWitt (5 iterations)	35° 18.6' S	005° 27.0' E	
Metcalf	L = 35° 18.6' S, λ = 005° 27.0' E (DR not required)		
PSO	L = 35° 18.6' S, λ = 005° 27.0' E (DR not required)		

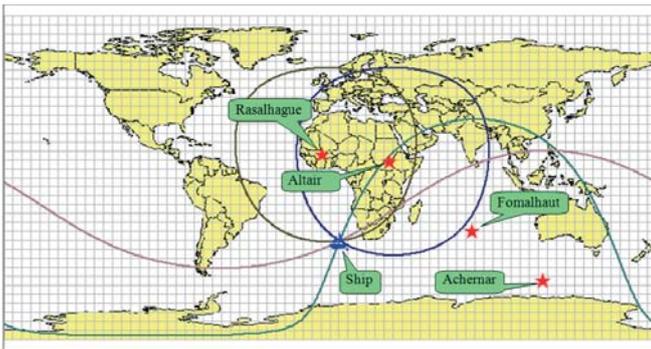


Fig. 3. Results of the multi

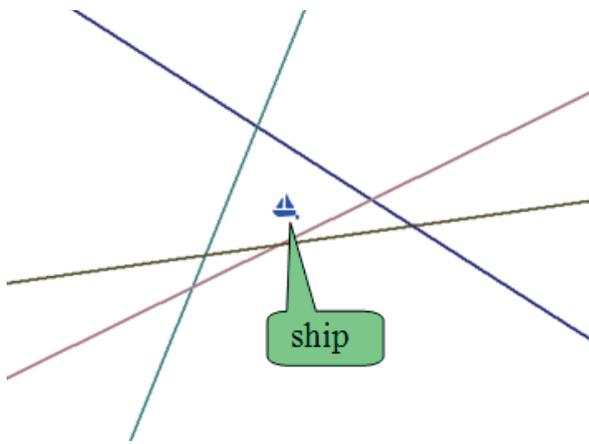


Fig. 4. Vessel position with uncorrected weights

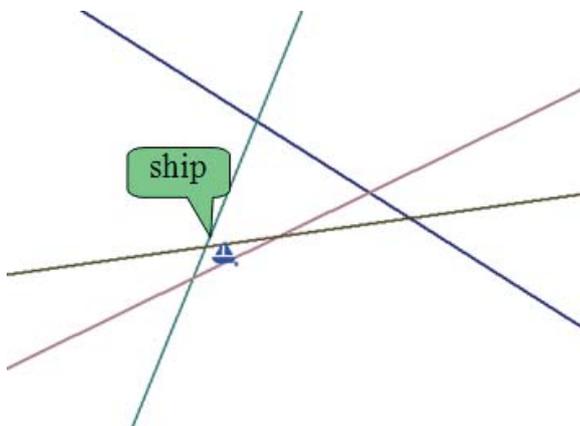


Fig. 5. Vessel position with corrected weights

## HIGH-ALTITUDE OBSERVATION FIX

In this example, a high-altitude observation of the Sun was performed three times before and after its transit. When the Sun is in a low latitude region within a few minutes of the transit, its altitude can reach above 88°, and its azimuth changes remarkably fast. In this situation, the curvature of the LOP makes the IM inappropriate to use. A high-altitude graphical method or a computational method must be used. Therefore, if 2–3 observed altitudes of the Sun before and after the transit time are available, then the vessel position can be obtained from a high-altitude observation of the Sun. Table 5 contains observation data of the Sun's altitude, and experimental results are shown in Table 6.

Tab. 5. High-altitude observations

Course: 290° Speed: 15 kts				
Body	ZT (1996/9/8)	Ho	GHA	Dec
Sun	11-56-13	89° 19.4'	269° 38.9'	05° 34.6' N
Sun	11-58-19	89° 36.2'	270° 10.5'	05° 34.5' N
Sun	12-00-41	89° 19.2'	270° 46.0'	05° 34.5' N

Tab. 6. High-altitude fix positions

12-00-41 ZT DR 1: L = 05° 45' N, λ = 090° E DR 2: L = 05° 33' N, λ = 090° E				
Method	L <sub>1</sub>	λ <sub>1</sub>	L <sub>2</sub>	λ <sub>2</sub>
DeWitt (Iterations)	05° 58.1' N	089° 47.6' E	05° 11.0' N	089° 47.7' W
COP Graphical Method	L = 05° 59.0' N, λ = 089° 47.8' E			
Metcalf	L = 05° 58.1' N, λ = 089° 47.6' E (DR not required)			
PSO	L = 05° 58.1' N, λ = 089° 47.6' E (DR not required)			

The experimental results of using PSO are shown in Figure 6. The COPs from three observations of the Sun give two points of intersection, A and B. The two points are very close, and therefore an inappropriate DR position will lead the algorithm to converge to a non-optimal position. It can be seen from Table 6 that the PSO and Metcalf methods do not require a DR position as reference. As a result, they can find the correct vessel position, or point A.

On the other hand, DeWitt's method is sensitive to the initial DR position. When the DR position is DR 1, the correct vessel position can be obtained. However, when the DR position is DR 2, DeWitt's method finds point B as the vessel position, which is incorrect. Since B and DR 2 are only about 20 NM apart, the algorithm is mistaken in finding the correct vessel position when in fact it has fallen into a local optimum.

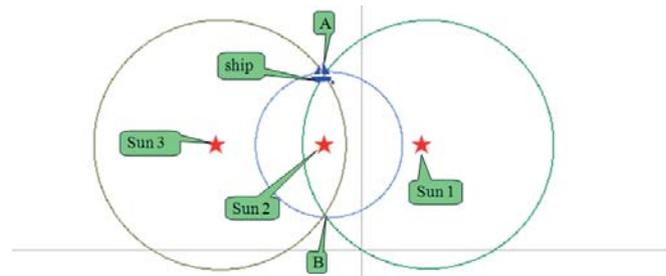


Fig. 6. Results of a high-altitude observation using PSO

## COMPARISON AND DISCUSSION

By comparing the three test cases, it can be determined that, although González's and Hsu's methods are accurate and do not need an AP to solve the problem, they are limited to only a two-body celestial fix. DeWitt's method, Metcalf's method, and the PSO method proposed in this study can perform the fix with any number of celestial bodies and therefore have a wider applicability. Metcalf's method is very convenient, fast, and accurate at solving a multi-body fix and high-altitude observations. However, it becomes inconvenient for solving a two-body fix, in which case a virtual star must be assumed based on the DR position. The accuracy of the positioning is affected by the DR position. It also violates Metcalf's claim that positioning can be performed without needing a DR position as a reference. DeWitt's method has wide applicability and can produce an accurate fix. However, it is still an IM in nature, which means that it has some inherent limitations. The accuracy must be improved by increasing the number of iterations. It is also possible for the algorithm to converge to a non-optimal solution.

The PSO method proposed in this study is applicable in all cases. It is not only accurate and robust, but also relies less on the DR reference position. However, the PSO consumes more computational time than the other methods. Though, a correct vessel position can be obtained within 10 iterations (less than 2 seconds) that can still meet the real time requirement for modern marine navigation.

## CONCLUSIONS

In this work, we utilized Particle Swarm Optimization from the field of swarm intelligence, due to its superior optimization and searching ability, to compute the astronomical vessel position. This technique, used in combination with the GIS and the principle of using circles of position in a celestial navigation fix, allows for a fast, direct, and accurate calculation of a vessel's optimal position. Test results showed that the proposed method not only can be applied in two-body fix, multi-body fix, and high-altitude observation problems, but it is also less reliant on the initial dead reckoning position. In combination with an electronic nautical almanac module, it can be seen as a prototype for integration with an ECDIS. It can serve as an ancillary positioning option and as a useful tool in celestial navigation-related education.

## REFERENCES

1. Chen, C.-L., Hsu, T.-P. & Chang, J.-R. (2003) A Novel Approach to Determine the Astronomical Vessel Position. *Journal of Marine Science and Technology*, 11(4), 221-235.
2. Chiesa, A. and Chiesa, R. (1990) A Mathematical Method of Obtaining an Astronomical Vessel Position. *The Journal of Navigation*, 43, 125-129.
3. DeWitt, C. (1974) Optimal Estimation of a Multi-Star Fix. *NAVIGATION, Journal of The Institute of Navigation*, 21(4), 320-325.
4. Eberhart, R. and Kennedy, J. (1995) A new optimizer using particle swarm theory. *Proceedings of the 6th International Symposium on Micro Machine and Human Science*, Nagoya, Japan, 39-43.
5. Gibson, K. (1994) On the Two-Body Running Fix. *The Journal of Navigation*, 47, 103-107.
6. González, A. R. (2008) Vector Solution for the Intersection of Two Circles of Equal Altitude. *The Journal of Navigation*, 61, 355-365.
7. Hsu, T.-P., Chen, C.-L. & Chang, J.-R. (2005) New Computation Methods for Solving Problems of the Astronomical Vessel Position. *The Journal of Navigation*, 58, 315-335.
8. Kaplan, G. H. (1995) Determining the Position and Motion of a Vessel from Celestial Observations. *NAVIGATION, Journal of The Institute of Navigation*, 42(4), 631-648.
9. Kennedy, J. and Eberhart, R. (1997) A Discrete Binary Version of the Particle Swarm Optimization. *Proceedings of IEEE International Conference on Neural Networks*, Perth, Australia, 4104-4108.
10. Metcalf, T. R. and Metcalf, F. T. (1991) On the Overdetermined Celestial Fix. *NAVIGATION, Journal of The Institute of Navigation*, 38(1), 79-89.
11. Peacock, A. (2011) *Astro Navigation – The Admiralty Manual of Navigation Volume 2* (2011 edition) Para. 0524, Nautical Institute, London, U.K.
12. González, A. R. (2011) Use of Rotation Matrix to Plot A Circle of Equal Altitude. *Journal of Maritime Research*, 8(3),
13. Severance, R. W. (1989) Overdetermined Celestial Fix by Iteration. *NAVIGATION, Journal of The Institute of Navigation*, 36(4), 373-378.
14. Spencer, B. (1990) Astronomical Fixes Without an Assumed Position. *The Journal of Navigation*, 43, 449-451.
15. Tsou, M.-C. (2012) Genetic algorithm for solving celestial navigation fix problems. *Polish Maritime Research*, 19(3), 53-59.
16. Van Allen, J. A. (1981) An Analytical Solution of the Two Star Sight Problem of Celestial Navigation. *NAVIGATION, Journal of The Institute of Navigation*, 28(1), 40-43.
17. Watkins, R.. and Janiczek, P. M. (1978) Sight Reduction with Matrices. *NAVIGATION, Journal of The Institute of Navigation*, 25(4), 447-448.

18. Wu, G. (1991) An Optimal Estimating Method for Celestial Navigation. *The Journal of Navigation*, 44, 266-269.
19. Yallop, B. D. and Hohenkerk, C. Y. (1990) *Compact Data for Navigation and Astronomy for the Years 1991-1995*, Cambridge University Press, Cambridge.

## **CONTACT WITH AUTHOR**

Ming-Cheng Tsou

National Kaohsiung Marine University  
482, Chung-Chou 3rd Rd.  
Chi-Jin District Kaohsiung City  
80543, Taiwan

Email: [d86228006@yahoo.com.tw](mailto:d86228006@yahoo.com.tw)

# NUMERICAL AND ANALYTICAL APPROACHES FOR ROLL MOTION ANALYSIS IN REGULAR LONGITUDINAL WAVES

Emre Peşman<sup>a</sup>

Deniz Can Kolukısa<sup>a</sup>

Metin Taylan<sup>b</sup>

<sup>a</sup>Karadeniz Technical University, 61080, Trabzon, Turkey

<sup>b</sup>Istanbul Technical University, Turkey

## ABSTRACT

*In this study numerical and analytical approaches were investigated in terms of accuracy of their results, practicality of solution and ability to reproduce the main features of the parametric roll phenomenon such as loss of stability and bifurcations in parametric roll motion analysis of ships. In general, single-degree-of-freedom analytical approach is based on reducing number of degrees of freedom from 3 to 1 by using the quasi-static Froude-Krylov assumption, incorporating heave and pitch effects by means of a time varying restoring moment. On the other hand, numerical approaches to motion of six and four degrees of freedom are based on three dimensional diffraction/radiation and potential flow theories. In summary, this paper reveals that analytical approaches are sufficiently adequate to obtain accurate practical results for this relatively complex phenomenon.*

**Keywords:** Parametric roll motion; bifurcations; panel method; AQWA

## INTRODUCTION

The first publications about parametrically excited roll motion appeared in 1930s and 1940s by Watanabe and Kempf [1,2], respectively. Roll motion of a ship in longitudinal waves have been studied by a number of researchers including Graff, Heckscher [3], Kerwin [4], Paulling and Rosenberg [5]. The first experimental observation of parametric roll was done by Paulling et al in San-Francisco Bay [6]. Although its theoretical existence has been known for a long time, parametric roll attracted a great deal of interest in recent years because of the incidents that resulted in damages [7,8]. Results of studies about Post-Panamax C11 class containership which sustained extensive loss and damage to deck-stowed containers in October 1998 showed that Post-Panamax containerships tend to experience parametric roll motion in extreme weather [7]. These casualties led designers, researchers and regulatory authorities to initiate further research and investigations. Among the researchers Spyrou [9], Neves, Rodrigues [10]

and Bulian *et al.* [11] focused attention on nonlinear aspects and effects of changing tuning factors on parametric roll motion. In addition, some researchers focused attention on probabilistic properties of parametric roll [12,13,14,15,16]. The state of the art in development of methodology and regulations for assessment of ship intact stability, can be found in [17].

Ships sailing in longitudinal waves are excited by the hydrostatic and hydrodynamic forces. Differing the geometry of ship's wetted hull form with respect to wave crest position has an important role on roll motion. In regular waves, the excitation is periodical with a finite period and certain ratios of encounter and natural frequencies which may lead to loss of stability. The most dangerous situation usually occurs in the first parametric resonance region in which wave length is approximately equal to the ship length at an encounter frequency that is twice greater than the roll natural frequency. In this case, the variation of restoring moment causes the

roll angle to increase drastically unless other factors such as damping come into play. This particular state is called the parametric roll resonance phenomenon. A detailed explanation of the physics behind parametric roll has been given in [12] and assessment of parametric roll resonance in design of container carriers has been presented in the guide of American Bureau of Shipping [13].

In the literature, parametric roll resonance phenomenon has been investigated with the use of both numerical and analytical approaches. Both the solution methods have their own advantages depending on a purpose of application. Numerical approaches are suitable for solving coupled motions in time domain. Neves and Rodriguez used a two-dimensional analysis for a set of equations for coupled heave, pitch and roll motion with 2nd and 3rd order nonlinearities describing the restoring action [18]. Levadou and Van't Veer used coupled nonlinear equation of motion in time domain with 3 and 5 degrees of freedom (DOF) [19]. The nonlinear excitations are incorporated by pressure integration over the actual wetted surface while diffraction forces are considered linear. Hydrodynamic effects are calculated in the frequency domain by a 3D panel code and are incorporated in the time domain by adopting the impulse response function method. France et al. [7] and Shin et al. [12] used a similar approach with a hybrid singularity based on Rankine source in the near field and transient Green's function in the far field.

On the other hand, analytical approaches have some practical advantages such as possibility of determining roll amplitudes and bifurcations above the stability threshold. An alternative simplified approach is possible for reducing number of DOF from 3 to 1 with the quasi-static Froude-Krylov assumption [20,21]. Based on these assumptions, both direct and indirect effects of waves may be evaluated. It is believed that the quasi-static approximation may be a good model which is able to assess the main features of the phenomenon such as loss of stability and bifurcations [21].

In this paper, both numerical and analytical approaches were applied to predict large roll amplitudes in time and frequency domain. 6 -DOF and 4- DOF time domain simulations based on 3D panel code and single- DOF time domain simulation considering heave and pitch effects by means of time varying restoring moment were used for numerical DOF approaches. 6 DOF and 4 DOF time domain simulation approaches were applied by using AQWA-LINE and AQWA-NAUT software which are capable of calculating motions by using 3D diffraction/radiation and potential flow theory. The single - DOF time and frequency domain models based on reducing number of DOF from 3 to 1 by using the quasi-static Froude-Krylov assumption [20,21,22] were used for analytical approaches. In the models, heave and pitch effects were considered by means of time varying restoring moment [20,21,22]. Basically, three analytical approaches were used in the present study. The first is Bulian's PolyFour Model, the other one is the simplified model which may be considered a simplified version of the first model [21,22], and the other one is linear model [23]. The accuracy of numerical and analytical approaches were evaluated by comparing

experimental results obtained for the sample ship form [21].

## DESCRIPTION OF MODELS

ANSYS AQWA program suite was used for 6 -DOF, 5 -DOF (surge, sway, heave, pitch and roll) and 4 - DOF (surge, heave, pitch and roll) analyses. AQWA-NAUT uses frequency - domain results for added mass and damping coefficients to obtain time-history analysis results of ship motions in regular waves by calculating hydrodynamic forces acting on the ship at each time step. AQWA-LINE program module calculates harmonic ship motion amplitudes and hydrodynamic coefficients in frequency domain. The program calculates hydrodynamic forces acting on the ship by integrating hydrodynamic pressure over the wetted surface of the ship. It calculates Froude-Krylov forces due to unscattered - incident wave potential, diffraction forces due to diffracted wave potential and radiation forces due to the resultant potential of 6 - DOF motion of the ship hull. Furthermore, the potentials are handled by a boundary surface method, distributing pulsating sources on the ship wetted surface by assuming a mean equilibrium position in the calm water by means of Green's theorem [24].

In analytical approaches, models should be simplified to obtain steady state solution in frequency domain. The effect of non-hydrostatic pressure can usually be neglected for regular waves longer than half of the ship's length. Therefore, non-hydrostatic pressure referred as Smith effect is neglected. Sway and yaw motions are neglected by considering the capability of the ship to keep her course. It is assumed that the ship is able to maintain a constant speed, accordingly surge motion is neglected. On these assumptions, only roll motion is taken into account in the present analysis. The influence of roll on pitch and heave could be modeled as an explicit forcing, whereas the influence of heave and pitch on roll can be modeled as a parametric excitation by utilizing quasi-static Froude-Krylov assumption [20,21]. The essence of quasi-static assumption is balancing heave and pitch statically in a wave. In this study, non-hydrostatic effects under a wave profile are neglected and restoring moment curves are computed with standard hydrostatic calculation software. Quasi-static assumption is carried out in order to solve the model analytically without unnecessary cumbersome calculations. A detailed explanation of the procedure may be found in [21,22].

Assuming that the displacement of the ship is constant, roll motion equation is written in the following form:

$$(I_{xx} + \delta I_{xx})\ddot{\phi} + B(\dot{\phi}, \phi) + \Delta GZ(\phi, t) = 0 \quad (1)$$

where  $(I_{xx} + \delta I_{xx})$  is moment of inertia,  $\phi$  is roll angle,  $B(\dot{\phi}, \phi)$  is damping function and  $\Delta GZ(\phi, t)$  is restoring function. Eq. (1) may be re-written as;

$$\ddot{\phi} + b(\dot{\phi}, \phi) + \frac{\omega_0^2}{GM_0} GZ(\phi, t) = 0 \quad (2)$$

In Eq. (2),  $\Delta GZ(\phi, t)$  may be approximated by different

expressions. Bulian used Eq. (3) to identify the GZ surface analytically [21].

$$GZ(\phi, t) = \sum_{j=1,3,5,7,9} \left( Q_{j0} + \sum_{n=1}^{N_h} Q_{jn}^c \cos\left(\frac{n\omega_e t}{\cos(\chi)}\right) + Q_{jn}^s \sin\left(\frac{n\omega_e t}{\cos(\chi)}\right) \right) \phi^j \quad (3)$$

$$\left\{ \begin{array}{l} Q_{j0} = A_{j0} \\ Q_{jn}^c = A_{jn}^c \cos(n\psi_{c0}) + A_{jn}^s \sin(n\psi_{c0}) \\ Q_{jn}^s = -A_{jn}^c \sin(n\psi_{c0}) + A_{jn}^s \cos(n\psi_{c0}) \\ \psi_{c0} = 2\pi \frac{x_{c0}}{\lambda_w} \end{array} \right. \quad (4)$$

In the above-mentioned study,  $A_j(x_c)$  values are polynomial coefficients for each wave crest position and can be expressed as Fourier series in the variable  $x_c$  with main period equal to the wave length  $\lambda_w$ .  $N_h$  is the maximum number of harmonic components that can be estimated from polynomial coefficients by using Nyquist sampling theorem [25]. Coefficients of the Fourier series related to  $A_j(x_c)$  are calculated as follows:

$$\left\{ \begin{array}{l} Q_{j0} = A_{j0} \\ Q_{jn}^c = A_{jn}^c \cos(n\psi_{c0}) + A_{jn}^s \sin(n\psi_{c0}) \\ Q_{jn}^s = -A_{jn}^c \sin(n\psi_{c0}) + A_{jn}^s \cos(n\psi_{c0}) \\ \psi_{c0} = 2\pi \frac{x_{c0}}{\lambda_w} \end{array} \right. \quad (5)$$

In the simplified model, unlike the aforementioned model, GZ surface is approximated more simply by using only polynomial coefficients of wave-crest and wave-trough GZ curves. Furthermore, the values between aft and fore of a ship are approximated by a sinusoidal function.

$$GZ(\phi, t) = \sum_{n=1}^N (m_{2n-1} + k_{2n-1} \cos(\omega_e t)) \phi^{2n-1} \quad (6)$$

Coefficients “ $m$ ” and “ $k$ ” in Eq. (6) are obtained from restoring moment curves in wave crest and wave trough conditions.

$$m_{2n-1} = \frac{c_{2n-1, trough} + c_{2n-1, crest}}{2} \quad (7)$$

$$k_{2n-1} = \frac{c_{2n-1, trough} - c_{2n-1, crest}}{2} \quad (8)$$

$$b(\phi, \dot{\phi}) = 2\mu\dot{\phi} + \beta\dot{\phi}|\dot{\phi}| + d\dot{\phi}^3 \quad (9)$$

In Eq. (7) and (8), “ $c_{2n-1, crest}$  and  $c_{2n-1, trough}$ ” show the coefficients of polynomials fitted to the restoring moment curves in wave-trough and wave-crest conditions. 7<sup>th</sup> degree polynomials are used for developing the restoring moment surfaces. In this study, quadratic damping and linear damping terms are used.

Steady state solution of the equation of roll motion in longitudinal waves was carried out by means of the averaging method in order to determine bifurcations in frequency domain. The averaging method was introduced by Bulian [21] to estimate roll amplitudes in frequency domain. Bulian’s PolyFour model and simplified model are solved by using the same procedure as that applied in [21]. Time domain solution of the simplified model was carried out by using the 4.5 order Runge-Kutta (Dormand-Prince) method.

1- DOF linear model is based on linear Mathieu-Hill equation and its solution Ince-Strutt diagram [23]. Resonant frequencies can be estimated easily by using Ince-Strutt diagram. But predicting roll amplitudes at these frequencies is not possible. Stability of pitchfork bifurcations at minimum and maximum frequencies, in other words, ship speeds can be predicted by using a method detailed in [22].

## SAMPLE SHIP AND ENVIRONMENTAL CONDITIONS

The sample ship used throughout the analysis is of a frigate form whose experimental tests were carried out at the towing tank of DINMA [21]. The experiments were carried out for 3 DOF (heave, pitch and roll) [21]. The sample ship has no bilge keels and appendages. The form and main characteristics of the sample ship, named F1, is given in Fig. 1 and Tab. 1, respectively.

Tab.1. Main characteristics of the sample ship

	Type	LBP	B	T	KG
F1	Frigate	120.00 m	14.25 m	4.060 m	6.557 m

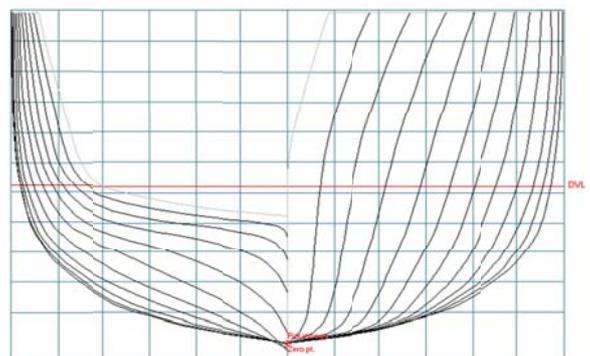


Fig. 1: Frame body plan of the sample ship

In this study, wind force was neglected in order to create results which are comparable with the experimental ones. AQWA mesh model and pressure contours of the sample ship are presented in Fig. 2.

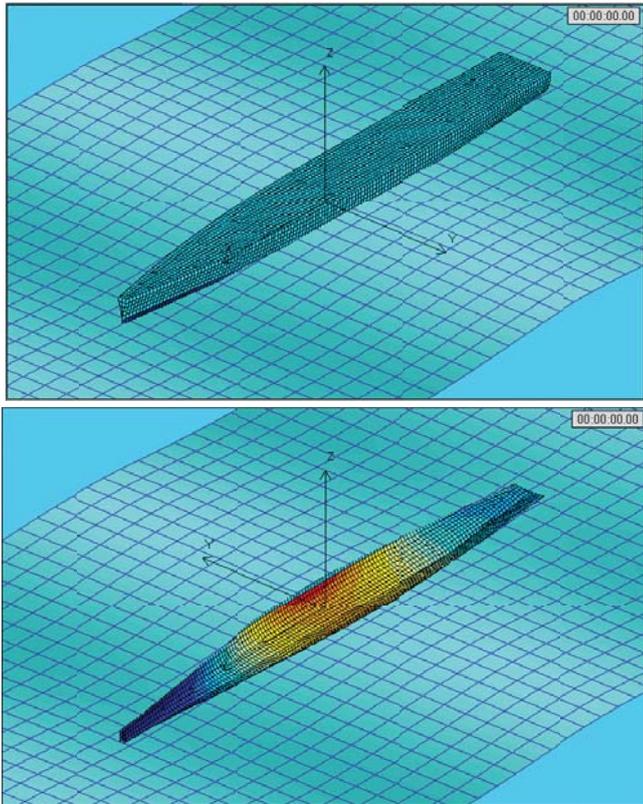


Fig. 2 AQWA mesh model and pressure contours of the sample ship.

The wave length and wave height were chosen equal to 120 m and 2.4 m, respectively. In AQWA program suitable thrust forces were chosen to include ship speed and altering wave effect. Non-dimensional linear damping coefficient is chosen equal to 0.212. Bulian's values with respect to Froude number given in Fig. 3 and 4 are used for nonlinear damping coefficient.

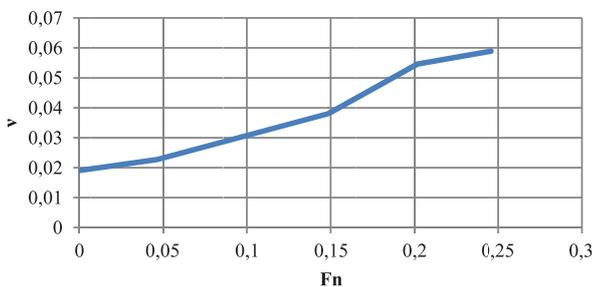


Fig. 3 Linear part of the damping term in Eq. (9)

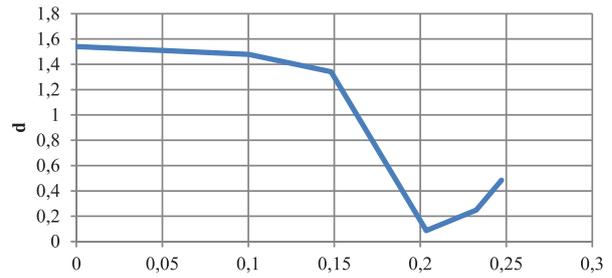


Fig. 4 Nonlinear part of the damping term in Eq. (9)

## RESULTS OF THE ANALYSES

As explained above, the maximum roll amplitudes with respect to ship speed were determined by using different approaches. Roll amplitudes over ship speeds were simulated in time domain by using AQWA-NAUT™ for various thrust forces. Sway and yaw motions are neglected in 4-DOF model by considering course keeping capability of ship. Maximum roll amplitudes and mean ship speeds obtained by means of AQWA-NAUT™ (4, 5 and 6 DOF) were plotted in Fig. 5 in comparison with the experimental results obtained in DINMA [21]. It can be perceived that the results of AQWA-NAUT (4 DOF) and (5 DOF) are in good agreement with the experimental results in general, however, there is no results of AQWA-NAUT (6 DOF) where the ship has no course keeping capability at speeds higher than 3 m/s. It is observed that periodic variation of ship speed has no unfavorable effect on results.

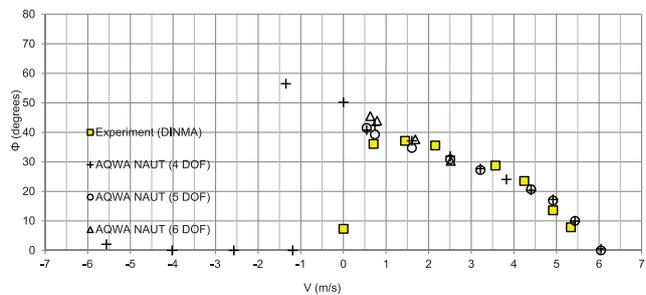


Fig. 5 Comparison of AQWA-NAUT results with the experimental results

Comparison of the analytical results with the experimental results was given in Fig. 6. Roll amplitudes from Bulian's model are approximately by 10° lower than the experimental and AQWA-NAUT results (Fig. 6). However, this fact does not suggest strong conclusive evidence that they may be disregarded in preliminary design. The simplified model was solved for both linear and nonlinear damping conditions (Fig. 6). The amplitudes acquired from the simplified model for nonlinear damping are approximately by 5° lower than the experimental results. However, when linear damping is used, the results of the simplified model are in good

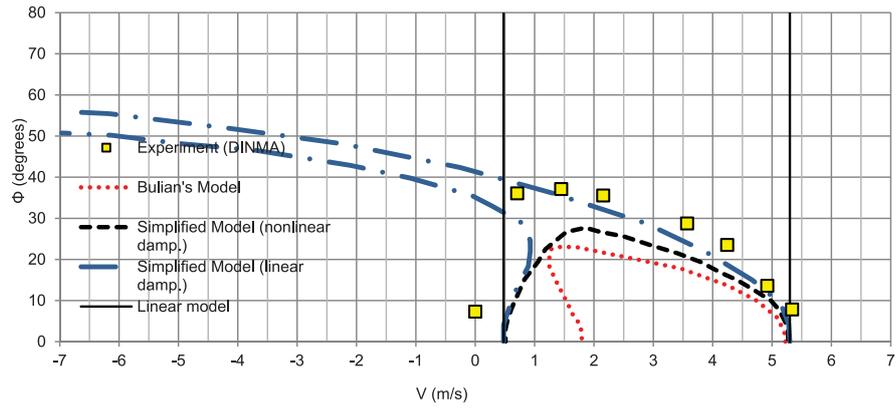


Fig. 6 Comparison of approximation method results and experimental results

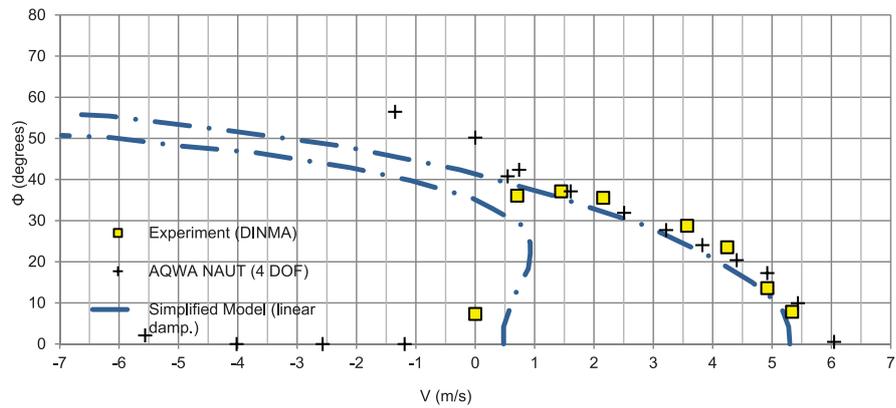


Fig. 7 Comparison of the simplified model (linear damping) and AQWA-NAUT (4 DOF) results with the experimental results

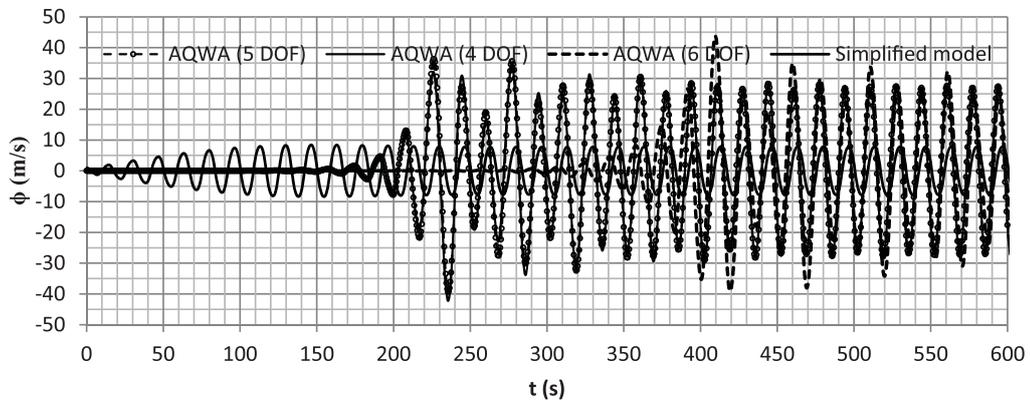


Fig. 8 Variation of roll angles in time domain at neighborhood of 0.7 m/s ship speed

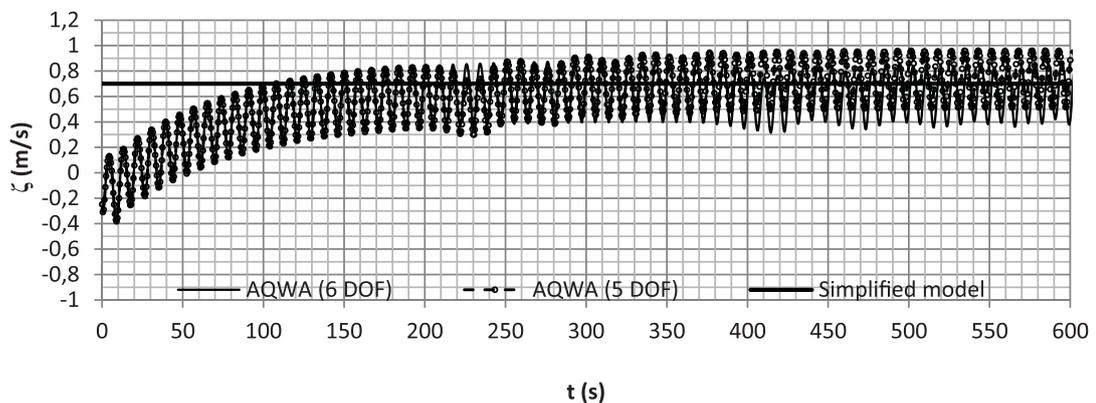


Fig. 9 Variation of ship speed in time domain

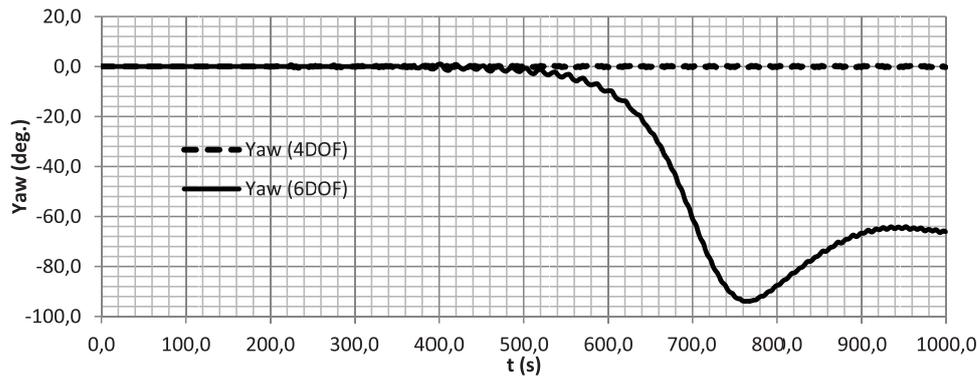


Fig. 10 Variation of yaw motion in time domain

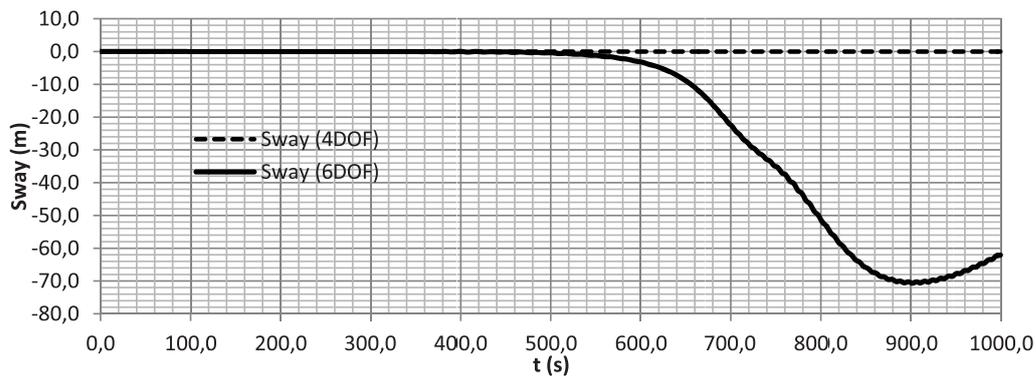


Fig. 11 Variation of sway motion in time domain

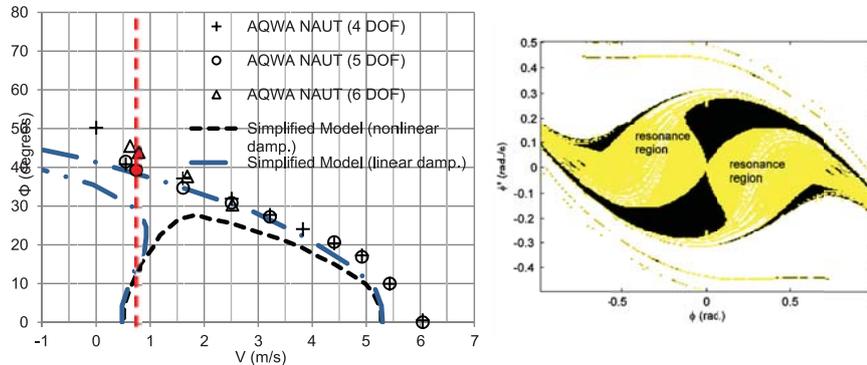


Fig. 12 Results obtained from AQWA-NAUT and the simplified model in frequency domain , and the resonance region diagram with respect to initial values

agreement with the experimental results. Results of linear model (Mathieu-Hill Eq.) are also a stable trivial solution of the simplified model. Unstable trivial solution means a boundary of speeds in which parametric roll motion starts (minimum and maximum ship speeds).

Results of AQWA-NAUT (4 DOF) and simplified model with linear damping were compared in Fig. 7. It can be perceived that results are in good agreement in a region where fold bifurcation occurs. Amplitudes obtained from AQWA-NAUT (4 DOF) were approximately by 10° higher than the amplitudes resulting from the simplified model (linear damping) at following waves. The experiments were made for only head waves therefore fold bifurcation could not be proved experimentally.

The time domain simulation procedure applied in 4- DOF approach is repeated in 5- and 6- DOF calculations with the use of AQWA-NAUT in a specific condition. The results of AQWA-NAUT calculations given in Fig. 8 are compared with the results of analytical simplified model in time domain. Variation of ship speed is presented in Fig. 9. As it can be observed from Fig. 8, the results of AQWA-NAUT calculations with 4 and 5 DOF are compatible with each other but results of calculations with 6 DOF differ a little from 4 -DOF and 5 -DOF results due to the degrees of freedom. Results of sway and yaw motion in 6- DOF simulations are shown in Fig. 10 and Fig. 11. Results of the analytical model in time domain are obtained by solving the equation numerically. It is observed that maximum roll amplitudes do not exceed 15° unlike

AQWA-NAUT results since there is fold bifurcation. The fold bifurcation means that there are three different solutions and the obtained result is the smallest one due to initial values, as shown in Fig. 12. Resonant and non-resonant regions are also plotted with respect to initial values in Fig. 12.

## CONCLUSIONS

In this study, results obtained from numerical and analytical approaches were compared with experimental results for parametric roll evaluation of a sample ship. Selection of the solution method has a great influence on setting up the model. A ship sailing in longitudinal regular waves experiences 6 - DOF motions under wind and wave effects. Therefore, the model has to be set up by considering 6 - DOF motion to obtain more accurate results. For this reason, AQWA-NAUT™ program was used for the numerical analysis. The program was applied in three modes, namely: of 6 DOF, 5 DOF and 4 DOF. The results of analyses with 6 DOF, 5 DOF and 4 DOF are highly complying with the experimental results. It may be concluded that 5-DOF and 4- DOF time domain simulations are good tools for determining roll amplitudes. It should be noted that the results of 6 - DOF simulations naturally differ from 4-DOF results because of the presence of sway and yaw motions. Oscillation of ship speed is included as a second excitation term in restoring moment for roll motion in longitudinal waves, as expressed in Eq. (10). In Eq. (10),  $V(t)$  can be assumed a sinusoidal function (Fig.9), hence it can be said that excitation is doubled depending on amplitude of  $V(t)$ . However results of this study revealed that oscillation of ship speed has no unfavourable effect (Fig.7). Comparison of constant speed and oscillating speed indicates that amplitudes of about 0.3 m/s in value are not adequate to change results.

$$GZ(\phi, t) = \sum_{n=1}^N \left( m_{2n-1} + k_{2n-1} \cos \left( \left( \omega_w + \frac{\omega_w^2}{g} V(t) \right) t \right) \right) \phi^{2n-1} \quad (10)$$

Numerical simulations usually start with a particular initial condition that lies in a specified domain of attraction. Thus, co-existence of any other steady state solutions cannot be estimated without changing the initial condition. Numerical solutions are somewhat inadequate to give a global picture of the response curve and bifurcations involved in the phenomenon. However, an analytical approach is usually able to give such an overall view in a more practical way. In addition, analytical approaches have some practical advantages such as determining roll amplitudes and also bifurcations above the stability threshold. It is strongly believed that the quasi-static approximation maybe a good model producing the main features of the phenomenon such as loss of stability and bifurcations [21,22]. In this study, two analytical approaches were used: Bulian's PolyFour and simplified models [21,22] in which nonlinear damping term was applied. It is found that the results of Bulian's PolyFour and simplified models

containing nonlinear damping term are in good agreement, comparatively. However, they are not in a close agreement with the experimental results: PolyFour model result is approximately by 10° lower and the simplified model result is approximately by 5° lower than the experimental results. The simplified model was also run with linear damping term. As seen from Fig. 4, the result of the simplified model with linear damping term complies satisfactorily with AQWA-NAUT and experimental results. Damping coefficients have great effect on results, but unfortunately they cannot be predicted accurately without full-scale experiments. Ships designed for high speed have lower block coefficients and these ship forms tend to encounter parametric roll motion more than others. In this study, a frigate ship form of block coefficient equal to 0.453, was chosen. Behaviour of other ship forms can be found in [21] and [22]. The most important factor of parametric roll motion is restoring moment term. The shape of wave crest-wave trough restoring moment curves and roll responses of container ships and frigates are almost the same as indicated in [22].

Finally, this study indicates that using the linear damping term gives comparable results and the nonlinear damping term is not demanded. The most important advantage of the analytical approach is ability of determining types of bifurcation. It is critically important to determine the type of bifurcations since they lead to jump phenomenon. It should be also stressed that subcritical bifurcations may lead to large roll amplitudes or capsizing [26,22].

## BIBLIOGRAPHY

1. Watanabe Y.: *On the dynamic properties of the transverse instability of a ship due to pitching*. J. Soc. Nav. Archit. Jpn 53, 1934, pp. 51–70
2. Kempf G.: *Die Stabilität Beanspruchung der Schiffe Durch Wellen und Schwingungen*. Werft Reederei Hafen 19, 1938, pp. 200–202
3. Graff W, Heckscher E.: *Widerstand und Stabilität Versuche mit Drei Fischdampfer Modellen*. Werft Reederei Hafen 22, 1941, pp.115–120
4. Kerwin J.E. : *Note on rolling in longitudinal waves*. Int. Shipbuild. Prog. 2(16),1955, pp. 597–614
5. Paulling J.R., Rosenberg R.M.: *On unstable ship motions resulting from nonlinear coupling*. J. Ship Res. 3, 1959, pp. 36–46
6. Paulling J.R., Kastner S., Schaffran S.: *Experimental Studies of capsizing of intact ships in heavy seas*. U.S. Coast Guard Technical Report,1972 (also IMO Doc. STAB/7, 1973)
7. France W.N., Levaduo M., Treakle T.W., Paulling J.R., Michel R.K., Moore C.: *An investigation of head-sea parametric rolling and its influence on container lashing systems*. Mar. Technol. 40(1), 2003, pp. 1–19

8. BSU: *Fatal accident on board the CMV CHICAGO EXPRESS during typhoon "HAGUPIT" on 24 September 2008 of the coast of Hong Kong*. Bundesstelle für Seeunfalluntersuchung, Federal Bureau of Maritime Casualty Investigation Report 510/08, 2009
9. Spyrou K.J.: *Designing against parametric instability in following seas*. Ocean Eng. 27, 2000, pp. 625–653
10. Neves MAS., Rodriguez C.A.: *Influence of non-linearities on the limits of stability of ships rolling in head seas*. Ocean Eng. 34, 2006, pp. 1618–1630
11. Bulian G., Francescutto A., Lugni C.: *On the nonlinear modeling of parametric rolling in regular and irregular waves*. Int. Shipbuild. Prog. 51, 2004, pp. 205–220
12. Shin Y.S., Belenky V.L., Paulling J.R., Weems K.M., Lin W.M.: *Criteria for parametric roll of large container ships in longitudinal seas*. SNAME Trans. 112, 2004, pp. 14–47
13. ABS : *Guide for the assessment of parametric roll resonance in the design of container carriers*. American Bureau of Shipping, Houston, 2004 (as amended 2008)
14. Belenky V.L.: *On risk evaluation at extreme seas*. In: Proceedings of the 7<sup>th</sup> international stability workshop, Shanghai, China 2004, pp. 188–202
15. Hashimoto H., Umeda N., Matsuda A.: *Experimental and numerical study on parametric roll of a Post-Panamax container ship in irregular wave*. In: Proceedings of STAB'06 9th international conference on stability of ships and ocean vehicles, Rio de Janeiro, Brazil, 2006, pp. 181–190
16. Bulian G., Francescutto A., Lugni C.: *Theoretical, numerical and experimental study on the problem of ergodicity and 'practical ergodicity' with an application to parametric roll in longitudinal long crested irregular sea*. Ocean Eng. 33, 2006, pp. 1007–1043
17. Francescutto A.: *Intact stability of ships - recent developments and trends*. In: Proceedings of 10th international symposium on practical design of ships and other floating structures PRADS'07, Vol. 1, Houston, 2007, pp. 487–496
18. Neves MAS., Rodriguez C.A. : *A coupled third order model of roll parametric resonance*. In: Proceedings of the maritime transportation and exploitation of ocean and coastal resources, London, 2005, pp. 243–253
19. Levadou M., Van't Veer R.: *Parametric roll and ship design*. In: Proceedings of the ninth international conference on stability of ships and ocean vehicles STAB'06, Vol. 1, 2006, pp. 191–206
20. Bulian G.: *Approximate analytical response curve for a parametrically excited highly nonlinear 1-DOF system with an application to ship roll motion prediction*. Nonlinear Anal. Real World Appl. 5(4), 2004, pp. 725–748
21. Bulian G.: *Development of analytical nonlinear models for parametric roll and hydrostatic restoring variations in regular and irregular waves*. Ph.D. thesis, University of Trieste, Trieste, 2006
22. Pesman E., Taylan M.: *Influence of varying restoring moment curve on parametric roll motion of ships in regular longitudinal waves*. J. Mar. Sci. Tech. 17, 2012, pp. 511-522
23. Mathieu E.: *Mémoires sur le mouvement vibratoire d'une membrane de forme elliptique*. Journal des Mathématiques Pures et Appliquées 13, 1868, pp. 137-203
24. Garrison C.J.: *Hydrodynamic loading of large offshore structures: Three dimensional source distribution methods*. Numerical methods in offshore engineering, John Wiley, 1978, pp. 87-140
25. Nyquist H.: *Certain topics in telegraph transmission theory*. Trans AIEE, 47, 1928, pp. 617–644 (reprint as classic paper in: Proc. IEEE, Vol. 90, No. 2, Feb. 2002)
26. Thompson J.M.T., Rainey R.C., Soliman M.S.: *Ship stability criteria based on chaotic transients from incursive fractals*. Philos. Trans. R. Soc. 332 London, 1990, pp.14

#### CONTACT WITH THE AUTHOR

Emre Peşman

Karadeniz Technical University  
Faculty of Marine Science

Deniz Can Kolukısa

Karadeniz Technical University  
Faculty of Marine Science

Metin Taylan

Istanbul Technical University  
Faculty of Naval Architecture and Ocean Engineering

**TURKEY**

## USE OF SATELLITE DATA IN MONITORING OF HYDROPHYSICAL PARAMETERS OF THE BALTIC SEA ENVIRONMENT

Adam Krężel, Ph. D.  
Katarzyna Bradtke, Ph. D.  
Agnieszka Herman, Assoc. Prof.  
University of Gdańsk, Poland

### ABSTRACT

*Intensive development of infrastructure for fast processing of outsized amount of space-borne data enables now to use the satellite data for operational controlling the state of its environment. In our presentation we show some examples of analysis of processes in marine environment which are possible due to satellite data and algorithms of its processing developed in SatBaltic Project. It concerns supporting of modelling of solar energy inflow to the sea with space-borne input data, identification and analysis of sea ice cover, supporting of oil spill detection, and identification of phenomena which modify spatial distribution of the sea surface temperature.*

**Keywords:** satellite remote sensing, Baltic, SST, solar energy, sea ice, oil spill

### INTRODUCTION

Systematic observations of the Earth environment from space started in the late seventies of the 20th century. Due to these data our knowledge about some environmental processes and phenomena has increased rapidly. It concerns both global processes and, in not smaller degree, local ones like, for instance, those occurring in the Baltic Sea. Fast development of infrastructure for efficient processing of the extensive amounts of data enables now to use the space-borne information for continuous, operational monitoring of the state of the environment.

Effective use of satellite data comprises two components. On the one hand, it requires new, complex algorithms that allow converting the electromagnetic signal registered by sensors onboard of the satellite to physically meaningful quantities characterizing the state of the environment. On the other hand, these new, specific information has to be incorporated into the analysis, which often had been limited or even impossible with the use of traditional sources of data. These were the assumptions on the basis of which the SatBaltic Operational System has been set up. Its main goal is to accomplish, in a possibly wide range, the above-indicated activities in relation to the Baltic Sea. In this work we show some selected examples of the application of the satellite data to the analysis of physical processes in marine environment, based on the algorithms created within the SatBaltic Project at the Institute of Oceanography, University of Gdańsk. In Section 1 we briefly present examples of particularly challenging and important groups of processes analyzed in the project: detection of clouds, oil spills and sea ice. The analysis of solar radiation and sea surface temperature (SST),

shown in Section 2, demonstrates how the data obtained as a result of our algorithms can be applied in oceanographic research.

The presentation of the algorithms begins with those devoted to cloudiness. Among the components of the terrestrial atmosphere, cloudiness is the one that, with its high temporal and spatial variability, influences the solar irradiance at the sea surface to the highest degree. The solar irradiance at a given place depends on whether the Sun is obstructed by clouds totally, partly or not at all, and in what degree the clouds are transparent to the light. The related challenges lie not only in reliable estimation of the influence of clouds on the amount of solar radiation reaching the sea surface, but also in detecting totally cloudless regions, in which it is possible to determine such variables as SST or the surface chlorophyll concentration. Even small cloud-masking errors can result in serious errors in the analysis of these quantities.

Huge oil spills from oil platforms or tankers, being result of accidents, can seriously affect marine and coastal environment, but they are rare and relatively easy to detect. In the Baltic Sea, the main source of oil pollutants is illegal dumping of oily wastes from ships. Regular, small-scale spills, contribute to high overall amount of the hydrocarbons released to the environment, whose negative impact cannot be diminished [3]. They require frequent observation of large areas in order to identify as many events as possible. For this reason, the routine aerial surveillance nowadays is usually supported by the satellite SAR (Synthetic Aperture Radar) data.

The third of the presented algorithm examples concerns sea ice, the monitoring of which has practical importance for marine transport and fishery, as well as weather forecasting, as sea ice is an important component of winter climate of the large parts of the Baltic Sea region. Experiments with a mesoscale Weather Research and Forecasting (WRF) model, performed in the SatBaltic project, suggest that the quality of sea ice data has significant influence on the simulated weather not only over the sea, but also the surrounding land.

## DEVELOPMENT OF ALGORITHMS FOR SPACE-BORNE DATA PROCESSING

### CLOUDINESS

During daytime, the simplest method of observation of cloudiness from space is based on the visible part of the spectrum where white clouds are easy noticeable on the background of a dark sea. There are several radiometers which can observe the Earth's colour from space [1] [10], but because of the high temporal dynamics of this phenomenon it is important to choose the source that provides data as frequently as possible – preferably devices working on geostationary satellites. The area of the Baltic Sea is located in the field of view of SEVIRI (Spinning Enhanced Visible and InfraRed Imager) working on board of Eumetsat's Meteosat Second Generation (MSG) satellite. It observes the Earth in 12 spectral channels and produces its image every 15 minutes. Basing on these data, it is possible to calculate "cloud fraction" that corresponds with the meteorological cloudiness, and "cloud transmission" which characterises the amount of solar radiation transmitted by clouds. Details of the algorithm are presented in [6]. In practice, the only alternative to such information is atmospheric modelling, but as one can see in Fig. 1, very often a good agreement in results at one time instance is followed by a total inconsistency at another. Independent measurements of solar radiation do not leave doubt which data are closer to reality.

As described in Section 2.1, SEVIRI provides the first, crucial component to the set of input data necessary to operate the solar energy transmission model. It enables calculation of the most important parameter which influences the majority of processes in marine environment.

### SUPPORT OF OIL SPILL DETECTION

Oil spills may be detected in the SAR-C images because of the damping effect of the oil on the backscattered signals from the radar instrument. This local decrease of backscattered energy appears as a dark spot in the image.

SAR can collect data independently of weather and light conditions, what makes it an excellent tool to monitor illegal spillages. Before ENVISAT, European Space Agency's satellite which operated from 2002 to 2012, access to SAR data was limited due to their high costs. Currently, a similar instrument operates on the Sentinel 1 satellite.

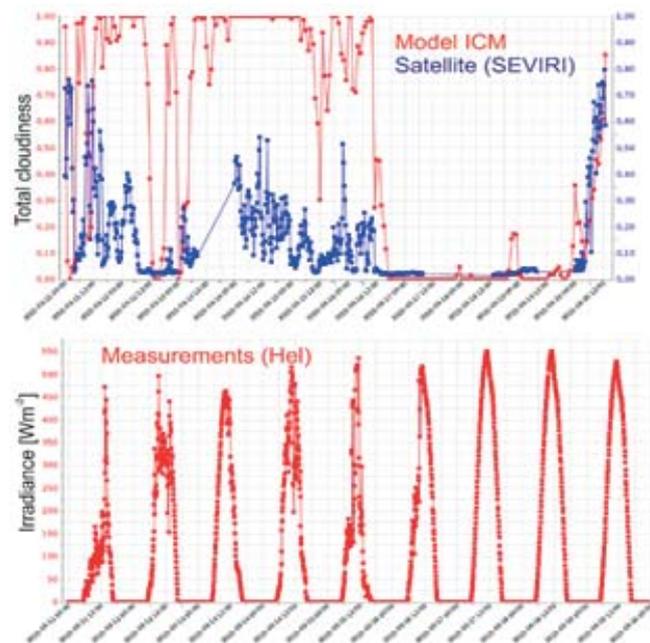


Fig. 1. Example nine days of variability of cloudiness observed from space and modelled (top) and solar irradiance measured (bottom) in the vicinity of Hel

The problem with interpretation of SAR data in the context of the oil pollution monitoring is the fact that oil spills are not the only phenomena causing dark spots in the images. Amongst phenomena causing misleading look alikes are also: grease ice, intensive algal blooms, windless areas known as "calm areas", or upwelling regions. To distinguish oil slicks from their look alikes, knowledge about their specific features and environmental conditions has to be taken into account. Thus, manual inspection of SAR images, by highly-qualified operators, is still the most popular technique. Operators use a software that can calculate some spot attributes, but they still have to go through the whole image manually. This is time consuming and many ideas for automation of this process have been proposed over time. Our team contribute to solve this problem [5].

The goal of our work, undertaken as a part of the SatBaltic Project, was to develop a methodology that would improve inspection of the SAR images. Particular emphasis was put on the adaptation to the specificity of the seas, like Baltic Sea, where small, elongated spillages take place due to shipping activities and their location is difficult to predict.

We proposed an object-oriented approach to the radar image interpretation using the ENVISAT ASAR (Advanced SAR) images and eCognition software for image interpretation (Fig. 2). The algorithm consist of four main steps: pre-processing of the data, hierarchical segmentation, dark spot recognition and their classification.

The algorithm utilizes a multi-resolution segmentation technique to obtain a hierarchy of objects at various scales. Such approach to the segmentation process enables to detect oil slicks, which vary in scales, and also to put dark objects in context of their surroundings defined by larger regions. The classification scheme is based on a 'decision tree' logic, which aims to eliminate look-alikes rather than classify all dark spots

into categories. A number of aspects is taken into account in the classification process, e.g., the shape characteristics and contrast to the surroundings, or position relative to the land, major shipping routes and the neighbouring objects on the sea surface.

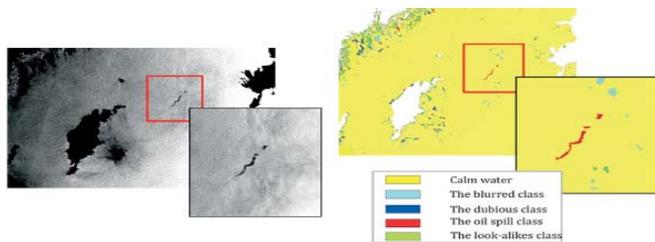


Fig. 2. Sample ASAR image with the confirmed oil spill (indicated in frames) and result of the algorithm – a classified image (source of data: ESA, ASAR WSM, VV, May 7, 2008)

The algorithm seems to be very reliable. Its performance was tested on a group of oil spills recorded by HELCOM. After automatic extraction of all dark spots from ASAR images and elimination of over 95 % of them as look-alikes, over 95% of known oil spills has been successfully identified. We plan to improve and adapt this algorithm for Sentinel-1 data, and to run it in operational mode in the nearest future.

### SEA ICE

Many challenges related to the satellite monitoring of sea ice are common with those that accompany determining other sea/land surface characteristics (e.g., lack of data in cloud-covered regions in visible imagery, or limited space/time data coverage in ASAR data [7]). Other are specifically related to sea ice properties. In order to fully characterize the state of the sea ice cover in a given area, a number of such properties have to be measured, including ice concentration (fraction of the sea surface covered with ice), thickness distribution, floe-size distribution, ice age and type (level ice, frazil ice, deformed ice etc.; Fig. 3), albedo, snow-cover thickness, velocity, temperature, and so on. Many of these properties are strongly scale-dependent, i.e., they change with changes of the resolution of observation/analysis, which makes comparisons of results obtained with different methods difficult if not impossible.

Whereas satellite sea ice monitoring is difficult in general, it is particularly challenging in the case of seasonal sea ice, which is thin, has relatively low albedo and high temperature (in the early phases of freezing close to the albedo and temperature of the surrounding sea water), and due to its low mechanical strength is easily deformed and broken during storm periods and susceptible to fast melting under favourable conditions. Consequently, the state of the ice cover may change rapidly in periods with strong wind and/or air temperature variability. These events – interesting scientifically and important from the point of view of marine transport and weather prediction in affected areas – pose a particularly demanding test for satellite algorithms and expose their limitations. A storm over

the northern Baltic Sea in early March 2010 provides a good example (Fig. 4). Under strong westerly winds associated with a fast-moving low-pressure system, the hitherto compact ice cover in the Gulf of Bothnia broke up and drifted to the east, leaving large areas of open water in the western parts of the basin (Fig. 4d, h). This significant change is resolved neither in the low-resolution (1°) Global Forecasting System (GFS) dataset (Fig. 4a, e), nor in the Group for High Resolution SST (GHRSSST) dataset dedicated to the Baltic Sea (Fig. 8b, f; resolution 0.03°) – both popular products with several applications, e.g., in numerical modelling. In both cases the sea ice extent after the storm is hardly different from that before the event. The results of the algorithm developed as part of the SatBaltic project (Fig. 4c, g) are much more realistic. The algorithm (suitable for operational use) produces daily maps of sea ice extent based on the whole set of SEVIRI data from a given day, i.e., 96 maps (available every 15 minutes) from 3 visible (VIS) and 4 infrared (IR) channels. This enables to reduce the number of gaps in the final product, because, first, all pixels that have been cloudless even during a very short time during a day are included in the analysis, and second, the time-variability of values from a given pixel can be used as an indicator of the presence of clouds. The core of the algorithm, which performs the essential classification of the data into 3 classes (water, clouds, and ice), operates on the VIS data; at further stages, correction to the initial results is performed based on the IR data in order to improve detection of thin, low-albedo ice. Importantly, each sea ice map is produced from the previous one by actualizing the values of those pixels that could be identified as either water or ice; in other words, the final products are not updated in permanently cloud-covered regions.

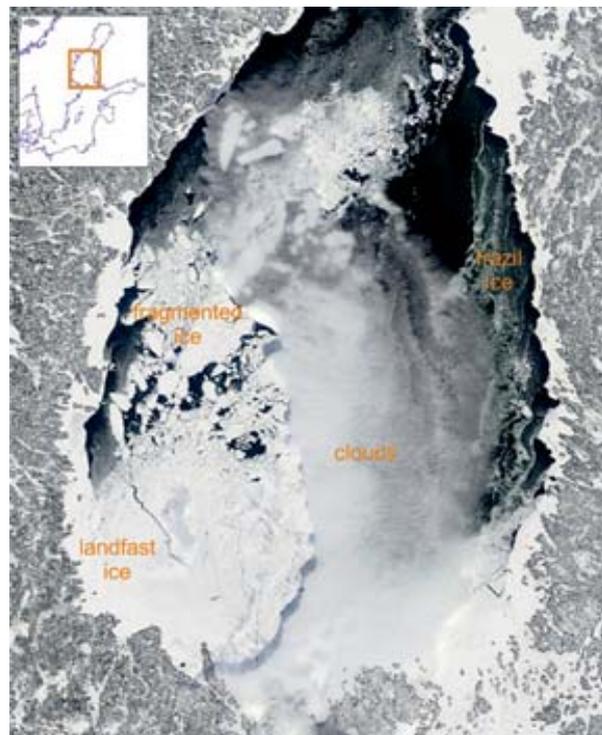


Fig. 3. Fragment of a MODIS Terra image of the Gulf of Bothnia from March 3, 2010

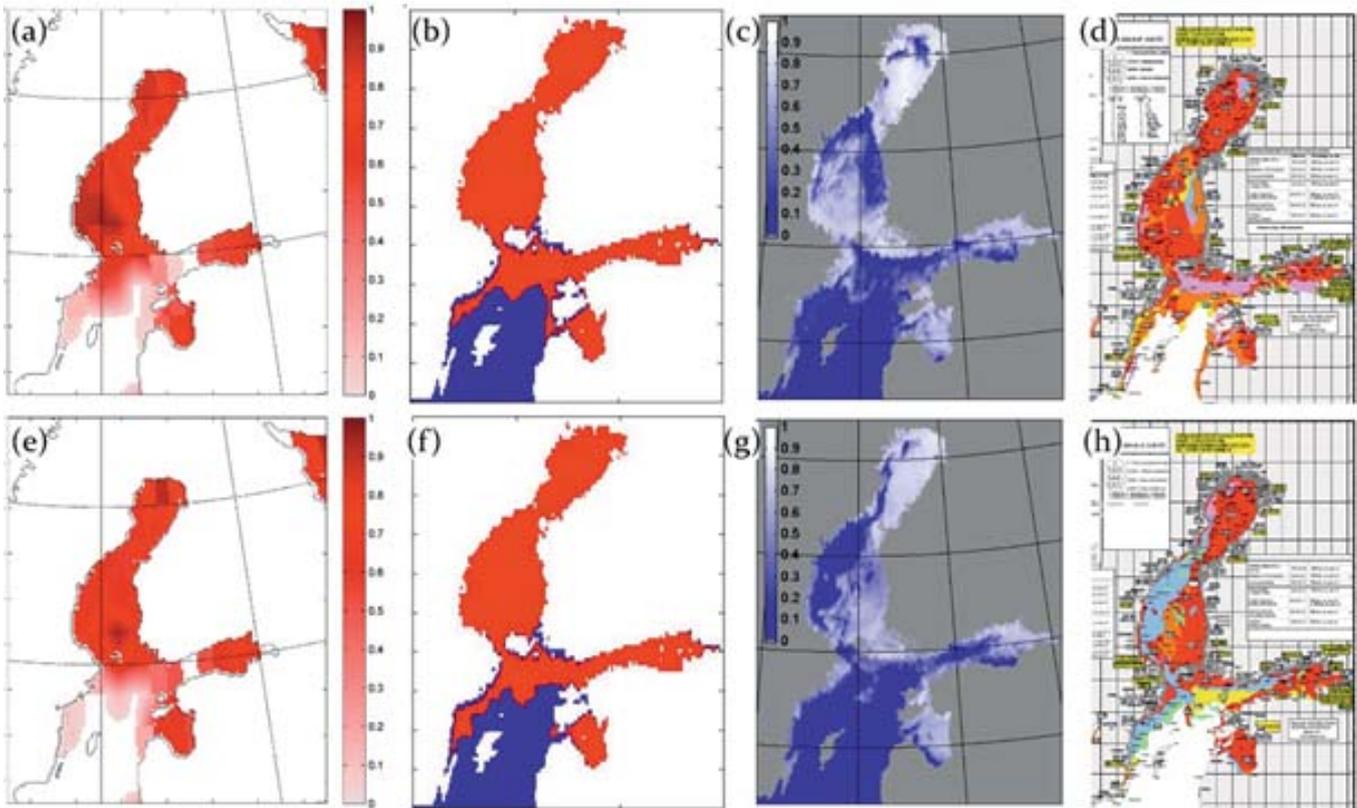


Fig. 4. Examples of satellite sea ice products (a-c, e-g) and SMHI ice charts (d, h) for the Baltic Sea, showing situation before and after a storm in March 2010 (a-d: 7th March, e-h: 12th March). GFS sea ice concentration (a, e), GHRSSST sea ice extent (b, f), SatBaltic sea ice albedo in the High Resolution Visible (HRV) channel (c, g)

Cross-validation with other sea ice products from the period 2010–2014 (not shown), as well as analysis of selected cases like the one presented above, confirm reliability and accuracy of the method, but also reveal its main limitation. Namely, in December and January – periods with short daytime and persistent cloudiness over the northern Baltic Sea – the average time interval between subsequent updates of the maps over the Gulf of Bothnia equals 5 days (as compared to 1.5–2 days in March and 0.5–1 day in April). Hence, although the SEVIRI-based algorithm is very attractive (apart from the ice extent and albedo, it provides a possibility to determine the temperature of the ice surface), it illustrates the need for combining data from many different sources in order to obtain sea ice information sufficiently complete in space and time.

## APPLICATION OF SATELLITE DATA IN ANALYSIS OF THE ENVIRONMENT

### 2.1. SOLAR ENERGY

The amount of solar energy reaching a given place at sea surface depends on the position of its source (solar zenith angle) and the state of Earth's atmosphere in terms of its components that influence solar radiation transmission on the way from its upper boundary. The amount of solar energy can be determined by means of direct measurements or calculations with the use of various models of different precision and complexity. Both measurements and

calculations, are particularly complicated over the sea area due to problems with the exact information about the main components of the atmosphere. In the case of models, their accuracy can be improved by use of information from space. In the simplest way, the solar irradiance at the sea surface  $E_s(\lambda)$  can be expressed as:

$$E_s(\lambda) = \cos \vartheta \frac{F_s}{\beta^2} T(\lambda) \quad (1)$$

where:  $\lambda$  denotes the wavelength;  $\vartheta$  – the solar zenith angle,  $F_s(\lambda)$  – the spectral density of the solar constant;  $\beta = R_s/R$  – the factor defining the annual variability of the distance between the Earth and the Sun ( $R$  and  $R_s$  – actual and mean distance between the Earth and the Sun, respectively) and  $T$  – transmission function depending on all important processes of light attenuation in the atmosphere. It can be expressed in the form of the product:

$$T(\lambda) = T_R(\lambda)T_a(\lambda)T_{O_3}(\lambda)T_{wv}(\lambda)T_G(\lambda)T_{cl} \quad (2)$$

where subscripts R, a,  $O_3$ , wv, G and cl denote dependency on molecular scattering, attenuation by aerosols, absorption by ozone, water vapour and other atmospheric gases, and attenuation by cloudiness, respectively. Optimal forms of some transmission functions in equation (2) can be worked out on the basis of space-borne data about such variables as cloudiness, aerosol optical thickness or ozone and water vapour content in the atmosphere [4]. Below, we present results of model calculations in which some input data were replaced by space-borne source.

## CLOUDS

Cloud transmittance was calculated based on measurements taken in channel 12 (HRV) of SEVIRI. It has spatial resolution of 1 km in the sub-satellite point, recalculated for Baltic to a 4-km grid. The map in Fig. 5 presents the mean diurnal cloud transmission. The ultimate result of the calculations, i.e., the daily dose of solar energy at the sea surface, shown later, was obtained as a sum of the SEVIRI data available every 15 minutes.

## AEROSOLS

From the point of view of the solar radiation transmission, aerosols are the second-important component of the atmosphere. Due to their impact on light, their monitoring from space is performed in a visible part of the spectrum, which means that is possible only in the daytime and in cloudless areas. Standard algorithms used to determine the aerosol optical thickness (AOT) are based on data from a single spectral channel of the AVHRR (Advanced Very High Resolution Radiometer) [9]. The data of AOT that are used to determine  $T_a$  (eq. 2) are operationally archived by NOAA CLASS (Comprehensive Large Array-data Stewardship System) based on one week's worth of data. Their spatial resolution is 100 km. An example of this product for the Baltic Sea is shown in Fig. 6.

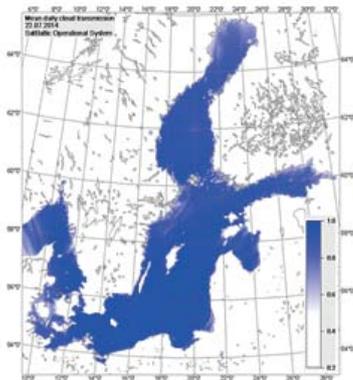


Fig. 5. Mean daily cloud transmission at July 22, 2014 on the basis of SEVIRI data

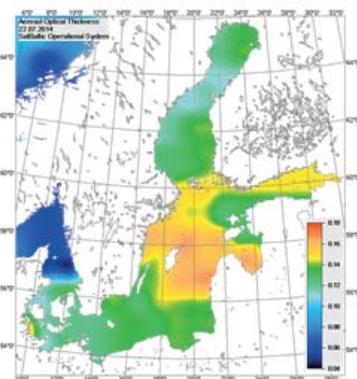


Fig. 6. Aerosol optical thickness on July 22, 2014 on the basis of AVHRR data

## OZONE

Absorption of solar radiation by ozone has a crucial importance for life on our planet. In standard conditions, its role consists in cutting off the UV-wavelength below 280 nm. For the total solar radiation at the sea level, absorption by ozone has minor importance. Nevertheless, the use of space-borne ozone information for determination of  $T_{O_3}$  in the model of light transmission in the atmosphere makes calculations more reliable. Satellite remote sensing of the total ozone in both the stratosphere and troposphere is based mainly on measurements of the ultraviolet albedo of the Earth. An example in Fig. 7 presents the results of analysis made within the framework of TOAST (Total Ozone Analysis using SBUV/2 - Solar Backscatter UltraViolet Version 2 and TOVS - Tiros-N Operational Vertical Sounder). It produces total column ozone in the troposphere and lower stratosphere on the basis of measurements in the UV-range, made by SBUV/2 instrument and in infrared in mid-to-upper stratosphere by TOVS.

## SOLAR ENERGY FLUXES AT THE SEA SURFACE

The above-mentioned continuous observations of the Earth's environment from space provide self-contained information about state of the terrestrial atmosphere. The knowledge obtained in this way makes the results of atmospheric modelling much more reliable, particularly when instantaneous, short-time results are of interest. A good example of information which can be obtained in that way is spatial distribution of photosynthetically active radiation (PAR). Beside the SST and the surface chlorophyll concentration, it is desirable to calculate intensity of such complex and important processes like the primary organic-matter production and the photosynthetically released oxygen in the water. An example result of the PAR calculation with a model working as a part of the SatBaltic Operational System [11] is presented in Fig. 8. An assessment of the accuracy of the model in regard to in situ measurements at the Hel station showed an increase of the linear correlation between the modelled and measured values from 0.87 to 0.94 when non-satellite input data were substituted by space-borne ones.

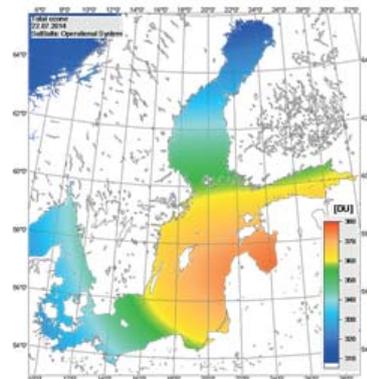


Fig. 7. Total column ozone in the atmosphere over Baltic Sea at July 22, 2014 on the basis of TOAST analysis (values in Dobson units)

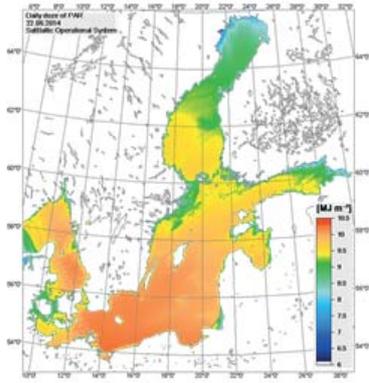


Fig. 8. Example of daily sum of photosynthetically active radiation at the sea surface at July 22, 2014 calculated within SatBaltic Operational System

### SPATIAL DISTRIBUTION OF THE SEA SURFACE TEMPERATURE

There were many radiometers providing SST data over the history of environmental remote sensing. One of them is the AVHRR radiometer operating since 1978 on board a series of American NOAA satellites and now also on the European METOP -A and -B. Its three thermal infra-red bands let to estimate SST in clear-sky conditions with errors less than 1 K and a resolution of 1 km.

In the Baltic Sea area, frequent cloudiness often partially or entirely precludes the use of satellite sensors for recording radiation leaving the sea surface in the thermal infra-red bands. Thus, precise algorithms for detection of cloudiness, as well as ice cover, are required to mask areas where SST estimation is not possible. Despite this limitation, nowadays, by combining historical satellite data derived from different radiometers (AVHRR, MODIS, SEVIRI, etc.) we have access to huge database suitable for analysis of long-term changes in SST. In situ data, of course, provide longer time series of water temperature observations, but they are usually limited to the few coastal monitoring stations. Satellite data enable to study spatial variations of the climate change effects (Fig.9).

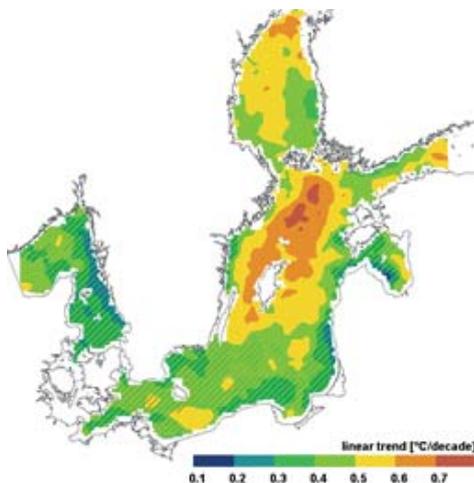


Fig. 9. Linear trends of the annual mean SST in the period 1986–2005 based on AVHRR and MODIS data from PO.DAAC (Physical Oceanography Distributed Active Archive Center) database

The general utility of the satellite data is even more apparent in studies of episodic phenomena such as upwelling currents or spreading of fresh-water inflows.

Coastal upwellings in the Baltic Sea [8] brings water from below the thermocline to the surface, where it mixes with the upper-layer waters. During summer and autumn, when the sea surface is warm, upwelling is seen on infrared satellite images as a local drop in temperature up to several degrees. Thermal fronts derived from SST maps show spatial scale of this phenomenon and its development in time (Fig. 10).

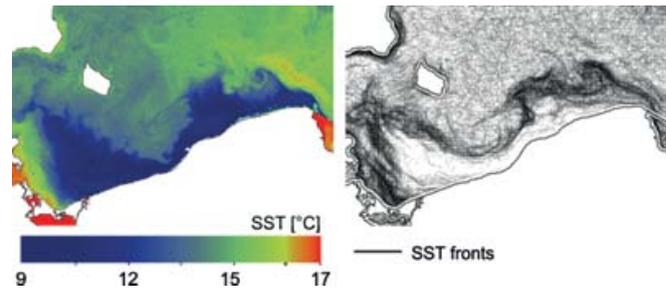


Fig. 10. Example of SST distribution (AVHRR June 30, 2008) with evidence of a coastal upwelling event along the Polish coast and location of thermal fronts determined on the basis of a gradient analysis on a sequence of images collected during 7 days

In the southern part of the Baltic, rivers bring to the sea a lot of suspension and dissolved organic matter, which influence the transparency of sea water. Due to the high dynamics of riverine plumes, investigation of the extent of offshore penetration of the fluvial waters is difficult. However, they usually differ from seawater in temperature, especially in spring and autumn, when inflow is also the highest. This fact enables us to use thermal radiometers for investigation of the extent of offshore penetration of the riverine waters. It gives contemporary information to maps of optical parameters (e.g. diffuse light attenuation coefficient  $K_d$  in Fig. 11) derived from ocean colour data, which are limited to daily observations (Fig.11).

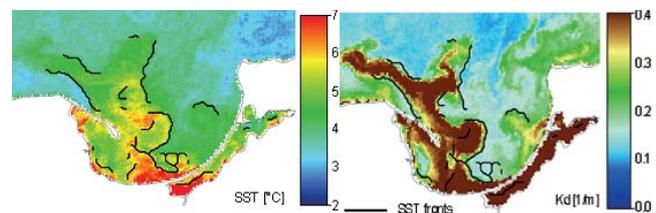


Fig. 11. Example of SST distribution with evidence of the fresh water inflow from the Vistula River and collocation of thermal fronts with spatial distribution of turbid waters (MODIS April 1, 2005)

### CONCLUSION

The problems presented in this paper demonstrate, on the one hand, an immense potential to increase our knowledge about environment if observed from space, and on the other hand, how much can be done yet in that matter. The examples shown were often based on data recorded by “classic” satellite

radiometers (AVHRR, MODIS, ASAR, SEVIRI). Some of them (ASAR/ENVISAT) have already ended their activity, others, for example MODIS, are approaching the end of their longstanding missions. Nevertheless, the very positive experience with them results in preparing new missions. One of them, in which our group is engaged, is a family of ESA's missions called Sentinels [2]. Its first member, Sentinel 1, was launched on April 3, 2014. Its activity gives us an opportunity to remove at least some of the problems related, e.g., to sea ice monitoring (our sea-ice algorithm based on archival ASAR data gave very promising results). The next one, Sentinel 3 which should start this year is a multi-instrument mission to measure sea-surface topography, sea- and land-surface temperature, ocean colour and land colour with high-end accuracy and reliability. Another one, Sentinel 4, is a payload devoted to atmospheric monitoring that will be embarked upon a Meteosat Third Generation-Sounder (MTG-S) satellite in geostationary orbit.

This is only a small example of the very fast development of the Earth observing system; it shows that every activity leading to the development of infrastructure suitable to process the ever-increasing amounts of data is justified and still needed. It can guarantee a better knowledge of phenomena and processes that occur in the ever faster changing environment, and give us a chance of better understanding the surrounding world.

## ACKNOWLEDGEMENT

This work was supported by the European Union through the European Regional Development Fund Project Satellite Monitoring of the Baltic Sea Environment (SatBałtyk) POIG 01.01.02-22-011/09 and DS/530-G210-D425-14.

## REFERENCES

1. S.A. Ackerman, K.I. Strabala, W.P. Menzel, R.A. Frey, C.C. Moeller, L.E. Gumley, Discriminating clear sky from clouds with MODIS, *J. Geophys. Res.*, 103, D24, (1998), 32,141–32,157.
2. European autonomy in space, C. Al-Ekabi (ed.), Springer International Publishing, (2015), 198.
3. J. Fabisiak: Ecological risks associated with chemical pollution of the Baltic Sea – hydrocarbons, (in Polish), *Zeszyty Naukowe Marynarki Wojennej*, XLIX, 3 (174), (2008), 7-24.
4. M. Konik, K. Bradtke: Characteristics of the oil spills detected in the Baltic Sea area in ENVISAT ASAR Images, *Earth observation for ocean-atmosphere interactions science*, Conference, 28-31.10.(2014), Frascati (Italy).
5. A. Krężel, K. Bradtke: Estimation of solar energy influx to the sea in the light of fast satellite technique development, in Rugescu Radu ed., *Solar Power*, Rijeka, InTech, (2012), ch. 10, pp. 171-192.

6. A. Krężel, Ł. Kozłowski, M. Paszkuta: A simple model of light transmission through the atmosphere over the Baltic Sea utilising satellite data, *Oceanologia*, 50 (2), (2008), 125-146.
7. A. Mazur, A. Krężel: Object-based classification of Baltic Sea ice extent and concentration in winter, *J. Earth Sci. Engng*, 2, (2011), 488-495.
8. K. Myrberg, O. Andrejev, Main upwelling regions in the Baltic Sea – a statistical analysis based on three-dimensional modelling, *Boreal Environment Research*, 8, (2003), 97-112.
9. L.L. Stowe., A.M. Ignatov, R.R. Singh: Development, validation, and potential enhancements to the second-generation operational aerosol product at the National Environmental Satellite, Data, and Information Service of the National Oceanic and Atmospheric Administration, *J. Geophys. Res.*, 102, D14, (1997), 16923-16934.
10. M.J. Uddstrom, W.R. Gray, Satellite cloud classification and rain-rate estimation using multispectral radiances and measures of spatial texture. *J. Appl. Met.*, 35: (1996), 839–858.
11. B. Woźniak, K. Bradtke, M. Darecki, J. Dera, J. Dudzińska-Nowak, L. Dzierzbicka-Głowacka, D. Ficek, K. Furmańczyk, M. Kowalewski, A. Krężel, R. Majchrowski, M. Ostrowska, M. Paszkuta, J. Stoń-Egiert, M. Stramska, T. Zapadka: SatBałtyk – A Baltic environmental satellite remote sensing system – an ongoing project in Poland. Part 1: Assumptions, scope and operating range, *Oceanologia*, 53 (4), (2011), 897-924.

## CONTACT WITH AUTHOR

Adam Krężel  
Katarzyna Bradtke  
Agnieszka Herman

Institute of Oceanography  
University of Gdańsk  
Marszałka Piłsudskiego 46 Ave.  
81-378 Gdynia  
Poland

e-mail: oceak@ug.edu.pl

# PRACTICAL APPLICABILITY AND PRELIMINARY RESULTS OF THE BALTIC ENVIRONMENTAL SATELLITE REMOTE SENSING SYSTEM (SATBAŁTYK)

Mirosława Ostrowska<sup>a</sup>

Mirosław Darecki<sup>a</sup>

Adam Krężel<sup>b</sup>

Dariusz Ficiek<sup>c</sup>

Kazimierz Furmańczyk<sup>d</sup>

<sup>a)</sup> Institute of Oceanology Polish Academy of Sciences, Poland

<sup>b)</sup> University of Gdańsk, Institute of Oceanography, Poland

<sup>c)</sup> Pomeranian University in Słupsk, Poland

<sup>d)</sup> University of Szczecin, Poland

## ABSTRACT

*The SatBałtyk (Satellite Monitoring of the Baltic Sea Environment) project is being realized in Poland by the SatBałtyk Scientific Consortium, specifically appointed for this purpose, which associates four scientific institutions: the Institute of Oceanology PAN in Sopot – coordinator of the project, the University of Gdańsk (Institute of Oceanography), the Pomeranian Academy in Słupsk (Institute of Physics) and the University of Szczecin (Institute of Marine Sciences). The project is aiming to prepare a technical infrastructure and set in motion operational procedures for the satellite monitoring of the Baltic Sea ecosystem. The main sources of input data for this system will be the results of systematic observations by metrological and environmental satellites such as TIROS N/NOAA, MSG (currently Meteosat 10), EOS/AQUA and Sentinel -1, 2, 3 (in the future). The system will deliver on a routine basis the variety of structural and functional properties of this sea, based on data provided by relevant satellites and supported by hydro-biological models. Among them: the solar radiation influx to the sea's waters in various spectral intervals, energy balances of the short- and long-wave radiation at the Baltic Sea surface and in the upper layers of the atmosphere over the Baltic, sea surface temperature distribution, dynamic states of the water surface, concentrations of chlorophyll a and other phytoplankton pigments in the Baltic waters, spatial distributions of algal blooms, the occurrence of coastal upwelling events, and the characteristics of primary production of organic matter and photosynthetically released oxygen in the water and many others. The structure of the system and preliminary results will be presented.*

**Keywords:** bearing capacity factor, displacement, compression, tension, depth factor

## INTRODUCTION

Highly significant part of the ongoing transformation of the Earth's natural environment worldwide is being played by processes taking place in marine ecosystems. It leads researchers to intensify the studies and monitoring of the state and productivity of different seas and oceans. For many years now, various environment state parameters were measured at sea using traditional *in situ* methods on board of research vessels. Such traditional methods are costly, not very effective and do not satisfy some of the present-day requirements. Monitoring of the vast expanses of seas and

oceans, characterized usually by high temporal variability requires more systematic approach. Development of remote methods to investigate the functioning of marine ecosystems, particularly those based on satellite observations of sea colour, enables the efficient global monitoring of the state of the sea. The paper presents some of the most essential results from SatBałtyk project. Its aim is to develop monitoring system for the Baltic, based mostly on the space born information.

## SHORT HISTORY OF REMOTE SENSING OF THE BALTIC SEA

The beginnings of the remote sensing of the Baltic Sea by Polish scientists go back to the early 1990s. It was a time when they start to develop on the one hand the modelling of bio-optical phenomena taking place in the sea [8,13,14,20,22,23,25], and on the other hand, remote optical methods for studying the functioning of marine ecosystems, in particular techniques based on satellite observations [6,7,11,12,15,17,24].

In 2001–2005 the Institute of Oceanology PAN, together with the Institute of Oceanography (University of Gdańsk), the Pedagogical (now Pomeranian) Academy of Słupsk, and the Sea Fisheries Institutes of Gdynia, conducted research within the framework of a project commissioned by the Polish Committee for Scientific Research – DESAMBEM (DEvelopment of a SAtellite Method for Baltic Ecosystem Monitoring; project No. PBZ-KBN 56/P04/2001).

That project's objective was to develop scientific principles and methods for utilising remote sensing technology to monitor the Baltic, an enclosed sea with a high level of biological productivity. This cooperation yielded the first version of the DESAMBEM algorithm which enables to make, using remote sensing data, a set of map with spatial distribution of four significant characteristics of the Baltic Sea, namely, photosynthetically active radiation (PAR) at the sea surface, sea surface temperature, surface chlorophyll *a* concentration and total primary production in the water column [18,19].

It was obvious, that those results should be implemented in the interests of the efficient and systematic monitoring of the state of the Baltic environment and the forecasting of the changes taking place in it. This imposes the duty of conserving the natural environment of the Baltic in accordance with international conventions and legal regulations.

The implementation of the DESAMBEM methods has become possible in early 2010 thanks to EU funding. Thus came into being project No. POIG.01.01.02-22-011/09-00 entitled 'The satellite monitoring of the Baltic Sea environment' (acronym SatBałtyk) realized in Poland by the SatBałtyk Scientific Consortium, specifically appointed for this purpose, which associates four scientific institutions: the Institute of Oceanology PAN in Sopot - coordinator, the University of Gdańsk (Institute of Oceanography), the Pomeranian Academy in Słupsk (Institute of Physics) and the University of Szczecin (Institute of Marine Sciences).

The main aim of the SatBałtyk project is to prepare the technical infrastructure and set in motion practical operational procedures for the efficient, routine monitoring of the state of the Baltic environment described at the moment by more than 90 structural and functional characteristics. The expected end result of the project is to set up an operational system for the rapid and efficient determination of these characteristics of the Baltic Sea in the form of maps showing their distribution in this sea. This complex task is being implemented (2010-2015).

## THE THEORETICAL BASIS OF OPERATIONAL SYSTEM SATBAŁTYK

The use of remote sensing methods in studies of the sea is relatively simple only with respect to the waters of the central oceanic regions. The great majority of substances affecting the colour of the sea in those regions are autogenic, that is, formed by the local ecosystem – photosynthesis by phytoplankton and the metabolism and decay of marine organisms. In consequence, the spectrum of the light emerging from these waters and reaching satellite sensors is correlated with the concentration of phytoplankton and its pigments, principally chlorophyll *a*, the commonest plant pigment.

Most of the algorithms now in common use for characterizing the state and functioning of marine ecosystems on the basis of remote sensing data are thus applicable to these waters: they utilize the correlations of their optical properties with the chlorophyll *a* concentration in surface waters and the correlation of this concentration with other properties of the aquatic environment (e.g. [1,3-5,9,16,21]). In contrast, the construction of analogous algorithms for the optically complex waters of Baltic Sea basin, is much more complicated. Apart from chlorophyll and other products of the local ecosystem, such waters contain many substances entering it from the exterior (from rivers, the land, the atmosphere, the sea bed and shores), which have complex optical properties, not directly correlated with the chlorophyll *a* concentration [26]. The use of remote sensing techniques to monitor such waters requires the application of different algorithms, purpose-designed for a particular sea region. A serious problem hampering the design and use of these algorithms is also the high variability of atmospheric states, which distort the light spectrum bearing information from the sea to the satellite. Work on the development of suitable algorithms for the Baltic Sea has been going on in Poland for the last 20 years by the teams of researchers from mother institutions of authors of this presentation [27,28]. We now present only the most essential information characterizing the progress of this modelling.

Figure 1 illustrates the simplified general block diagram of main components of the SatBałtyk operations system. This system consists of two independent but coordinating subsystems:

**The DESAMBEM Diagnostic System**, upon which the entire SatBałtyk operations system is founded, enables current structural and functional parameters of the marine environment to be determined on the basis of the relevant calculations, for which the input data are the results of current remote sensing data acquisition.

**The BALTFOS Forecasting System** based on a packet of prognostic hydrodynamic and ecological models.

Both, DESAMBEM and BALTFOS contain sets of algorithms (in view of the sequence of the calculations to be carried out and their complexity, in Fig. 1 some blocks of subalgorithms were distinguished) enabling diagnose current or forecasted the future sea states on the basis of appropriate input data, the sources of which are principally satellite

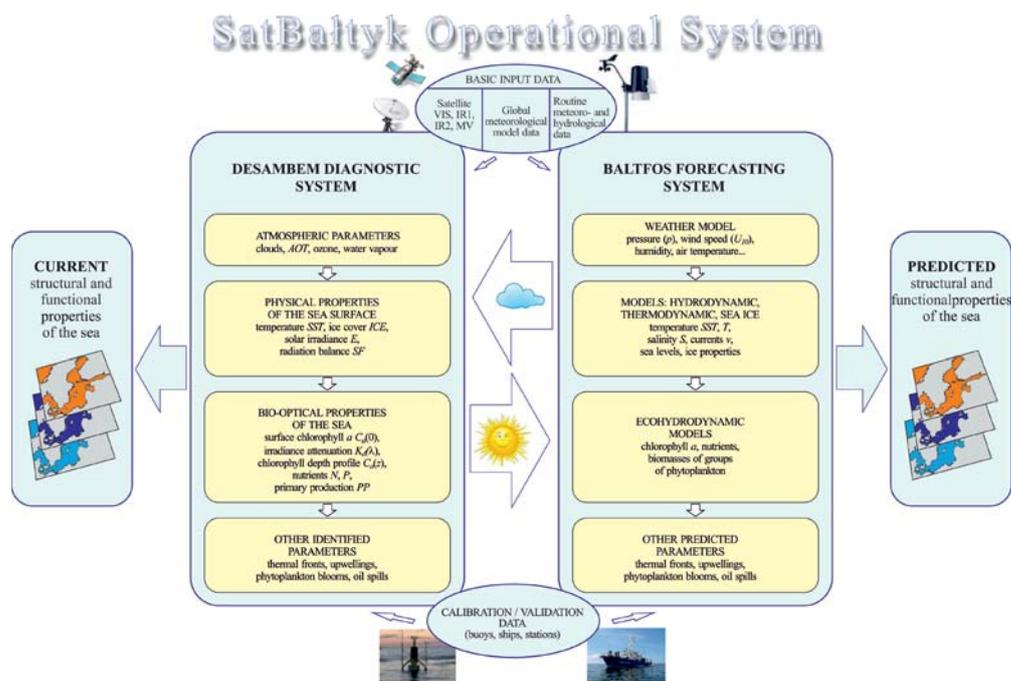


Fig. 1. Simplified general block of the Operational System SatBałtyk

radiometers and/or hydrometeorological data supplied by common routine services.

These two cooperating data processing subsystems are complementary within the framework of the SatBałtyk Operational System. So, for example, if direct remote sensing of the sea is not possible because the sky is overcast, especially if the sea surface temperature (SST) and the surface chlorophyll *a* concentration  $C_a(0)$  cannot be retrieved, the values of these data in the DESAMBEM subsystem serving to determine various current values of marine ecosystem parameters based on SST and  $C_a(0)$  data are based on forecasts made by the BALTFOS subsystem (see the horizontal arrows from right to left between the subsystems in Fig. 1). On the other hand, the forecast values of parameters determined by the component models of the BALTFOS subsystem can be rectified by the assimilation of the actual values of these parameters determined by the DESAMBEM algorithm (see the horizontal arrows from left to right between the subsystems in Fig. 1).

## THE GENERAL SCHEME AND SELECTED RESULTS OF THE SATBAŁTYK OPERATIONAL SYSTEM FUNCTIONING

The SatBałtyk Operational System (SatBałtyk OS) is a complex set of operational measurement instruments, mathematical models and procedures enabling the routine monitoring of different characteristics of the Baltic Sea environment as well as their presentation (mostly via the Internet) in the form of maps and tables, as well as several-day forecasts. The instrumentation used by SatBałtyk OS comprises receivers of data recorded by selected satellites, sets of measurements from marine and meteorological sondes, loggers, converters and transmitters of the data from

these sondes, as well as laboratory analytical apparatus and a network of computer servers with installed databases, the relevant mathematical models and algorithms, to be used in accordance with stipulated procedures. Figure 2 illustrates the main components of the SatBałtyk operational system and of how it is ultimately expected to function.

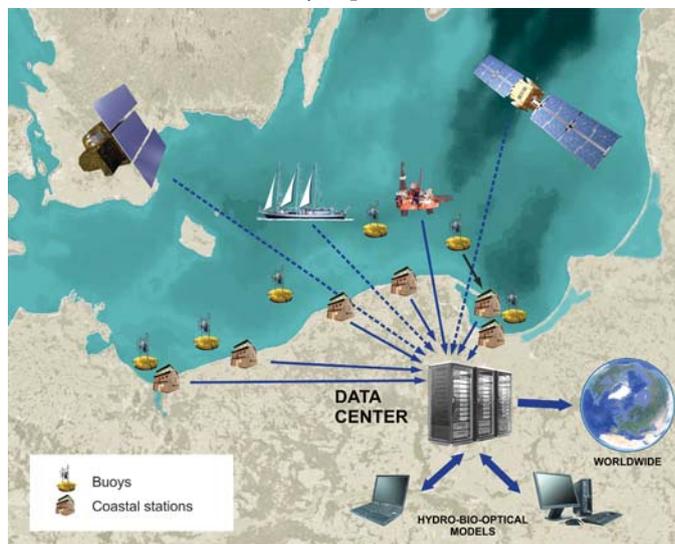


Fig. 2. Main components of the Operational System SatBałtyk

The system utilizes data systematically recorded by environmental satellites and also special-purpose ones such as TIROS N/NOAA, MSG (currently Meteosat 9), EOS/AQUA and Sentinel -1, 2, 3 (in the future) regularly passing over the Baltic, and various in situ data measured on buoys and ships, shore stations and analyzed in laboratories belonging to Institutes from the Consortium SatBałtyk.

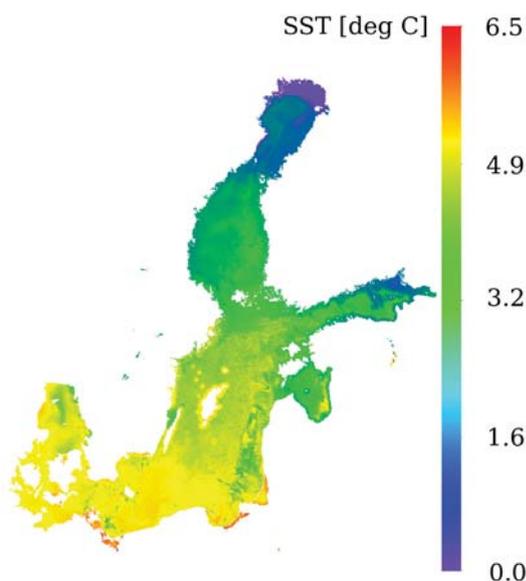


Fig. 3. Sea surface temperature (SST) of the Baltic Sea constructed on the basis of satellite data and completed by the model (17 03 2015)

As we have already stated, in the case of the Baltic, frequent cloudiness partially or entirely precludes the use of satellite sensors for recording radiation in the visible and thermal infra-red bands for diagnosing various parameters of the marine environment. Because of this the applicability of the DESAMBEM algorithms were expanded by linking them up with the set of algorithms from the BALTFOS Forecasting System. The latter are based on prognostic models and procedures for their adjustment by the assimilation of space born and other data and other data obtained using the diagnostic subalgorithms of the DESAMBEM (see Fig. 1). Presented in Fig. 3 map of the surface temperature SST of whole area of the Baltic Sea is constructed on the basis of satellite data and supplemented by the results of modelling.

The next example of the maps provided by the system, is phytoplankton assimilation number, characterizing the condition of phytoplankton, in particular its potential photosynthetic abilities.

Fig. 4 represents maximum possible rate of photosynthesis in waters for a fixed amount of nutrients in those waters and a particular sea water temperature expressed in numbers of atoms or moles of carbon assimilated in unit time by phytoplankton of unit chlorophyll content. Such maximum rates of photosynthesis are usually recorded at intermediate (photosynthetically optimal) depths, at which irradiance levels are still sufficiently high not to limit the rate of light reactions, yet not so high that destructive photoinhibition of the photosynthetic apparatus comes into play. In the Baltic such optimal conditions usually (in ca 66% of cases) prevail at depths from 1 to 5 m [29].

List of the various characteristics of the Baltic Sea environment, provided by SatBałtyk OS is constantly expanding and includes, for example: fluxes and budgets of radiant energy reaching the various regions of the Baltic and the amount of energy driving the Baltic ecosystem including the inflow and characteristics of radiant energy

PAR (photosynthetically ctive radiation (~400–700 nm)) and UV (ultraviolet radiation), temperature distributions, the occurrence of upwelling events, the dynamic state of the sea surface, salinity of surface waters, the speed and direction of surface currents, upwelling events, water transparency indices and other optical properties of the sea, surface ice cover, the concentrations of chlorophyll and other phytoplankton pigments, blooms of toxic algae, the appearance of patches of pollutants such as oil slicks, characteristics of the primary production of organic matter and the release of oxygen into the water during photosynthesis, the humidity and ozone content in the atmosphere over the Baltic.

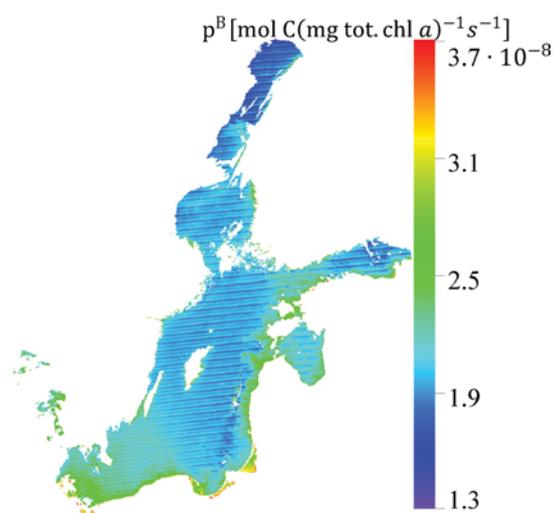


Fig. 4. Maximum rates of photosynthesis of the Baltic Sea constructed on the basis of satellite data (17 03 2015)

Integrated part of SatBałtyk system is a module called SatBałtyk-Brzegi, which allows to monitor two types of parameters of the coastal zone conditions and their prediction for two days in advance. Values of the first type of parameters are measured for one kilometer long sections of the coast for three pilot areas of a dune coast: Dziwnów Spit (12 km), Spit of Jamno and Bukowo Lakes (10 km) and Hel Peninsula (6 km), in perspective to cover the whole length of the Polish coast. Two days prediction consists of the following parameters: beach width inundation in percent; dry corridor of the beach; possibility of rip current occurrence in the nearshore. The second type of parameters shows consequences of the storm event and they are calculated after the process ends. The parameters are: volume of the erosion along the coastal profile: topography and bathymetry; volume of dune erosion at the one kilometer section along the coast; amount of inorganic suspended matter concentration. The calculation of parameters were done on a basis of hydromorphodynamic XBeach model which was adopted for pilot areas by [2] The model was also used to construct the Early Warning System of Storm Impact, developed by Szczecin University team in the framework of MICORE project [10]. Additionally two

parameters: width inundation in percent; dry corridor of the beach, are calculated using statistical methods for the whole Polish coast. The values of parameters calculated in a framework of SatBałtyk project are crucial for research of coastal processes and management. Predicted parameters are especially important for safety use of the coastal zone.

An important component of the SatBałtyk system is a systematic validation of the products provided by the service. In situ data, collected on every day basis at the platform and our buoy as well data from regularly organised cruises (usually 5 per year) are used to check accuracy of the products. Also a third-party data e.g. from ICES have been used to validate environmental parameters delivered by our system. The recent validation showed that accuracy of some of them are as follow: sea surface temperature – systematic error  $-0.03^{\circ}\text{C}$ , statistical error  $-0.91^{\circ}\text{C}$ ; surface chlorophyll a concentration: – systematic error  $-9\%$ , statistical error  $-140\%$ ; downwelling photosynthetically active radiation – systematic error  $-0.11\text{Wm}^{-2}$ , statistical error  $-40\text{Wm}^{-2}$ .

More information and characteristics of the Baltic Sea environment, can be find at the website [www.satbaltyk.eu](http://www.satbaltyk.eu).

## CONCLUSIONS

The article brings the summary of the results achieved during implementation of SatBałtyk project founded on the scientific outcomes from earlier, long-term studies of the Baltic environment by the Institutes forming the SatBałtyk Consortium. This system, based mainly on satellite data, enables making of maps of spatial distributions of many state parameters of this environment, as well as certain state parameters and optical properties of the atmosphere. The establishing the scientific foundations and methodology for employing remote sensing techniques to monitor the Baltic as an inland sea with a high biological productivity and yet under serious as a consequence of economic development is great Polish achievements. The final result of the project is the creation and setting in motion of the SatBałtyk Operational System for the comprehensive and effective monitoring of the current state of the marine environment, and also to a large degree for the forecasting of a whole range of natural phenomena taking place in Baltic waters and in the atmosphere above. The end of the work is schedule for December 31, 2015 after 72 months of implementation.

The magnitudes of ecosystem parameters determined using the set of SatBałtyk algorithms from NOAA, Meteosat, the MSG/SEVIRI and MODIS AQUA satellite data are compared with the magnitudes of the same parameters measured in situ. For these comparisons the relevant errors have been calculated. These errors are relatively small, typical of remote, spatial estimates. Such results can be regarded as satisfactory. Nevertheless, improvement of all the components of this complex algorithm will continue.

## ACKNOWLEDGEMENT

The work was supported within the framework of the SatBałtyk project funded by the European Union through European Regional Development Fund, (contract No. POIG.01.01.02-22-011/09 entitled 'The Satellite Monitoring of the Baltic Sea Environment') and IO PAN statutory research.

## REFERENCES

1. Antoine D., J. M. André, A. Morel: Oceanic primary production: 2. Estimation at global scale from satellite (Coastal Zone Color Scanner) chlorophyll, *Global Biogeochemical Cycles* 10 (1) (1996) 56-69.
2. Bugajny, N., Furmańczyk, K., Dudzińska-Nowak, J. i Paplińska-Swerpel, B., 2013. Modelling morphological changes of beach and dune induced by storm on the Southern Baltic coast using XBeach (case study: Dziwnow Spit). *Journal of Coastal Research*, 65, pp. 672-677.
3. Campbell J., D. Antoine, R. Armstrong, K. Arrigo, W. Balch, R. Barber, M. Behrenfeld, R. Bidigare, J. Bishop, M-E. W. Carr, Esaias, P. Falkowski, N. Hoepner, R. Iverson, D. Keifer, S. Lohrenz, J. Marra, A. Morel, J. Ryan, V. Vedemikov, K. Waters, C. Yentsch, J. Yoder: Comparison of algorithms for estimating ocean primary production from surface chlorophyll, temperature, and irradiance, *Global Biogeochemical Cycles* 16 (2002) 74-75.
4. Carr M-E., M. A. Friedrichs, M. Schmeltz, M. N. Aita, D. Antoine, K. R. Arrigo, I. Asanuma, O. Aumont, R. Barber, M. Behrenfeld, R. Bidigare, E. T. Buitenhuis, J. Campbell, A. Ciotti, H. Dierssen, M. Dowell, J. Dunne, W. Esaias, B. Gentili, W. Gregg, S. Groom, N. Hoepner, J. Ishizaka, T. Kameda, C. Le Quere, S. Lohrenz, J. Marra, F. Melin, K. Moore, A. Morel, T. E. Reddy, J. Ryan, M. Scardi, T. Smyth, K. Turpie, G. Tilstone, K. Waters, Y. Yamanaka: A comparison of global estimates of marine primary production from ocean color, *Deep Sea Res.* 53 (2006) 741-770.
5. Darecki M., D. Ficek, A. Krężel, M. Ostrowska, R. Majchrowski, S.B. Woźniak, K. Bradtke, J. Dera, B. Woźniak: Algorithms for the remote sensing of the Baltic ecosystem (DESAMBEM). *Part 2: Empirical validation*, *Oceanologia* 50(4) (2008) 509-538.
6. Darecki M., D. Stramski: An evaluation of MODIS and SeaWiFS bio-optical algorithms in the Baltic Sea, *Remote Sens. Environ.* 89 (3) (2004) 326–350.
7. Darecki M., S. Kaczmarek, J. Olszewski: SeaWiFS chlorophyll algorithms for the Southern Baltic, *Int. J. Remote Sens.* 26 (2) (2005) 247–260.

8. Ficek D., B. Woźniak, R. Majchrowski, M. Ostrowska: Influence of nonphotosynthetic pigments on the measured quantum yield of photosynthesis, *Oceanologia* **42(2)** (2000) 231–242.
9. Ficek D., R. Majchrowski, M. Ostrowska, S. Kaczmarek, B. Woźniak, J. Dera: Practical applications of the multi-component marine photosynthesis model (MCM), *Oceanologia* **45(3)** (2003) 395–423.
10. Furmańczyk, K., Bugajny, N., Dudzińska-Nowak, J., Andrzejewski, P., Trefenko, P., Giza, A., Benedyczak, R. i Cieszyński, Ł., 2012. Projekt MICORE – podsumowanie zrealizowanych zadań. ZZOP w Polsce – stan obecny i perspektywy. Tom 4. Zagrożenia i systemy Ostrzegania, str. 77–86.
11. Krężel A.: Solar radiation at the Baltic Sea surface, *Oceanologia* **21** (1985) 5–32.
12. Majchrowski R., M. Ostrowska: Influence of photo- and chromatic acclimation on pigment composition in the sea, *Oceanologia* **42(2)** (2000) 157–175.
13. Majchrowski R., M. Ostrowska: Modified relationships between the occurrence of photoprotecting carotenoids of phytoplankton and Potentially Destructive Radiation in the sea, *Oceanologia* **41(4)** (1999) 589–599.
14. Ostrowska M., R. Majchrowski, D.N. Matorin, B. Woźniak: Variability of the specific fluorescence of chlorophyll in the ocean. Part 1. Theory of classical ‘in situ’ chlorophyll fluorometry, *Oceanologia* **42(2)** (2000) 203–219.
15. Pelevin V. N., B. Woźniak, O. J. Koblenz-Mishke: Algorithm for estimating primary production in the sea from satellite sensing, *Oceanologia* **31** (1991) 57–72.
16. Platt T., S. Sathyendranath, A. Longhurst: Remote-sensing of primary production in the ocean – promise and fulfilment, *Philosophical Transactions of the Royal Society of London, Series B* **348** (1995) 191–201.
17. Rozwadowska A.: Influence of aerosol vertical profile variability on retrievals of aerosol optical thickness from NOAA AVHRR measurements in the Baltic region *Oceanologia* **49(2)** (2007) pp. 165–184.
18. Woźniak B., A. Krężel, J. Dera: Development of a satellite method for Baltic ecosystem monitoring (DESAMBEM) - an ongoing project in Poland, *Oceanologia* **46(3)** (2004) 445–455.
19. Woźniak B., A. Krężel, M. Darecki, S.B. Woźniak, R. Majchrowski, M. Ostrowska, Ł. Kozłowski, D. Ficek, J. Olszewski, J. Dera: Algorithm for the remote sensing of the Baltic ecosystem (DESAMBEM). Part 1: Mathematical apparatus, *Oceanologia* **50(4)** (2008) 451–508.
20. Woźniak B., J. Dera, D. Ficek, M. Ostrowska, R. Majchrowski: Dependence of the photosynthesis quantum yield in oceans on environmental factors, *Oceanologia*, **44(4)** (2002) 439–459.
21. Woźniak B., J. Dera, D. Ficek, R. Majchrowski, M. Ostrowska, S. Kaczmarek: Modelling light and photosynthesis in the marine environment, *Oceanologia* **45(2)** (2003) 171–245.
22. Woźniak B., J. Dera, D. Ficek, R. Majchrowski, S. Kaczmarek, M. Ostrowska, O.I. Koblenz-Mishke: Modelling the influence of acclimation on the absorption properties of marine phytoplankton, *Oceanologia* **41(2)**, (1999) 187–210.
23. Woźniak B., J. Dera, D. Ficek, R. Majchrowski, S. Kaczmarek, M. Ostrowska, O.I. Koblenz-Mishke: Model of the ‘in vivo’ spectral absorption of algal pigments. Part 1. Mathematical apparatus, *Oceanologia* **42(2)** (2000) 177–190.
24. Woźniak B., J. Dera, R. Majchrowski, D. Ficek, O. I. Koblenz-Mishke, M. Darecki: ‘IOPAS initial model’ of marine primary production for remote sensing application, *Oceanologia* **39(4)** (1997) 377–395.
25. Woźniak B., J. Dera, R. Majchrowski, D. Ficek, O.I. Koblenz-Mishke, M. Darecki: Statistical relationships between photosynthesis and abiotic conditions in the ocean – the IOPAS initial model for remote sensing application, *Proc. SPIE* **3222** (1997) 516–528.
26. Woźniak B., J. Dera: *Light Absorption in Sea Water*, Atmospheric and Oceanographic Sciences Library 33, Springer (2007) 454 pp.
27. Woźniak B., K. Bradtke, M. Darecki, J. Dera, J. Dudzińska-Nowak, L. Dzierzbicka-Głowacka, D. Ficek, K. Furmańczyk, M. Kowalewski, A. Krężel, R. Majchrowski, M. Ostrowska, M. Paszkuta, J. Stoń-Egiert, M. Stramska, T. Zapadka: SatBaltic – a Baltic environmental satellite remote sensing system- an ongoing project in Poland. Part 1: Assumptions, scope and operating range, *Oceanologia* **53(4)** (2011) 897–924.
28. Woźniak B., K. Bradtke, M. Darecki, J. Dera, J. Dudzińska-Nowak, L. Dzierzbicka-Głowacka, D. Ficek, K. Furmańczyk, M. Kowalewski, A. Krężel, R. Majchrowski, M. Ostrowska, M. Paszkuta, J. Stoń-Egiert, M. Stramska, T. Zapadka: SatBaltic – a Baltic environmental satellite remote sensing system- an ongoing project in Poland. Part 2: Practical applicability and preliminary results, *Oceanologia* **53(4)** (2011) 925–958.
29. Woźniak B., R. Hapter, J. Dera: Light curves of marine phytoplankton photosynthesis in the Baltic, *Oceanologia* **27** (1989) 61–78.

## CONTACT WITH THE AUTHORS

Mirosława Ostrowska

Institute of Oceanology Polish Academy of Sciences  
Powstańców Warszawy 55, 81-712 Sopot  
e-mail: ostra@iopan.gda.pl

Mirosław Darecki

Institute of Oceanology Polish Academy of Sciences  
Powstańców Warszawy 55, 81-712 Sopot  
e-mail: darecki@iopan.gda.pl

Adam Krężel

University of Gdańsk, Institute of Oceanography  
Marszałka Piłsudskiego 46, 81-378 Gdynia  
e-mail: oceak@ug.edu.pl

Dariusz Ficek

Pomeranian University in Słupsk  
Bohaterów Westerplatte 64, 76-200 Słupsk  
e-mail: ficek@apsl.edu.pl

Kazimierz Furmańczyk

University of Szczecin  
Mickiewicza 16, 70-383 Szczecin  
e-mail: kaz@univ.szczecin.pl

**POLAND**

# DAILY RADIATION BUDGET OF THE BALTIC SEA SURFACE FROM SATELLITE DATA

Tomasz Zapadka<sup>a</sup>  
Adam Krężel<sup>b</sup>  
Marcin Paszkuta<sup>b</sup>  
Mirosław. Darecki<sup>c</sup>

<sup>a</sup>Pomeranian University in Słupsk, Institute of Physics, Poland

<sup>b</sup>University of Gdańsk, Institute of Oceanography, Poland

<sup>c</sup>Institute of Oceanology Polish Academy of Sciences, Poland

## ABSTRACT

Recently developed system for assessment of radiation budget for the Baltic Sea has been presented and verified. The system utilizes data from various sources: satellite, model and in situ measurements. It has been developed within the SatBałtyk project (Satellite Monitoring of the Baltic Sea Environment - [www.satbaaltyk.eu](http://www.satbaaltyk.eu)) where the energy radiation budget is one of the key element. The SatBałtyk system generates daily maps of the all components of radiation budget on every day basis. We show the scheme of making daily maps, applied algorithms and empirical data collection within the system. An empirical verification of the system has been carried out based on empirical data collected on the oil rig placed on the Baltic Sea. This verification concerned all the components of the surface radiation budget. The average daily NET products are estimated with statistical error ca.  $13 \text{ Wm}^{-2}$ . The biggest absolute statistical error is for LWd component and equals  $14 \text{ Wm}^{-2}$ . The relative error in relation to the average annual values for whole Baltic is the biggest for SWu and reaches 25%. All estimated components have correlation coefficient above 0.91.

**Keywords:** Surface radiation budget, satellite, Baltic Sea

## INTRODUCTION

The exchange of radiative energy between the atmosphere and the sea surface plays an important role in shaping the climate of the Baltic Sea. Monitoring of the radiation budget components at the sea surface allows to assess possible changes in the Baltic Sea environment and its directions. The surface radiative budget  $NET$  consists of four fluxes: downward  $SW_d$  and upward  $SW_u$  shortwave (solar radiation range  $0.3 \mu\text{m} - 4 \mu\text{m}$ ), and downward  $LW_d$  and upward  $LW_u$  (atmospheric and sea surface thermal radiation range  $4 \mu\text{m} - 100 \mu\text{m}$ ) longwave radiation. The sum of all fluxes gives the surface radiation budget at surface  $NET$ . Details of these fluxes are shown on Fig.1.

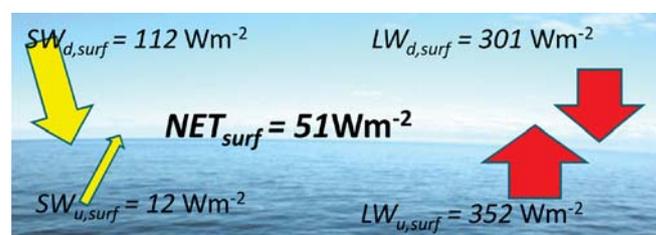


Fig. 1. Components of surface radiation budget with approximate values for the Baltic Sea estimated on the basis of data from SatBałtyk system

For the Baltic Sea seasonal changes of the radiation budget at sea surface are calculated mainly based on numerical models [1], [9] or from hydro-meteorological ship observations

[6]. The question of the energy budget of the Baltic Sea was developed within BALTEX (Baltic Sea Experiment [2], [3], [8], [12]). Depending on the model used, the final results are different [1], [9]. Due to the lack of direct measurements of the radiation fluxes at the sea, assessment of accuracy of such obtained data is very difficult. Therefore, other methods are desirable. Such a method commonly applied for many years is to assess radiative budget with the use of satellite observations. Satellite data allow the estimation of the surface radiation budget covering large areas at the one moment in time. Such estimations have been developed for many years by CM SAF (Climate Monitoring Satellite Application Facility on Ocean and Sea Ice) and OSI SAF ([www.cmaf.eu](http://www.cmaf.eu), [www.osi-saf.org](http://www.osi-saf.org)). However, their study have a rather global character and possibilities of applying it to specific regions are often limited. In the SatBałtyk project [14], [15], local satellite algorithms have been used to create daily maps of radiation components. The algorithms were developed for the Baltic Sea [6], [16], [17] and improved within SatBałtyk project based on empirical data collected directly on the sea surface. The choice of satellite data source and appropriate algorithm determines an accuracy of estimated values. In the case of a local area like the Baltic Sea, the space and time resolution of data used is also critical. It is important that the used algorithms have been developed specifically for this region. On the basis of these algorithms maps of radiation budget components are created every day and presented on the website <http://satbaltyk.iopan.gda.pl/>. The SatBałtyk project has launched actinometric stations along the Polish coast and one station on the oil rig placed around 70 km from shore. The data from the station at sea are very valuable and unique for the Baltic. The data enable to validate daily radiation budget components for real marine conditions for each day of the year. For the Baltic this type of analysis has not been conducted. Most of this type of validation was done on the basis of data from land stations or by indirect methods. The validations based on empirical data collected on the vessel have been made for instantaneous values [16].

The aim of this work is to show possibility of using satellite data to estimate the daily average of the surface radiation budget components for the Baltic Sea and to show accuracy of the presented methods. An empirical verification of the final products were carried out against empirical data from pyranometers and pyrgeometers placed on the oil rig [19]. The paper focuses on the overall scheme of modeling with emphasis on the type used input data and assessment of accuracy. Detailed information about the formulas used can be found in the cited papers [6], [16], [19] and in SatBałtyk project reports [18]. Most of them are still being modified and presented results should be treated as a preliminary.

## METHODS AND INPUT DATA

Figure 2 shows the general block diagram for calculating components of the radiation budget. The input data come from the four independent sources: AVHRR (Advanced Very High Resolution Radiometer) radiometers working on board the

American TIROS-N/NOAA meteorological satellites (NOAA 15, 16, 17), SEVIRI (Spinning Enhanced Visible and InfraRed Imager) radiometer working on board MSG (Meteosat Second Generation 9, 10) and from prognostic model UMPL and from ecohydrodynamic model 3D CEMBS [4] as auxiliary data. The data from these sources are generated by SatBałtyk service. The maps from each source were imported into a 1 km resolution format (1280×1408 pixels), in Lambert Azimuthal Equal Area projection. The final products have the same format. The main parameters determining the fluxes (sea surface temperature (SST), cloudiness) are defined based on satellite information. The SST were determined by using the split-window method for AVHRR thermal channels 10.8 μm and 12 μm [5]. The number of SST maps was from six to eight for a day. Most of the maps were empty or only partially completed due to the presence of clouds. The cloudiness parameter was estimated separately for  $SW_d$  and for  $LW_d$  algorithms. This is due to the different properties of the  $SW$  and  $LW$  radiation. In the case  $SW_d$  cloudiness coefficient was determined based on HRV (High Resolution Visible 0.4-1.1 μm) channel from SEVIRI [6]. In the case  $LW_d$  an impact of clouds on radiation reaching sea surface was determined by the functions of cloudiness different for night and daytime. For the daytime the function depends on HRV [18]. For the night the function is dependent on the cloud fraction parameter. Cloud fraction is determined based on thermal channels from SEVIRI [10].

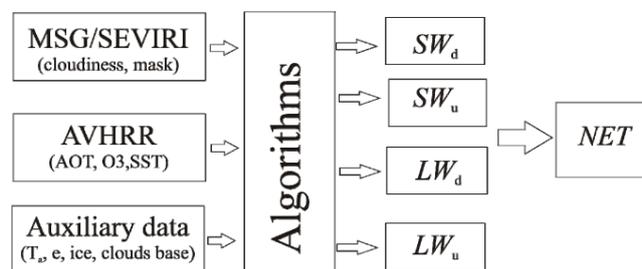


Fig. 2. Block diagram

The used models are described below. Most of them were developed or improved within the SatBałtyk project. The empirical data for satellite algorithms development were collected during cruises on s/y Oceania on the Baltic Sea in different seasons [16].

### DOWNWARD SHORTWAVE RADIATION $SW_D$

The  $SW_d$  model described in detail by Krężel et al. [6] is a physical parametrization in which transmittances of the atmospheric column are computed separately for cloudless and cloudy situations. The model was developed within SatBałtyk project. For cloudless part  $SW_{d,0}$  is computed by adding contributions from gaseous absorption, Rayleigh scattering and aerosols absorption and scattering. The model includes the absorption effect of the ozone (total ozone

column daily from TOAST<sup>1</sup>/NOAA or OMI/Aura), the attenuation by aerosols (AOT - aerosol optical thickness of the atmosphere from AVHRR) and the water vapour content in the atmosphere (from UMPL model). The clouds transmission calculations is based on a cloudiness parameter from visible HRV channel from SEVIRI.

$$SW_d = SW_{d,0} T_{cloud} \quad (1)$$

where  $SW_{d,0}$  - the irradiance for a cloudless atmosphere,  $T_{cloud}$  - cloud transmittance computed on the basis of satellite algorithm as a function of cloudiness coefficient.

The instantaneous maps are computed every 15 minutes. Map of dose is created based on these maps. Both products are in the SatBałtyk operational service. The average daily maps are computed from:

$$\overline{SW_d} = \frac{\sum_{t_{sunrise}}^{t_{sunset}} SW_d}{8640} \quad (2)$$

where 8640 – the number of seconds in a day.

#### UPWARD SHORTWAVE RADIATION $SW_u$

The  $SW_u$  flux is computed based on the function from Payne [11] modified by Rozwadowska [13] for Baltic Sea and developed in SatBałtyk project [19]:

$$SW_u = SW_d \begin{cases} A_{sea} & \text{for water} \\ A_{ice} & \text{for ice} \end{cases} \quad (3)$$

where  $A_{sea}$ ,  $A_{ice}$  - sea and ice albedo dependent on the transmission of the atmosphere, solar zenith angle and ice thickness concentration.

The information about the ice concentration is taken from ecohydrodynamic model 3D CEMBS [4]. For ice areas albedo is calculated for melting ice, snow and pure ice.

The instantaneous  $SW_u$  maps are computed every 15 minutes and based on these maps, the daily average product is created:

$$\overline{SW_u} = \frac{\sum_{t_{sunrise}}^{t_{sunset}} SW_u}{8640} \quad (4)$$

#### DOWNWARD LONGWAVE RADIATION $LW_d$

The  $LW_d$  model described in detail by Zapadka et al. [16], [17] is a semi-empirical formula which depends on water vapour  $e$  concentration in the atmospheric column and greenhouse gases, air temperature at the sea surface and cloudiness. Air temperature  $T_a$  and water vapour  $e$  were taken from the UMPL model. Cloud parameters were calculated separately for daytime (cloudiness from visible HRV channel

from SEVIRI) and for night (cloud fraction as a combination IR channels from SEVIRI).

$$LW_d = LW_{d,0} f(c_i) \quad (5)$$

where  $LW_{d,0}$  - the longwave irradiance for a cloudless atmosphere,  $f(c_i)$  - a satellite cloud function different for night and daytime. The instantaneous maps are computed every hour and based on these maps, the average daily product is created. The average daily downward longwave radiation is calculated based on:

$$\overline{LW_d} = \frac{\sum_{t=1}^{t=24} LW_{d,0}}{n} \quad (6)$$

where  $n$  – number of maps.

#### UPWARD LONGWAVE RADIATION $LW_u$

The  $LW_u$  flux is computed on the basis of sea surface temperature  $T_{SST}$  obtained as a combination of two IR channels from AVHRR [5] and ice surface temperature  $T_{ice}$ .

$$LW_u = \begin{cases} \varepsilon_w \sigma T_{SST}^4 \\ \varepsilon_{ice} \sigma T_{ice}^4 \end{cases} \quad (7)$$

where  $\varepsilon_w$ ,  $\varepsilon_{ice}$  - emissivity of the sea and ice surfaces respectively,  $T_{SST}$  - sea surface temperature from AVHRR and  $T_{ice}$  - ice surface temperature from the model.

The daily maps are created on the basis of all available scenes from up to five days back.

$$\overline{LW_u} = \frac{\sum_{i=1}^n LW_u}{n} \quad (8)$$

where  $n$  – the number of cloudless pixels.

Almost 2/3 of a year the sky over Baltic Sea area is covered by clouds. Obtaining the momentary temperature map for the whole Baltic Sea is not easy and requires analysing satellite maps from wider time range. Creating a satellite daily map for whole Baltic Sea often was done by merging many momentary maps. Used methodology gives a possibility to fill upward radiation map in 90% for almost 300 days a year [19].

### EMPIRICAL MATERIAL

In this work we are presenting the verification of the mean daily values of every components. The verification was carried out against an empirical material collected in the years 2013 – 2014 on an actinometrical station located on the oil rig (PetroBaltic Platform on the Baltic Sea). The fluxes were measured using CG3, CGR4 pyrgeometers and CMP3 pyranometers (Kipp&Zonen) with one minute time resolution. Data were averaged for every day. The database contains 250 days from September 2013 to September 2014 without February, half March, half June, July 2014. Gaps in the data are connected with a break in the measurements. However,

<sup>1</sup> Total Ozone Analysis using SBUV/2 and TOVS

the range of values can be considered as representative for the Southern Baltic.  $SW_d$  changes from 6 to 370  $Wm^{-2}$ ,  $SW_u$  from 4 to 26  $Wm^{-2}$ ,  $LW_d$  from 255 to 392  $Wm^{-2}$  and  $LW_u$  from 302 to 427  $Wm^{-2}$ . In the Fig. 3 there is presented oil rig location on the background of exemplary  $NET$  map for 20<sup>th</sup> September 2014. This figure shows a print screen taken from SatBałtyk website where one can find analyzed products. The analysis concerns one pixel within which the actinometrical station occurs.

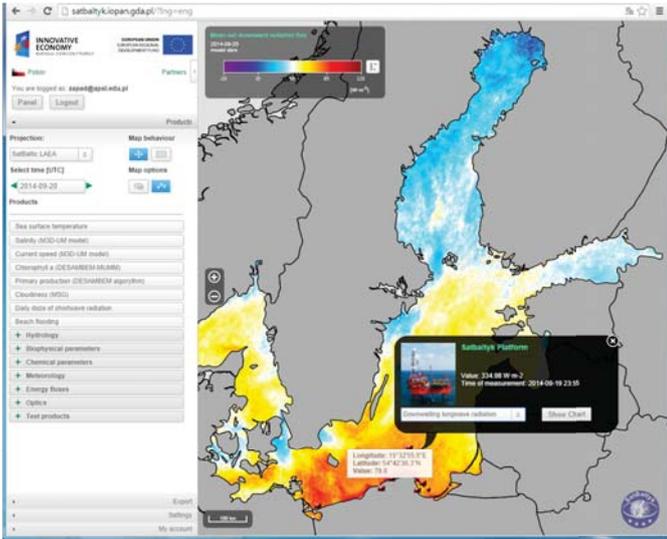


Fig. 3. Print screen from SatBałtyk website service with the oil rig localization

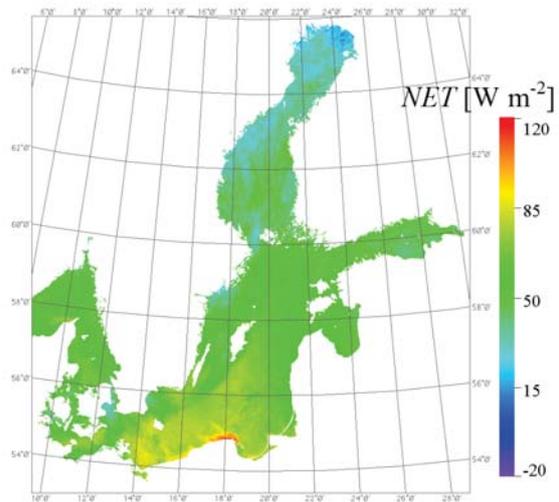
## RESULTS

The methods described in paragraph 2 allow to create maps of each component of the radiation budget. The examples of daily average maps of  $SW_d$ ,  $SW_u$ ,  $LW_d$ ,  $LW_u$  for a chosen day 20<sup>th</sup> September 2014 are shown in Fig. 4. As one can see in the case of shortwave products  $SW_d$ ,  $SW_u$  distribution of values is similar but  $SW_u$  values are much smaller. It is because  $SW_u$  depends on  $SW_d$ . In the case of longwave radiation fluxes values, distribution differs and depends on many factors such as sea and atmosphere temperature, water vapour pressure and cloudiness.  $LW_u$  is usually smaller than  $LW_d$ . The daily map  $NET$  (Fig. 5a) for 20<sup>th</sup> September 2014 is a sum of four components where  $LW_u$  and  $SW_u$  are taken with the opposite sign. The smallest impact on the  $NET$  has  $SW_u$  but negligence of this variable could lead to big systematic errors especially in seasons of low sun position. The components are estimated with different accuracy. Comparisons between modelled and measured values of  $SW_d$ ,  $SW_u$ ,  $LW_d$ ,  $LW_u$  respectively are presented in Fig. 4 on the right side. The first plot concerns  $SW_d$ . In this case, the correlation coefficient between the modeled and measured values is very high:  $r=0.99$ . Model works correctly in the whole range of values (for all seasons) with a slight underestimation of the value (systematic error - 2  $Wm^{-2}$ ). In the case of  $SW_u$  the model works correctly for values below 20  $Wm^{-2}$ . It means that for high positions of the sun model overestimates. For  $LW_d$  the spread of values is the

biggest and the statistical error reaches 14  $Wm^{-2}$ . The main reason of so high error is wrong interpretation of cloudless and cloudy situations. Moreover, this version of model does not use information on the height of the clouds base. This is important in the case of used cloud fraction parameter especially for the night. The statistical error of  $LW_u$  reaching 7  $Wm^{-2}$  is associated with underestimating mask of clouds. This error occurs for the edge of clouds. For instantaneous values specially selected for cloudless sky  $LW_u$  is estimated with statistical error of 4  $Wm^{-2}$  [19].

Summarised results of empirical analysis are presented in Table I. The table contains statistical and systematic errors for each modelled component and correlation coefficients. Additional analysis was carried out for  $NET$  radiation balance. As one can see used model underestimates  $NET$ . Systematic error reaches value of -3  $Wm^{-2}$ . The relative error in relation to the average annual values for the whole Baltic is the highest for  $SW_u$  and reaches 25% and the lowest for  $LW_u$  - 2%. In other cases for  $SW_d$  this error reaches 9% and for  $LW_d$  - 6%.

a)



b)

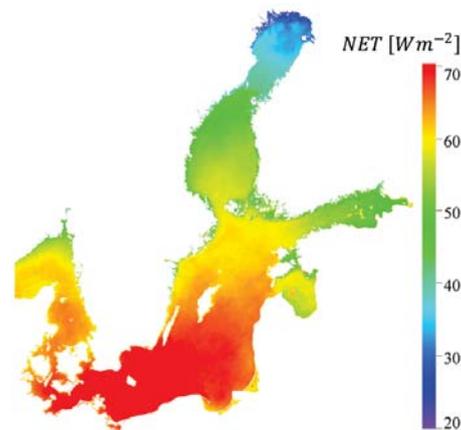


Fig. 5. The surface radiation budget  $NET$  for 20<sup>th</sup> September (a) and for 2013 year average annual (b)

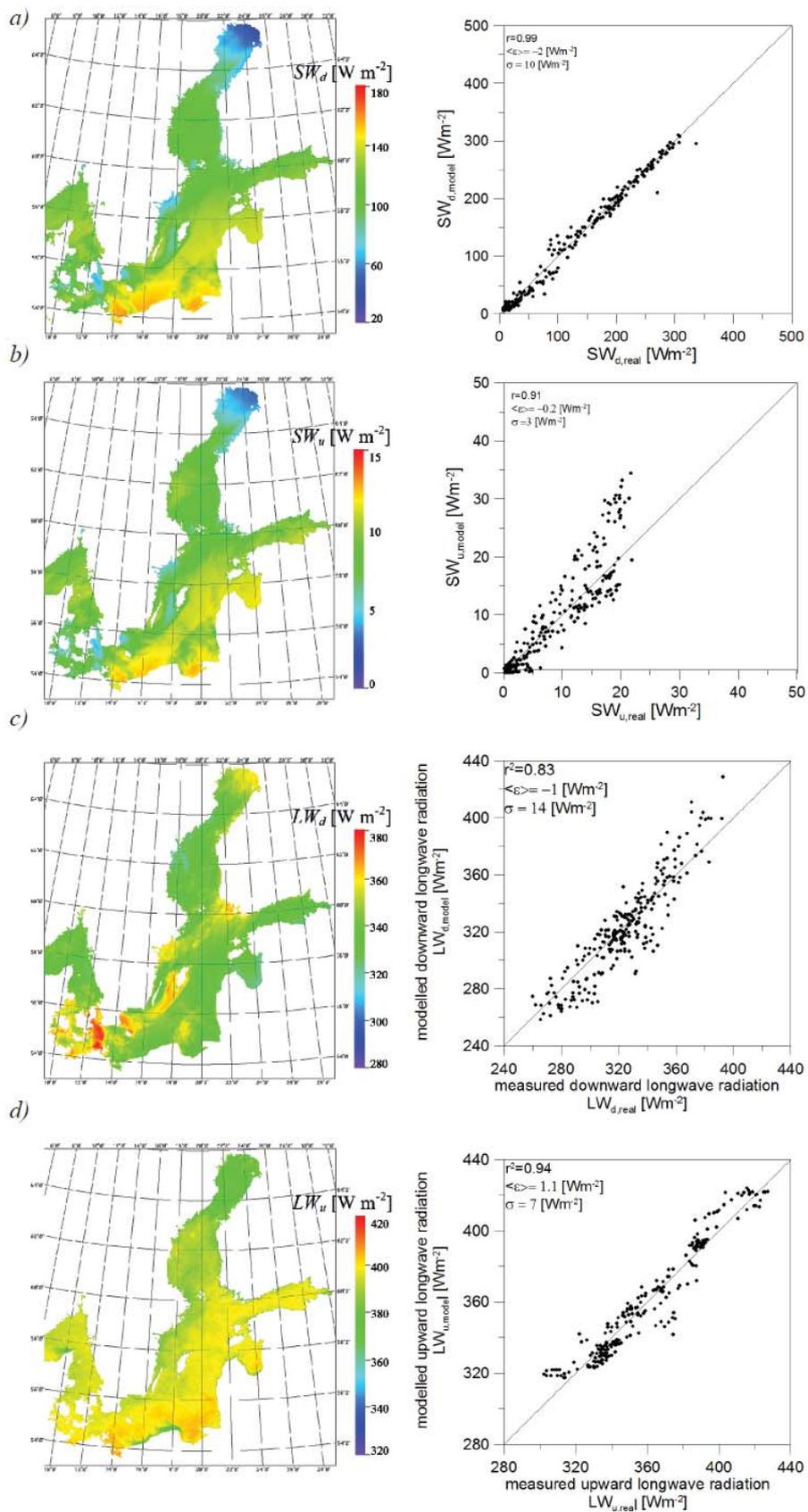


Fig 4. Daily average maps of  $SW_d$  a),  $SW_u$  b),  $LW_d$  c)  $LW_u$  d) for 20-09-2014 ,next to respectively for each flux comparisons between modelled and measured values

Table 1. Validation of the components of the radiation budget (Statistic error - standard deviation and Systematic error - BIAS, Mean - average values for all database)

flux	Mean	Syst. error [Wm <sup>-2</sup> ]	Stat. error [Wm <sup>-2</sup> ]	Correlation coefficient
SW <sub>d</sub>	139	-1.8	10.0	0.99
SW <sub>u</sub>	10	-0.2	3.1	0.91
LW <sub>d</sub>	323	-1.3	14.2	0.91
LW <sub>u</sub>	359	1.1	7.1	0.97
NET		-3.0	13.0	0.98

## CONCLUSIONS

In presented work possibilities of estimating of the net radiation flux at the sea surface and its components SW<sub>d</sub>, SW<sub>u</sub>, LW<sub>d</sub>, LW<sub>u</sub> for Baltic Sea elaborated within SatBałtyk project have been shown. We presented the scheme of creating daily maps and used algorithms. Described methods allow operationally to create maps of SW<sub>d</sub>, SW<sub>u</sub>, LW<sub>d</sub>, LW<sub>u</sub>, NET for one day and any time period (Fig. 4; Fig. 5). The algorithms operate on the basis of input data from SatBałtyk system. To a large part, their development depends on additional input data generated by SatBałtyk system.

The validation was carried out for every component based on empirical data collected at sea surface on the oil rig (Fig. 3). Such data are unique to the Baltic and the Earth. The empirical verification showed that presented method allows with a high accuracy to estimate radiation budget for the southern Baltic Sea region (table 1). Unfortunately, the lack of empirical data for typical areas covered with ice, does not allow assessment of upward components for northern regions of the Baltic. We realize that some algorithms still require some improvement, for example, of LW<sub>d</sub> and SW<sub>u</sub> models. In the first place the presented model requires improvement of a cloudiness function. In the SW<sub>u</sub> case the function of the solar elevation should be improved. LW<sub>u</sub> dependence needs more correct mask of clouds.

The radiation budget at sea surface can be a key parameter for climate monitoring and analysis of the Baltic region. Satellite data allow estimation of the surface radiation budget with high spatial and time resolution. As shown, SatBałtyk system produce data which may be used to mentioned climate analysis. The daily map allows to analyse of spatial change of the surface radiation budget NET for every day in a year (Fig. 5a). The average annual map (Fig. 5b) shows the spatial distribution NET for the whole year.

## ACKNOWLEDGEMENT

The study was supported by the Satellite Monitoring of the Baltic Sea Environment - the SatBałtyk project funded by the European Union through the European Regional Development Fund contract No. POIG 01.01.02-22-011/09.

## REFERENCE

1. L. Bengtsson, Numerical modelling of the energy and water cycle of the Baltic Sea, *Meteorology and Atmospheric Physics*, **77**, (2001); 9-17
2. BALTEX, BALTEX Phase I 1993–2002. State of the art report. *BALTEX Secr. Publ.*, **31**, (1977), 181 pp.
3. BALTEX, BALTEX Phase II 2003–2012. Science framework and implementation strategy, *BALTEX Secr. Publ.*, **34**, (2006a), 90 pp.
4. L. Dzierzbicka-Głowacka, M. Janecki, A. Nowicki, J. Jakacki, Activation of the operational ecohydrodynamic model (3D CEMBS)- the hydrodynamic part, *Oceanologia*, **55**(3), (2013), 519-541
5. A. Krężel, M. Ostrowski, M. Szymelfenig, Sea surface temperature distribution during upwelling along the Polish Baltic coast, *Oceanologia*, **47**(4), (2005), 415-432
6. A. Krężel, Ł. Kozłowski, M. Paszkuta, A simple model of light transmission through the atmosphere, *Oceanologia*, **50** (2), (2008), 125-146
7. R. Lindau, Energy and water budget of the Baltic Sea derived from merchant ship observation, *Boreal Environ. Res.*, **7**(4), (2002), 327-334
8. P. Lorenz, D. Jacob, BALTIMOS - a coupled modelling system for the Baltic Sea and its drainage basin, *Theor. Appl. Climatol.*, **118**, (2014), 715-727
9. A. Omstedt, Ch. Nohr, Calculating the water and heat balances of the Baltic Sea using ocean modelling and available meteorological, hydrological and ocean data, *Tellus*, **56A**, (2004), 400-414
10. M. Paszkuta, Cloud detection and cloud fraction estimation, *User manual - Final Report for SatBałtyk POIG.01.01.02-22-011/09*, UG Gdynia, (2015), (in Polish)
11. R.E. Payne, Albedo of the sea surface, *J. Atmos. Sci.*, **29**, (1979), 959-970
12. M. Reckermann, J. Langner, A. Omstedt, H. von Storch, S. Keevallik, B. Schneider, B. Arheimer, M. Meier, B. Hünicke, BALTEX - an interdisciplinary research network for the Baltic Sea region, *Environmental Research Letters*, **6**(4), (2011), 1-11
13. A. Rozwadowska, The variability in solar energy influx to the southern Baltic, Ph. D. thesis, *Gdańsk Univ.*, Gdynia, (in Polish), (1992)

14. B. Woźniak, K. Bradtke, M. Darecki, J. Dera, J. Dudzińska-Nowak, L. Dzierzbicka-Głowacka, D. Ficek, K. Furmańczyk, M. Kowalewski, A. Krężel, R. Majchrowski, M. Ostrowska, M. Paszkuta, J. Stoń-Egiert, M. Stramska, T. Zapadka: SatBaltic – a Baltic environmental satellite remote sensing system- an ongoing project in Poland. Part 1: Assumptions, scope and operating range, *Oceanologia*, **53(4)**,(2011), 897–924.
15. B. Woźniak, K. Bradtke, M. Darecki, J. Dera, J. Dudzińska-Nowak, L. Dzierzbicka-Głowacka, D. Ficek, K. Furmańczyk, M. Kowalewski, A. Krężel, R. Majchrowski, M. Ostrowska, M. Paszkuta, J. Stoń-Egiert, M. Stramska, T. Zapadka: SatBaltic – a Baltic environmental satellite remote sensing system- an ongoing project in Poland. Part 2: Practical applicability and preliminary results, *Oceanologia*, **53(4)**, (2011), 925–958.
16. T. Zapadka, B. Woźniak, J. Dera, A more accurate formula for calculating the net longwave radiation flux in the Baltic Sea, *Oceanologia*, **49(4)**, (2007), 449-470
17. T. Zapadka, A. Krężel, B. Woźniak, Longwave radiation budget at the Baltic Sea surface from satellite and atmospheric model data, *Oceanologia*, **50(2)**, (2008), 147-166
18. T. Zapadka, D. Stoltmann, The preliminary version of the algorithm to estimation the radiation budget components at the sea surface, *User Manual -Report for SatBałtyk* POIG.01.01.02-22-011/09, AP Słupsk, (XII 2012), (in Polish)
19. T. Zapadka, D. Stoltmann, M. Paszkuta, The daily upward longwave radiation flux for Baltic Sea from MSG/SEVIRI data, 2013 Eumetsat Meteorological Satellite Conference; 19th American Meteorological Society(AMS) Satellite Meteorology, Oceanography, and Climatology Conference, 16-20 September (2013), Vienna, Austria

## CONTACT WITH THE AUTHOR

T. Zapadka

Pomeranian University in Słupsk, Institute of Physics,  
Poland, ul. Arciszewskiego 22a, 76-200 Słupsk, Poland,  
e-mail: zapad@apsl.edu.pl

A. Krężel

University of Gdańsk, Institute of Oceanography, al.  
Marszałka Piłsudskiego 46, 81-378 Gdynia, Poland,  
e-mail: oceak@ug.edu.pl

M. Paszkuta

University of Gdańsk, Institute of Oceanography, al.  
Marszałka Piłsudskiego 46, 81-378 Gdynia, Poland,  
e-mail: ocompa@univ.gda.pl

M. Darecki

Institute of Oceanology Polish Academy of Sciences, ul.  
Powstańców Warszawy 55, 81-712 Sopot, Poland,  
e-mail: darecki@iopan.gda.pl

**POLAND**

# INFLUENCE OF PITTING CORROSION ON FATIGUE AND CORROSION FATIGUE OF SHIP AND OFFSHORE STRUCTURES, PART II: LOAD - PIT - CRACK INTERACTION

Marek Jakubowski, Assoc. Prof.  
Gdańsk University of Technology, Poland

## ABSTRACT

*In the paper has been discussed influence of stresses on general corrosion rate and corrosion pit nucleation rate and growth, whose presence has been questioned by some authors but accepted by most of them. Influence of roughness of pit walls on fatigue life of a plate suffering pit corrosion and presence of the so called „non-damaging” pits which never lead to initiation of fatigue crack, has been presented. Possibility of prediction of pit-to-crack transition moment by two different ways, i.e. considering a pit a stress concentrator or an equivalent crack, has been analyzed. Also, influence of statistical distribution of depth of corrosion pits as well as anticorrosion protection on fatigue and corrosion fatigue has been described.*

**Keywords:** Pitting corrosion, fatigue, ship structures, offshore structures.

## INTRODUCTION

Pitting corrosion constitutes the extreme form of material surface roughness due to corrosion. Melchers [1] has stated that influence of pitting corrosion on strength of well maintained structures such as ship ones is very low, but reduction of thickness due to general corrosion is much more important. However, according to common opinion, pits, if only present on material surface, constitute potential spots of fatigue crack initiation. The fact has been often, but not always, observed in research and practice [2]. At higher stress, pits in spot of corrosion crack initiation may not be present because of too short duration time. However presence of pits on material surface does not guarantee that initiation of cracks would occur in spots of the pits [3,4]. Thus the sequence of events: „first pit – then crack – and finally complete damage of a structure” not always takes place. There are sometimes exceptions but a great importance of the sequence of events can not be questioned.

Most publications concerning mutual pit-crack relation is focused on aircraft industry problems. Therefore they mainly concern Al-alloys with low resistance to seawater corrosion, especially of 2000 series [3, 5-7], or 7000 series [2, 8-10] of the alloys. As far as steels are concerned, majority of research papers deals with steels applied in machinery industry, i.e. medium- carbon steels after different heat treatment [11-15], or even high -carbon steels [16,17] or a martensitic stainless

steel [18]. Only a few papers concern weldable low-alloy, low- carbon steels applicable to ship and offshore structures [4, 19-22]. Although this paper concerns mainly ship and offshore structures, a limited amount of data on relevant materials makes that a model to be used for predicting fatigue life of such structures must be partly based on conclusions drawn from the tests on materials not applicable to ship and offshore structures. The existing models deal with aircraft materials and structures [6,12,23,24]. In the most complete model, initially proposed by Goswami T.K. and Hoepner D.W. (1995), and then adopted by Shi and Mahadevan [24], fatigue process runs through seven phases (Fig. 1).

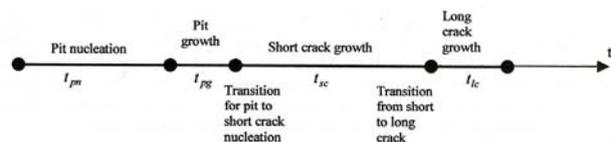


Fig. 1. Seven phases of pitting corrosion fatigue process [24]

In this approach 4 stages, each lasting a definite period, and 3 transient states: transition from pit nucleation to its growth, from pit to short crack and from short crack to long one, are considered. Phases of short crack growth and long

crack growth are assumed to be separated from each other because of different mathematical description of short and long cracks propagation. Total pitting corrosion fatigue life of a structure, when its damaging crack is initiated in a pit, can be calculated by summing the following four phases:

$$t_f = t_{pn} + t_{pg} + t_{sc} + t_{lc} \quad (1)$$

- $t_{pn}$  - pit nucleation duration time;
- $t_{pg}$  - pit growth duration time;
- $t_{sc}$  - short crack growth duration time;
- $t_{lc}$  - long crack growth duration time;

All the above mentioned phases and transition states, or only some of them, are described deterministically [6,12] or probabilistically [23,24]. Akid et al. [25], does not consider the pit- to- crack transition to be a transition state in the above presented sense but a stage which lasts a definite period. In such approach, the pit-to-crack transition time ( $t_{p \rightarrow c}$ ) should be included into the above presented model described by Eq.1 which takes the following form:

$$t_f = t_{pn} + t_{pg} + t_{sc} + t_{lc} + t_{p \rightarrow c} \quad (2)$$

Phenomena, mechanism and partly the modeling of particular pitting corrosion phases has been discussed previously [26]. This paper is focused on influence of mechanical loads on pit growth kinetics as well as a criterion for transformation of pit into fatigue crack, importance of statistical distribution of pit depth with respect to corrosion fatigue life, as well as importance of anticorrosion protection in these conditions, and influence of grooving corrosion - a specific kind of pitting corrosion - on fatigue of welded joints. All the phenomena discussed in this paper are controlled mainly by electrochemical action of the environment together with an additional influence of fatigue loads.

## INFLUENCE OF STRESSES ON GENERAL AND PITTING CORROSION

Although this paper concerns interrelation between pitting corrosion and fatigue process, there is certain, previously described in [26], relation between general corrosion and pitting one. It is common knowledge that dynamic (monotonic or cyclic) stresses, and especially strains, increase corrosion rate of passive metallic materials such as Al-alloys and corrosion resistant steels, mainly by rupture of their passive film. However, the same was demonstrated by Evans [27] for low-carbon steel exposed to cyclic stress action in chloride solution, where the steel was covered by a very weak passive film. Corrosion rate was increasing together with increasing cyclic stresses even below fatigue limit where plastic strains are close to zero. It was probably possible due to high loading frequency, hence specimen's surface strain rate was also high. In service, ship's hull is usually exposed first of all to low-frequency loads and only from time to time to a greater change in the mean load level. Some observations indicate that every stress increase and – in smaller degree – the stress drop accelerates corrosion

as a transient effect only [28]. Melchers [28] stated that at high level of strains in plate surface, rust layers are able to fall down periodically. At extremely high strains a continuous removal of rust may occur and the initially high corrosion rate may be maintained for a significant part of life of a structure. At lower strain levels rust is removed only from time to time, hence only a transient increase of corrosion rate may be expected. Melchers concluded that “the above mechanisms appears to be likely to occur only in highly strained regions, at strains well in excess of what would normally be expected in operational ships”.

Kobzaruk et al. [29] tested a low-carbon steel exposed to action of natural sea environment (1,8% salinity, 5500 h exposure) both without stresses and under cyclic stresses (1 cycle per day). Corrosion of stressed coupons was faster, but more uniform with a lower number of less sharp pits than that of non-stressed ones. The fact was attributed by him to a stress-induced removal of corrosion products and marine growths, that led to a more uniform surface of the stressed specimens.

Various authors formulate different opinions about influence of stresses and strains on pitting corrosion.

Booth et al. [30] observed that the weld toe grinding, which very effectively increases fatigue life in the air for welded joints made of a high-strength low-alloy (HSLA) offshore steel, only slightly improved corrosion-fatigue life of the joints in sea water. This phenomenon resulted from pitting corrosion of the grinded surface. The mentioned slight increase of corrosion-fatigue life corresponded to a short time necessary to generate pits deep and sharp enough to become crack initiation sites. They concluded that for this reason an increase of corrosion fatigue life due toe grinding is almost independent on the stresses. It means that they consider pitting corrosion rate as almost independent on stresses. For a high-carbon steel [17], under short exposure time to corrosion, growth rate of pits appears independent on applied mechanical loads. For longer exposure time values pit growth rate depends on applied loads. However, Linder and Blom [17] believe that other parameters such as flow velocity of corrosive solution and its temperature affects to a greater extent growth rate of pits. Hence, in this case the influence of cyclic stresses on pit growth rate may be neglected in engineering applications. Majority of authors, however, note a role of stresses and/or strains in pitting corrosion process.

For an austenitic steel [18] plastic strains associated with fatigue loading evidently accelerated pit nucleation in slip bands. For non-alloyed steels, influence of applied cyclic strains on pit nucleation period is not so distinct or is even negligible. Akid et al. [25] stated that the first pit appearance which was identified to be a breakdown in a moderately passive layer on steel surface, was controlled rather by electrochemical phenomena acting on metal-electrolyte interface than degree of mechanical damage. In established time periods, the mechanical damage degree was increasing along with increasing number of deformation cycles resulting from increasing load frequency ( $f$ ). Akid et al. stated that the pit initiation period  $t_i$  has been independent on the frequency  $f$ . The present author has reanalyzed the data read from an appropriate diagram published in [25] and

stated that a weak relationship did occur and the period  $t_i$  may be considered approximately proportional to  $f^{-0.105}$ . It means that though electrochemical control of pit nucleation dominates, mechanical loading may also play a certain role but it is not primary one. As stated in [20], the corrosion pits growing without of stresses, in an offshore structural steel, had a form close to circular on specimen surface. However along with fatigue process progress the form of the pits changed from circular to elliptical one with its main axis always perpendicular to direction of applied stresses. It proves that the stresses affected growth of pits not nucleation.

In a medium-carbon steel [13], driving force for pitting, i.e. corrosion current density inside the pit, is dependent on a stress level and its state. The pit current density, hence consequently the pit growth rate drops if the pit approaches a structural barrier of the growth, i.e. when plastic deformation range drops ahead of defect. The pit current density increases by an order of magnitude when the value of applied shear strains increases from 0.49% to 1,2%, typical values for low cycle fatigue. Nakajima and Tokaji [14] stated that the average depth of pits in a similar steel under fatigue corrosion is greater than under stress-free corrosion and increases along with stress level increasing. They proposed the following relation between growth of pit depth ( $a_p$ ) and action of stress ( $\sigma$ ):

$$a_p = (0.028\sigma + 0.404)t^B \quad (3)$$

$B = 0.199 - 0.224$   
 $a_p : \mu\text{m} ; t : \text{min}$

Other tests [6] of a similar medium-carbon steel showed that the pit growth rate at 100 MPa stress amplitude was from 5 to 10 times greater than that in load-free conditions. This order of magnitude of the pit growth rate acceleration by stresses is in agreement, for the same stress amplitude, with the value of 7 resulting from Eq. (3).

In another corrosion fatigue tests [11] on a medium-carbon steel, pits were initiated in the very early stage of the tests, i.e. approximately after 5000 load cycles (the total corrosion fatigue life was in the range from 20000 to 500000 cycles), almost independently of stress level. In non-loaded specimens the nucleation period of pits was the same but the pits ceased from growing very early. Therefore the pit growth rate increased along with stress amplitude increasing. The higher stress amplitude the greater density of pits on specimen surface. Worth reminding that in the Kobzaruk's investigations [29] influence of cyclic stresses on the density of pits on steel surface was opposite, that was probably associated with the extremely low load frequency (1 cycle per day) as well as marine growth presence.

Akid et al. [25] tested a carbon steel under fatigue loading in a synthetic sea water. They determined a ten-times increased value of pit growth rate as a result of stress range increasing. However, two threshold ranges (two plateaux) were observed where pit growth rate was almost independent of applied load range. The first of them corresponded to fatigue limit or a little above it where elastic state of surface stresses prevails. The other plateau corresponded to large stresses and the shortest fatigue lives well below 105 cycles, where "it might

be considered that there is a trend towards a saturation of plasticity in the surface layers" [25]. Hence the increase of pit growth rate may be attributed mainly to the increasing range of plasticity in steel surface layers between the two plateaux. The last statement is not in compliance with some results of the tests which revealed a significant difference in pit growth rate between these observed in a load-free steel and a loaded steel below fatigue limit) [6,14].

For 2024-T3 Al-alloy [6], pit growth rate depended both on the stress amplitude ( $\sigma$ ) and stress frequency ( $f$ ). However the stress amplitude effect was distinctly dominant. The authors proposed the following formula for the growth rate of the pit depth ( $a_p$ ):

$$a_p = A(\sigma) t^B f^C \quad (4)$$

where :

C - a constant assumed equal to 0, but its real value amounts to 0.01,

B = 0.36

$A(\sigma) = 2.34 \cdot 1.014^\sigma$

Wang and Akid [15] stated that the corrosion at the inclusion-matrix interface in a high-strength, medium-carbon steel, after application of cyclic stresses, run more suddenly than in the stress-free state. It leads to pit generation in this place. On the basis of nine different tests, Akid et al. [25] stated that the time to initiation of corrosion fatigue crack, i.e. the total time for nucleation and growth of a pit and its transformation into crack, was strongly decreasing along with cyclic stress increasing. Kendo et al. [31] examined a ferrite-pearlite steel in which general corrosion was non-uniform, and with corrosion pits due to non-homogenous microstructure of the steel, as well as a ferrite-bainite steel with uniform general corrosion and almost free of pits. The initiation period of corrosion fatigue crack of the first steel (with pits) was strongly shortened along with increasing level of cyclic stresses, and that of another steel (without pits) was much longer and only weakly dependent on stress level. The above mentioned results [15, 25,31] as well as results for Al-alloys [8,32] show that pitting corrosion is accelerated by cyclic stresses.

## PIT-TO-CRACK TRANSITION

When a pit reaches its critical size, often called threshold one, in its bottom a surface fatigue crack is initiated. Almost all researchers are unanimous that pit's critical size is decreasing along with increasing cyclic stress level, see e.g. [5,6,10,11]. At low stress amplitudes a small number of relatively large pits is generated as a result of their nucleation and growth, and fatigue cracks are always initiated in the pits, whereas at higher stress amplitudes a large number of small pits is generated in slip bands, and the test duration time is probably too short to cause pits to reach their threshold size as a result of their growth process, therefore cracks start in slip bands and only from time to time in small pits. The above described behaviour cases were the same for the steel [11] and Al-alloy [2].

Some authors have introduced the notion of the so called non-damaging depth of pit which never leads to fatigue crack initiation. Goto and Nisitani [11] considered pit diameter to be pit threshold size, but noted that some pits whose diameter values were greater than threshold ones, did not undergo cracks because pit depth is a more important parameter than pit diameter. Most authors mainly consider pit depth despite difficulties met in measuring pit depth in real structures. The non-damaging pit depth values are very different even for similar materials: e.g. for weldable HSLA steels the values amounted to 110  $\mu\text{m}$  [19] or below 40  $\mu\text{m}$  [21]; for Al-alloys the smallest depth of a pit in which fatigue cracks were initiated, amounted to 60  $\mu\text{m}$  [2], 20÷30  $\mu\text{m}$  [8] or 2  $\mu\text{m}$  [5], hence the non-damaging pits for Al-alloys were smaller than for steels. The above mentioned scatter of the data suggests that not only the pit depth but also another factors may affect pit-to-crack transition.

Cornet and Dolan [16] investigated an influence of salt water temperature on corrosion fatigue life of two high-carbon steels. The life was increasing along with increasing temperature of the water, because at higher temperature values pits were more of a hemispherical shape and less sharp. Therefore pit sharpness should be also taken into account. The stress concentration at a pit is determined in global scale by general pit curvature, but on the local scale by the curvature strictly in the site where crack is initiated [3]. In order to examine a role of roughness of pit walls for crack initiation in a weldable HSLA steel, many pits were chemically polished to make their bottom more smooth [19]. The life of specimens with more smooth pits was longer by about 50% than those with rough (non-polished) pits.

In global scale pit's shape is characterized by the pit aspect ratio, i.e. the ratio of the pit depth ( $a_p$ ) and pit breadth (the diameter  $D$ ). The following formula has been proposed [19] for the stress concentration factor ( $K_t$ ) at a pit:

$$K_t = 1 + 1.25 \sqrt{\frac{2}{1 + \left(\frac{D}{a_p}\right)^2}} \quad (5)$$

In the formula the reciprocal of the pit aspect ratio appears. In the tests [19] the pit aspect ratio values changed in the range from 1/3.6 to 1/2.1 with their average value of about 1/3. Hence the stress concentration factor values changed in the range from 1.85 to 2.2 with their average value equal to 1.98. Smaller values of pit aspect ratio are often observed for steel after its very long exposure to sea environment action. For an Al-alloy the pit aspect ratio values were contained within the range from 0.5 (i.e. for hemispherical pit) to 3.7, and most pits were more deep than broad, hence the maximum value of stress concentration factor amounted to  $K_t = 2.75$ . The pit aspect ratio for Al-alloys shows a tendency to increasing along with increasing corrosion exposure time. It means that the real pit approximation proposed in the subject-matter literature, by using a hemispherical form of the same depth as that of real pit and the factor  $K_t = 2.25$ , appears to be a conservative approach in the case of steels, but non-conservative assumption in the

case of Al-alloys. An additional support of the conclusion is the fact that walls of pits in steels are more smooth than those in Al-alloys where there are many sharp corners resulting from a specific form of pit growing due to crystallographic tunneling [1]. Though it cannot be expected that micro-unevenness of pit walls plays a significant role in steels, but the fatigue life reduction by 30÷40% as a result of pit surface roughness [19] cannot be considered as negligible. The above considerations show that on the basis of the tests on Al-alloys only qualitative conclusions, but not quantitative ones, concerning pit-crack transition in steel structures, may be drawn.

The most general approach to the problem of critical depth of pit and its transition to crack was presented by Jones and Hoepner [33]. The approach was elaborated on the basis of Al-alloy tests, but it seems to be qualitatively correct also for other materials. Interaction between different factors seems to play an important role, among which the following can be counted:

- material micro-structure, its properties (especially fracture toughness) and thickness,
- pit depth,
- pit surface area,
- pit size and shape,
- vicinity of other pits.

However for practical purposes pits on metal surface are often considered simply to act as effective stress concentrators [3,9,19,34]. It seems to be well-founded in the case when the pits are generated by corrosion, but later cracks in them are initiated and grown as a result of purely mechanical fatigue as in the case of aircraft fuselage structures. If a crack in a pit is initiated under corrosion fatigue conditions in a liquid or humid environment, then not only the stress concentration factor but also electrochemical activity in pit's tip should be taken into account [18,35]. The approach based on the stress concentration factor seems especially productive in determining fatigue life of pit-corroded material by using S-N diagrams.

The stress concentration factors for pits, as determined with the use of Eq. (5), are independent of pit depth. However, the tests of HSLA steels [19] showed that the pit depth rather than its aspect ratio affects fatigue and corrosion fatigue properties of specimens. The life to initiation of corrosion fatigue crack and total corrosion fatigue life are decreasing when pit depth is increasing [5,10,11]. The fact suggests that the quantities may be mutually dependent through the stress intensity factor which constitutes the main parameter of linear-elastic fracture mechanics. Another reason for which the method is presently more and more often used, is the direct using of pit and crack dimensions in it, therefore it can be integrated with results of inspection of structures.

In this approach a corrosion pit, in the moment of its transition to crack, is simply taken as a semi-elliptical crack of the same length as the pit depth ( $a_p$ ) [5,6,10,12, 13,17, 23, 24]. A question may be put how a surface defect of three-dimensional geometry and a relatively blunt tip can be approximated by a sharp two-dimensional crack. However it is common knowledge that period of crack initiation from pit's tip is negligibly short

[5,10]. The equating of pit growth to crack propagation is more justified since Smith and Miller [13] have shown that a crack grew from bottom of surface notch of the depth  $D$  and bottom radius  $\rho$  to reach the depth  $l$  approximately equal to  $0.13(D\rho)^{1/2}$ , then the effective crack depth is equal to  $(D+l)$ . The value of  $l$  of the order of  $1.2 \mu\text{m}$  is obtained after substitution of typical values of  $D$  and  $\rho$ , equal to  $20 \mu\text{m}$  and  $4 \mu\text{m}$ , respectively, where the assumed value of the radius  $\rho$  accounts probably for pit bottom micro-geometry. It is rather a non-important component in crack growth calculations when pit-to-crack transition is considered [13].

Kawai and Kasai [12] applied the approach in which a pit is considered equivalent to crack, to fatigue limit determination for material with corrosion pits. They approximated Kitagawa's diagram which represents a relation between fatigue limit and crack length, by using a simple El-Haddad's relation, and determined a relation between single pit depth and fatigue limit for low-carbon steel, medium-carbon steel and a high-alloy corrosion-resistant steels. The determined relations were in a satisfactory conformity with test results. However, the discussed approach in which a pit is considered as equivalent to crack, is applied much often to procedures based on linear fracture mechanics.

Range of the stress intensity factor ( $\Delta K$ ) value in the moment of pit-to-crack transition must be greater than, or equal to, the threshold value for propagation of fatigue or corrosion fatigue crack, ( $\Delta K_{th}$ ):

$$\Delta K \geq \Delta K_{th} \quad (6)$$

Kawai and Kasai proposed a relation (that can be found in [13]) for the critical depth of a pit from which a fatigue crack can subsequently grow as a function of the stress range,  $\Delta K_{th}$ , and the fatigue limits for smooth and of corroded material.

Relation (6) is valid under the assumption that initially pit grows faster than a crack in pit's root [13]. The pitting corrosion process controls pit nucleation and growth as well as a dominating part of pit-to-crack transition phase [25]. Pitting corrosion is then much faster than crack growth. Preliminary tests performed by Akid et al. [25] showed that current density in a pit dropped when crack has been initiated in it. The current density drop may be used for detecting the moment of crack initiation and determining the threshold stress intensity factor [25]. However it is not clear if the current density drop (i.e. drop of corrosion rate inside the pit) results from the crack initiation or is a condition which makes initiation of the crack inside the pit, possible. Therefore must be introduced another important condition taking into account that there is a competition between the pit growth and the crack growth. In the moment of pit-to-crack transition the pit growth rate ( $da_p/dt$ ) must be smaller than the crack propagation rate ( $da/dt$ ) [8,10]:

$$da_p/dt < da/dt = f \cdot da/dN \text{ at } \Delta K = \Delta K_{th} \quad (8)$$

The above discussed approach is questioned by Akid et al. [25] since the pit-to crack transition is controlled not only by stress intensity but also electrochemical conditions. The

linear-elastic fracture mechanics parameter  $\Delta K_{th}$  is determined from examination of long cracks but: (i) local electrochemical conditions inside such cracks greatly differ from the conditions inside short cracks and pits, and (ii) the parameter in question concerns elastic state whereas a micro-plasticity state is associated with a pit which forms a notch.

Certainly, cracks which propagate from pits are micro-structurally, physically and chemically short ones which, as it is common knowledge, propagate faster than long ones at the same value of the stress intensity factor  $\Delta K$ , and their propagation is possible below the  $\Delta K_{th}$  value for long cracks. Thus  $\Delta K_{th}$  which appears in the eq. 8, would mean the threshold value for short cracks [8,10]. It seems that, in order to satisfy the condition, the crack propagation rate could be determined according to an appropriate crack growth curve for long cracks in the range of validity of Paris law but extrapolated to the threshold and near-threshold range. Such approach seems to be strongly conservative but simple. The conditions determined by Eq. (6) and/or (8) make it possible to predict an initial crack length necessary for determining the crack propagation period with the use of the methods of linear-elastic fracture mechanics. However the condition  $da_p/dt < da/dt$  is to be fulfilled during the whole crack propagation period. If it is not satisfied then pits will be formed on the crack path in the early phase of crack propagation, that was observed in some tests [8,36].

Ishihara et al. [6] applied the Murakami's approach to calculate the stress intensity factor for surface cracks of an irregular shape. The authors determined the following formula for the stress intensity factor value corresponding to the moment of pit-to-crack transition:

$$K_{p \rightarrow c} = 0.65 \sigma_a \{ \pi [(area)_{p \rightarrow c}]^{1/2} \}^{1/2} \quad (9)$$

where the areas of corrosion pits ( $area)_{p \rightarrow c}$  are those of pit traces left on the crack surfaces in the moment of their nucleation (for hemispherical pit this is the semi-circular area equal to  $\pi(a_p)^2/2$ ). The obtained  $K_{p \rightarrow c}$  values, independently on their scatter, were practically independent of a level of applied stresses and approximately equal to the threshold value of the stress intensity factor, obtained from the tests of propagation rate of corrosion fatigue cracks in the rotating bending conditions.

For 7075-T6 aluminum alloy [10] pre-pitted and then fatigued in the air, the fatigue life values determined from the tests generally agreed with the predicted ones when the mean pit size was assumed to be initial crack length. The fatigue life values predicted by assuming the initial crack length to be equal to the largest pit size, were smaller than measured ones. The authors stated that much greater number of pits having dimensions more close to their mean value than to maximum one, were present. Therefore the probability of crack initiation within a mean size pit is greater. Also, Medved et al. [8] informed that in many cases the largest pits did not initiate fatigue cracks.

## IMPORTANCE OF STATISTICAL PIT DEPTH DISTRIBUTION FOR CORROSION FATIGUE

For the modeling of pitting corrosion fatigue it is necessary to know statistical distribution of depth of pits on material surface because the pit depth distribution greatly affects statistical fatigue life distribution [5]. The main effort of researchers who analyze statistical pit depth distribution, is focused on depth distribution of the largest pits. However, pits of depths closer to their mean, are more frequent and hence more hazardous from the point of view of possible initiation of fatigue cracks [10]. Therefore not only the statistical pit depth distribution but also a rigorous quantitative analysis of probability of crack initiation in pits of different depth values [10] is important for assessing fatigue damage risk for structures

## INTERRELATION BETWEEN CRACK INITIATION DUE TO PITTING-CORROSION FATIGUE AND ANTICORROSION PROTECTION

In engineering practice, metal surfaces exposed to corrosive environment action are usually protected against corrosion. Cathodic protection effectively protects offshore structures against pitting corrosion [30,37] which usually leads to initiation of corrosion fatigue cracks in non-protected structures. The cathodic protection will much less effectively prevent initiation of corrosion fatigue cracks if certain pits are developed on the to-be-protected surface before the protection has been applied [38].

Kumakura et al. [21] exposed specimens made of KA32 steel (with blunt notches of  $K_t = 2$ ) for 20 days in the air outside laboratory, which was associated with sprinkling them twice a day, i.e. every morning and evening, in order to simulate steel material storing practice in shipbuilding industry. During the period the pits up to 40  $\mu\text{m}$  deep were generated. Then the specimens were suitably brushed and covered with a tar epoxy resin layer. In the higher nominal stress region, cracks in the protective coating were detected earlier than crack initiation in steel. Therefore, in spite of the coating, crack initiation in steel may have a corrosion fatigue nature and start probably from pits. In the of lower nominal stresses region and longer life values, fatigue cracks were initiated without salt water interaction in previously existing pits below the coating layer, and next the coating layer was cracking as a result of increasing crack opening displacement. Salt water could penetrate through the cracks in the protective coating, therefore, in contrast to crack initiation, crack propagation in the steel preceded under corrosion fatigue conditions through the pits produced during the initial corrosion exposure. The cracks can merge plural pits initiated and developed separately.

Anticorrosive coatings are able to prevent pitting corrosion effectively, hence they can cause delay of corrosion fatigue crack initiation. Defects of a porosity type in an Al-coating thermally sprayed could result in nucleation and growth of pits on damaged surface [39]. The observation seems to be valid also for other metal and non-metal coatings, because e.g. Kumakura et al. [21] stated that in the range of longer fatigue

life values some specimens had blisters in the coating layer (tar epoxy resin), this way the coating was gradually deteriorated its quality. In the situation, corrosion developed in the steel under the protective layer.

Sometimes, to delay nucleation of pits and slow down their growth a modification of corrosion environment is implemented. For a HSLA offshore steel [20] exposed to salt water action, presence of the anodic inhibitor (0.04 M  $\text{NaNO}_3$ ) did not prevent pitting corrosion due to sulfide inclusions, but the pits were less numerous and showed smaller growth rates than those developed in clear salt water. Addition of an oxidizing agent to salt water 1000 sec after application of a small anodic polarization to the pits on medium-carbon steel surface [13] resulted, after 1600 sec, in full repassivation of the pits which ceased to grow.

## GROOVING CORROSION VERSUS FATIGUE OF WELDED JOINTS

Among different factors, the stress concentration factor resulting from joint's geometry is considered a factor of the greatest impact on fatigue strength. The grooving corrosion phenomenon was described in the paper [26]. High stress concentration may be expected in the case when loading direction is perpendicular to grooving corrosion line. Yuasa and Watanabe [40,41] carried out tensile fatigue tests on welded butt, fillet and gusset joints made of a higher-strength steel, with grooving corrosion close to the weld edge, as well as the same tests on non-corroded joints. The initial corrosion period was equal to 3 months (Case 1) or 6 months (Case 2). Grooving corrosion rate slowed down after 3 months. Cross sections of the corroded joints are presented in Fig. 2. In the butt joints rather deep grooving corrosion (about 0.9 mm deep) occurred along line of the weld and close to its edge. Local corrosion grooves about 0.3 deep were observed close to the weld in the cruciform joints, but in the gusset joints almost no traces of grooving corrosion were found. The specimens were fatigued in the laboratory air atmosphere. The test results are presented in Fig. 3.

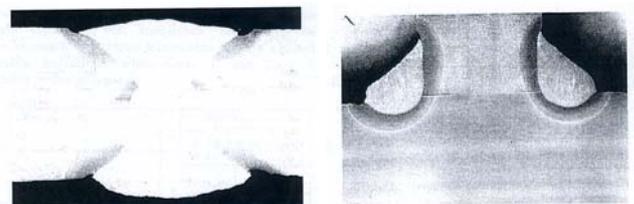


Fig.3 Cross sections of corroded specimens butt joint; cruciform joint [42].

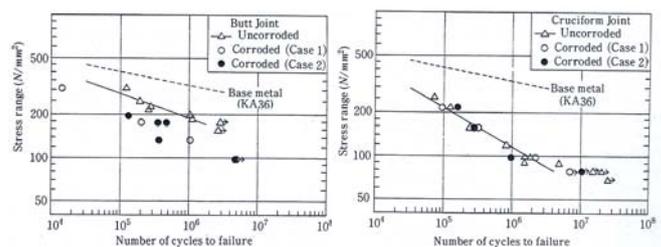


Fig.4 Results of fatigue tests in air of pre-corroded and non-corroded welded joints under tensile load of the stress ratio  $R=0.1$  and at the load frequency  $f = 0.1\text{Hz}$  [42].

It has appeared that only the fatigue strength of the butt joints decreased as a result of the initial corrosion. A stress analysis showed that in this case the mean value of the stress concentration factor increased from 2.13 to 3.83 due to corrosion. The stress concentration factor values in the non-corroded cruciform joints and gusset joints (fatigue results omitted in fig. 3) were relatively high and the corrosion was not able to sharpen the notches more (in reality it even decreased the mean value of stress concentration factor from 3.91 to 3.63). Thus the fatigue strength did not almost undergo any change as a result of the pre-corrosion.

The extended pre-corrosion period did not influence fatigue properties of the joints of any type. It seems reasonable to expect that in some cases the stress concentration factor at weld toe could be even reduced due to the prolonged grooving corrosion process. However in such cases a more important role may be associated rather with local reduction of plate thickness than stress concentration factor value. It depends on initial material thickness as well as loading mode.

The above given statements are correct for fillet joints with the weld leg greater than the so called critical dimension – term discussed in [42]. If the weld dimension is „subcritical” then the joint will undergo fatigue damage as a result of a fatigue crack initiated in weld root and propagating through the weld, where water has no access, hence stress concentration can not be changed under corrosion action. Then a very important role may be played by reduction of fillet weld leg, resulting from corrosion, but not by local notch sharpness at weld edge. Yuasa and Watanabe [43] modeled corrosion-induced material loss in fillet weld with the use of machining the plate and weld. Roughness of real corroded surfaces was considered unimportant because fatigue cracks were initiated in the weld root. Fatigue strength of specimens, expressed by nominal stresses, dropped markedly when weld and plate thickness has been reduced. When real, but not nominal, stress values were compared then the drop of fatigue strength of the joints, resulting from the decreased thickness of plate and weld, was much lower than in the case of nominal stresses. The above mentioned reduction of weld thickness affected also the stress concentration factor in weld root, therefore the strength of the joints with reduced thickness, expressed by real stress, showed a tendency to be somewhat smaller than that of the full - thickness joints. An inverse tendency observed in certain cases the authors attributed to the effect of releasing residual stresses as a result of reduction of weld and plate thickness.

## CONCLUSIONS

1. There is the almost common opinion [6,11,13-15,17,33] that growth rate of pits under stress is faster than without stress action, and that the rate increases along with applied load increasing. However in the publications the formulae for pit growth which take into account stress effect, were found only for an aluminum alloy and medium-carbon steel, both exposed to salt water for a short period under cyclic load of a relatively high frequency. A suitable formula for typical shipbuilding steels applicable in typical

working conditions, especially anaerobic ones after a long exposure period, is still lacking. The opinion of Melchers [28], in which stress effect importance in these conditions is depreciated, deals mainly with general corrosion, is only qualitative one and not based on test results.

2. There are different contradictory opinions concerning applied load influence on corrosion pits:
  - (a) stresses do affect pitting corrosion but the stress level is a factor of secondary importance [22];
  - (b) plastic stresses, resulting from loading, control pit nucleation [18] and growth [30]; this could mean that stress influence in elastic range is not a significant factor and is in line with argumentation of Melchers [28] concerning general corrosion;
  - (c) in the case of an aluminum alloy, when load frequency increases then pit growth rate also increases (this growth is small enough to be omitted in engineering applications) [6], but the rate evidently decreases for a HSLA steel [22];
  - (d) for medium-carbon steels: applied stress affects pit initiation but not its growth [14], or it does not affect pit nucleation but strongly affects its growth [6].
3. Some authors [2,5,8,19,21] determined the so called non-damaging pit depth, i.e. the maximum depth of pit which will never initiate a fatigue crack. The depth was greater for steel than for aluminum alloys, but for the same group of materials the depth differed significantly.
4. The critical pit depth in the moment of pit-to-crack transition decreases along with stress level increasing [2,6,10,11].
5. Pit-to-crack transition seems to be a complex process dependent on many variables [34], including: (i) material micro-structure and properties, (ii) pit depth, (iii) pit size and shape, including micro-shape, (iv) pit surface area and (v) proximity of other pits. In a simplified, more practical approach a pit may be considered a surface crack of the same length as the pit depth. The approach makes it possible to determine:
  - (a) fatigue limit in function of pit depth for a material with surface pits;
  - (b) total fatigue life of a material with surface pits, by using the methods of linear-elastic fracture mechanics.
6. The following working conditions should be checked in order to determine the critical pit depth:
  - (a)  $\Delta K \geq \Delta K_{th}$
  - (b)  $da_p/dt < da/dt = f \cdot dd/dN$  at  $\Delta K = \Delta K_{th}$
  - (c) to the relation (a) the threshold value of the stress intensity factor ( $\Delta K_{th}$ ) for short cracks should be inserted, or  $\Delta K_{th} = 0$  should be assumed; in the latter case the condition (b) will be decisive of pit depth critical value.
7. Statistical pit depth distribution was described by some authors, but the descriptions were focused on the distribution of maximum values only, whereas it seems [10] that the mean value of pit depths controls the mean value of fatigue life of specimens.

8. Cathodic protection effectively prevents pitting corrosion, but already existing pits developed during storing steel material in open air or during a break between protection periods can greatly reduce fatigue life of a protected structure.
9. Fatigue damages affect pitting corrosion: (a) degradation of protective coatings may lead, in corrosion fatigue conditions, to occurrence of blisters in the coatings and formation of pits under a metal or non-metal coating, (b) fatigue crack initiation in a pit results in a transient drop of corrosion current inside the pit.
10. Tar epoxy resin coating greatly extends corrosion fatigue life of steel [21], but does not guarantee full protection for it because:
  - (a) at higher nominal stress amplitude values (i.e. shorter life values) the coating layer frequently suffered cracks before a crack has been initiated in the steel; it means that the crack initiation and propagation processes in steel were running in corrosion fatigue conditions;
  - (b) at lower nominal stress amplitude values cracks were initiated in the pits developed under the protective coating during the pre-exposure of steel material in open air and after sprinkling it twice a day; then the coating was fractured as a result of increasing crack opening displacement and water penetrated into the cracks which continued to propagate but now in corrosion fatigue conditions.
11. In the case of the joints with a lower stress concentration, i.e. butt joints, grooving corrosion occurring in heat affected zone close to weld edge resulted in a marked reduction of fatigue strength. In the case of the joints with a higher stress concentration, i.e. cruciform and gusset joints, the notch at weld toe become even blunted by the grooves, hence the corrosion did not decrease fatigue strength of the joints. It seems that the corrosion grooves, similarly to pits, may be considered straight -line- front cracks or through-thickness cracks.
3. van der Walde K., Hillberry B.M., Initiation and shape development of corrosion nucleated fatigue cracks, *International Journal of Fatigue*, 2007, Vol.29, pp.1269-1281
4. Lu B.T., Luo J.L.: Crack initiation and early propagation of X70 steel in near-neutral pH groundwater, *Corrosion*, 2006, No 8, Vol.62, pp. 723-731
5. Dolley E.J., Lee B., Wei R.P.: The effect of pitting corrosion on fatigue life. *Fatigue & Fracture of Engineering Materials & Structures*, 2000, Vol.23, pp.555-560
6. Ishihara S., Saka S., Nan Z.Y., Goshima T., Sunada S., Prediction of corrosion fatigue lives of aluminium alloy on the basis of corrosion pit growth law, *Fatigue & Fracture of Engineering Materials & Structures*, 2006, Vol.29, pp.472-480
7. Pidapatri R.M., Patel R.R.: Correlation between corrosion pits and stresses in Al alloys. *Materials Letters* 2008, Vol.62, pp.4497-4499.
8. Medved J.J., Breton M., Irving P.E.: Corrosion pit size distribution and fatigue lives – a study of the EIFS technique for fatigue design in the presence of corrosion. *International Journal of Fatigue* 2004, vol.26, pp.71-80.
9. Pao P.S., Gill S.J., Feng C.R.: On fatigue crack initiation from corrosion pits in 7075-T7351 aluminum alloy, *Scripta Mater.*, 2000, Vol. 43, 391-396
10. Sankaran K.K., Perez R., Jata K.V.: Effect of pitting corrosion on the fatigue behavior of aluminum alloy 7075-T6: Modeling and experimental studies, *Materials Science and Engineering*, 2001, Vol. A297, pp223-229 10
11. Goto M., Nisitani H.: Crack initiation and propagation behavior of a heat-treated carbon steel in corrosion fatigue. *Fatigue & Fracture of Engineering Materials & Structures*, 1992, Vol.15, pp.353-363 11
12. Kawai S., Kasai K.: Considerations of allowable stress of corrosion fatigue (focused on the influence of pitting), *Fatigue & Fracture of Engineering Materials & Structures*, 1985, Vol.8, pp.115-127 12
13. Miller K.J., Akid R.: The application of microstructural Fracture Mechanics to various metal surface states. *Proc. Royal Society A*, 1996, Vol. 452, 1411-1432 13
14. Nakajima M., Tokaji K.: Fatigue life distribution and growth of corrosion pits on a medium carbon steel in 3%NaCl Solution. *Fatigue & Fracture of Engineering Materials & Structures*, 1995, Vol.18, pp.345-351 14
15. Wang Y., Akid R. Role of nonmetallic inclusions in fatigue, pitting and corrosion fatigue. *Corrosion*, 1996, Vol. 52, p.92-104 15

## ACKNOWLEDGEMENTS

This work has been performed in the scope of the project RISPECT “Risk-Based

Expert System for Through-Life Ship Structural Inspection and Maintenance and New-Build Ship Structural Design” which has been financed by the UE under the contract SCP7-GA-2008-218499

## REFERENCES

1. Melchers R.E.: Pitting corrosion of mild steel in marine immersion environment – Part 1: Maximum pit depth. *Corrosion*, 2004, Vol.60, No 9, pp.824-836.
2. Chlistovsky, Hefferman P.J., DuQuesnay D.L., Corrosion fatigue behavior of 7075-T651 aluminum alloy subjected to periodic overloads, *International Journal of Fatigue*, 2007, vol.29, pp.1941-1949

16. Cornet i., Golan S.: Influence of temperature on corrosion fatigue. *Corrosion*, 1959, No 5, p.262 16
17. Linder J., Blom R.: Development of a method for corrosion fatigue life prediction of structurally loaded bearing steel. *Corrosion*, 2001, Vol. 57, No.5, pp.404-412 17
18. Qian Y.R., Cahoon J.R.: Crack initiation mechanisms for corrosion fatigue of austenitic stainless steel. *Corrosion*, 1997, Vol.53, No.2, pp.129-135 18
19. Ahin S.-H. Lawrence Jr. F.V., Metzger M.M.: Corrosion fatigue of an HSLA steel. *Fatigue & Fracture of Engineering Materials & Structures*, 1992, Vol.15, pp.625-642 19
20. Boukerrou A., Cottis R.A.: Crack initiation in the corrosion fatigue of structural steels in salt solutions. *Corrosion Science*, 1993, Vol.35, pp.577-585 20
21. Kumakura Y., Takanashi M., Fuji A., Kitagawa M., Ojima M., Kobayashi Y.: Fatigue strength of coated steel plate in seawater. *Proc. Ninth Int. Offshore and Polar Engineering Conference, Brest, France, May 30 – June 4, 1999, Vol.4, pp. 108-113.* 21
22. Maximovich, Kobzaruk: Initiation and propagation of low-cycle fatigue cracks in 15HN5DMF steel in seawater. *Physical Chemical Mechanics of Materials*, 1985, Vol.20, No 5, pp.16-20 (in Russian) 22
23. Shi P., Mahadevan S.: Probabilistic corrosion fatigue life prediction. 8th ASCE Specialty Conference Probabilistic Mechanics and Structural Reliability. 2000 23
24. Zhang R., Mahadevan S.: Reliability based reassessment of corrosion fatigue life. *Structural Safety*, 2001, Vol.23, pp.77-91 24
25. Akid R., Dmytrakh I.M., Gonzales-Sanchez J.: Fatigue damage accumulation: the role of corrosion on the early stages of crack development. *Corrosion Engineering, Science and Technology*, 2006, Vol.41, No.4, pp.328-335. 25
26. Jakubowski M., Influence of pitting corrosion on fatigue and corrosion fatigue of ship structures. Part 1: Mechanisms and modeling of pitting corrosion of ship structures. *Polish Maritime Research*,
27. Evans U.R., Tohopandui Simnad M., The mechanism of corrosion fatigue of mild steel. *Proceedings of the Royal Society, Series A*, 1947, vol.188, pp.372-392. 27
28. Melchers R.E., Development of new applied models for steel corrosion in marine applications including shipping. *SAOS*, 2008, Vol.3, No2, pp.135-144. 28
29. Kobzaruk K.A.V., Marichev V.A.: Corrosion and corrosion fatigue resistance of steels in real marine and in laboratory. *Physical Chemical Mechanics of Materials*, 1981, Vol.16, No 2, pp.15-21 (in Russian) 29
30. Booth G.S.: Constant amplitude corrosion fatigue strength of welded joints. *Fatigue in Offshore Structural Steels (Proc. of a Conference London, 24-25 Feb., 1981), Paper No 2, pp.5-16* 30
31. Konda N., Suzuki S., Tada N., Kho Y., Kazushige A., Watanabe E., Yamamoto M and Yaima H.: Effect of microstructure on fatigue properties of steel in seawater – development of steels for high resistance to fatigue in ships, Part 2. *J Soc. Naval Architects of Japan*, 2001, Vol.191, pp.229-237. 32
32. Rajasankar J., Iyer N.Y., Gopinath S., Probabilistic modeling of fatigue crack initiation from pits and pit clusters in aluminum alloys, *Corrosion Engineering, Science and Technology*, 2007, Vol.42, No.3, pp.260-265 33
33. Jones K, Hoepfner D.W., Prior corrosion and fatigue of 2024-T3 aluminum alloy, *Corrosion Science*, 2006, Vol.48, pp.3109-3122 34
34. Grimes D., i in.: Corrosion fatigue strength of welded K-joints and HSLA-cast steel hybrid K-nodes at component-similar scale. *Steels in Marine Structures, Amsterdam 1987, pp.465-478.*
35. Ebara R.: Corrosion fatigue phenomena learned from failure analysis, *Engineering Failure Analysis*, 2006, Vol.13, pp.516-525 36
36. Sonsino C.M., Lipp K., Lachman E.: Corrosion fatigue of welded high-strength cast and structural steel joints under constant and variable amplitude loading. *Proc. Fifth Int. Offshore and Polar Conference, The Hague, June 11-16, 1995, pp.53-58.* 37
37. Jootsen M.W., Salama N.N.: Corrosion fatigue of aluminum sprayed, high strength steel immersed in seawater. *Material Performance*, 1984, Vol.23, No 7, pp.22-26 38
38. Booth G.S.: Techniques for improving the corrosion fatigue strength for plate welded joints. *Steel in Marine Structures, Amsterdam, 1987, pp.747-757* 39
39. Connolly B.J., Meng Q., Moran A.L., McCaw R.L., Mechanical and pre-corroded fatigue properties of coated aluminum aircraft skin system as function of various thermal spray processes, *Corrosion Engineering, Science and Technology*, 2004, Vol.39, No.2, pp.137-142 40
40. Yuasa M, Watanabe T.: Fatigue strength of corroded weld joints. *J. Society of the Naval Architects of Japan*, 1994, Vol.176, pp.481-490 (in Japanese) 41
41. Yuasa M, Watanabe T.: Fatigue strength of corroded weld joints. *ClassNK Technical Bulletin*, 1996, Vol.14, pp.51-61 (in English) 42

42. Gurney T.R.: Fatigue of welded structures. (monograph)  
Cambridge University Press, 1968 43
43. Yuasa M., Watanabe T.: The influence of corrosion wastage  
on the fatigue strength of fillet welded joints. NK. Tech.  
Bulletin, 1998, pp.21-31. 44

## **CONTACT WITH AUTHOR**

Marek Jakubowski

Gdańsk University of Technology  
Ocean Engineering and Ship Technology  
Narutowicza 11/12 Str.  
80-233 Gdansk  
POLAND

## COMPARING GUIDELINES CONCERNING CONSTRUCTION OF THE S-N CURVE WITHIN LIMITED FATIGUE LIFE RANGE

Przemysław Strzelecki, PhD

Janusz Sempruch, Prof.

Krzysztof Nowicki, PhD

University of Technology and Life Sciences in Bydgoszcz, Poland

### ABSTRACT

The article collates guidelines concerning experimental construction of the S-N fatigue curves within a limited fatigue life range. An attempt is made to compare these guidelines, based on experimental data recorded during rotating bending of a notched specimen made of 42CrMo4. The recorded differences in fatigue life values between the constructed curves reach the maximum of 12.2%. According to the above guidelines, the number of tests in particular test series varies from 6 to 28. Based on the performed analysis a conclusion was made that the increase in the number of tests leads to the increase of accuracy but, on the other hand, remarkably increases the time of the experiment and, consequently, its cost. In this context, it is the research worker who, taking into account a possible future use of the fatigue curve, should individually decide about its accuracy.

**Keywords:** high-cycle fatigue, S-N curve, fatigue tests

### INTRODUCTION

When designing new structural elements which are expected to be exposed to time-dependent loads, such as for instance load-bearing parts of ship hulls, their fatigue strength curve is calculated for the assumed lifetime [1]. These calculations require the information on the S-N fatigue characteristic.

For this purpose curves constructed according to normative guidelines [2] or those formulated by classification societies [3,4], for instance, are used. Specimen fatigue curves for different categories of welded joints, taken from [4], are shown in Fig. 1. The recommended characteristics refer to 75% confidence with 95% probability of reliability [2,5]. It is noteworthy that the dispersion of the results recorded in experimental examination is relatively large, as shown in Fig. 2 [7]. Since any change of material, geometry, heat treatment, etc., of the structural element results in its fatigue strength change, fatigue tests are to be repeated each time when new technology is introduced, or new material used.

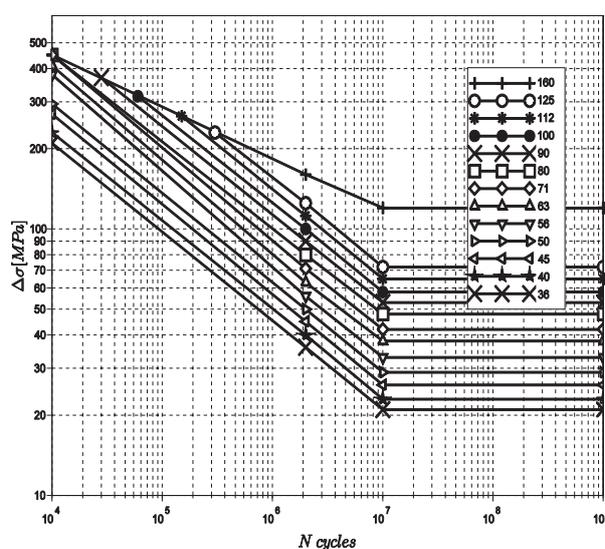


Fig. 1. Specimen fatigue curves for different categories of welded joints [4]

For instance, Ref. [8] presents the effect of the size of the tested object on its fatigue life.

Fatigue tests oriented on constructing the S-N curve within a limited fatigue life range can be performed based on guidelines given in [5,6,9,10,11], among other options. However, despite the fact that all these documents refer to the limited fatigue life range, each of them was prepared for an individually defined scope of applications and contains different guidelines.

The present analysis aims at comparing these guidelines and presenting experimental verification of the effects of their implementation when constructing the S-N characteristic.

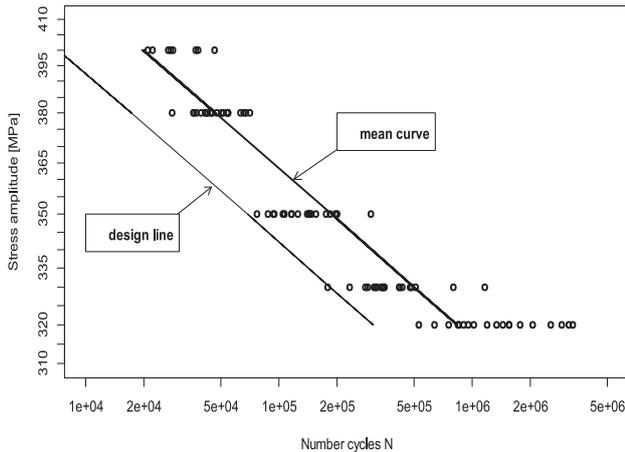


Fig. 2. Dispersion of results of fatigue tests for steel 0,72%C [7]

## NUMBER OF TESTED SPECIMENS

The fatigue tests are time-consuming and, consequently, expensive. For instance, obtaining 105 cycles at the loading frequency of 30 Hz takes 55 min, while 106 cycles require as much as 9 hours. This assessment does not take into account the time of specimen preparation. It is also noteworthy that fatigue test machines frequently enable to obtain much lower frequency, for instance 5 Hz or less, which even more extends the time of examination. That is why numerous research workers limit the number of the tested specimens to a minimum, in particular in case of multiaxial fatigue tests, such as those reported in Ref. [12]. Since different characteristics are frequently used for comparisons, for instance to determine the effect of different production technologies on the fatigue life, numerous normative guidelines were worked out which define the minimal acceptable number of specimens. Selected guidelines are discussed below.

According to the standard [11], tests oriented on determining the fatigue characteristic for a given reliability level, require 28 specimens as the minimum, while in preliminary tests 7 specimens are recommended as sufficient. The above guidelines do not define the number of the applied load levels, they only refer to the standard ASTM [6] which gives the relation for the replication percent:

$$PR = 100 \left[ 1 - \left( \frac{Sl}{n} \right) \right] [\%] \quad (1)$$

where:

$Sl$  – number of load levels,

$n$  – total number of specimens.

This standard includes a suggestion that the replication percent given in Tab. 1 is to be obtained.

It is noteworthy that the minimal number of specimens recommended in the standard [11] is larger than that given in the document [6]. Numbers of specimens suggested in the ASTM standard are given in Tab. 2.

Other guidelines can be found in the EN standard [9], where a proposal is formulated to start the test at the load level for which the fatigue life approximately equal to  $10^5$  cycles is expected. This guideline can be applied when the characteristic for a similar material is known. When the reliable information which would enable to assess this load level is missing, the following formula can be used instead:

$$S_5 = (UTS + 0.2\% PS) / 2, \quad (2)$$

where:

$S_5$  – stress amplitude for 105 cycles

UTS – ultimate tensile strength,

0.2% PS – proof stress.

The above notation is in accordance with the document [9]. It is noteworthy that the fatigue life based on tensile test results can be calculated using analytical methods, described in [13,14], among other references.

Tab. 1 Guidelines concerning the replication percent for various tests [6]

Type of Tests	Replication Percent
Preliminary and exploratory (research and development tests)	17 to 33
Research and development tests of components and specimens	33 to 50
Design allowable data	50 to 75
Reliability data	75 to 88

Tab. 2 Guidelines concerning the number of specimens [6]

Type of Test	Minimum number of specimens
Preliminary and exploratory (research and development tests)	6 to 12
Research and development tests of components and specimens	6 to 12
Design allowable data	12 to 24
Reliability data	12 to 24

The next level of load should be lower by 40 MPa than the initial level. If the specimen cracks before the number of  $10^7$  cycles is obtained, the next level of load should be decreased by 20 MPa. Then, the level of load should be increased by 40 MPa with respect to the initial level. If the specimen does not crack after about  $10^4$  cycles, the level of load should be increased by the next 20 MPa. As a final part, tests should be performed for the load levels situated between those previously applied. This testing procedure does not include the information on the recommended minimal number of specimens.

Other guidelines can be found in the PN standard [10]. This document recommends that the tests are to be performed for a minimum of 15 specimens and at least for 5 levels of load. The obtained S-N characteristic should cover the range from  $10^4$  to  $10^6$  cycles.

On the other hand, the guidelines of the International Institute of Welding [5] suggest that the tests are to be performed for a minimum of 10 specimens, and for at least 2 levels of load within the range from  $10^5$  to  $10^6$  cycles.

It is also noteworthy that the ISO standard [11] has the applicability range limited to determining fatigue properties of metal elements, while the ASTM document [6] refers generally to construction of the S-N curve when it can be estimated using linear regression for the probability not lower than 5%.

In turn, the EN standard [9] can be applied for metal materials for which high-cycle fatigue characteristics are to be constructed. It does not refer clearly to structural elements, but only suggests that this methodology can be applied for those elements.

The PN standard [10] refers to tests of metals at axial load, bending, and twisting. It can be used for both smooth and notched specimens. The above standard does not include tests of specimens with welded, or glued joints, etc.

The IIW document [5] is recommended in cases of tests of welded elements made of steel or aluminium.

### INTERVAL BETWEEN LOAD LEVELS

As shown in Ref. [15], the S-N characteristic can take different forms. Within the limited fatigue life range it can be most frequently estimated by a linear model, see Fig. 3. This model of S-N characteristic is applied in the majority of practically used methods to calculate limited fatigue life and strength within this cycle range. Fig. 3 shows a scheme of estimation of the fatigue characteristic based on experimental data. The relation between the fatigue life and the load is most frequently estimated using the linear regression method, which can be described as:

$$\log(N) = m \log(S) + b, \quad (3)$$

where:

$m$  – slope coefficient,

$b$  – intercept term.

The values of the slope coefficient  $m$  vary within a wide

range (Fig. 5, Fig. 6, Fig. 7, Fig. 8). When the value of  $m$  is close to the upper limit of its changeability, the characteristic becomes more horizontal and the intervals between the levels of load become smaller. On the other hand, when the slope coefficient  $m$  is small, the fatigue characteristic has a more vertical form and, consequently, the intervals between the levels of load are larger. As an example, Fig. 4 shows two fatigue characteristics for smooth specimens made of steels 4140 [16] and SUJ2 [17]. The S-N curves for these two materials have different slope coefficients, equal to 23.15 and 9.829 for the former and latter material, respectively. For the material 4140 the fatigue strength changes by about 30 MPa within the cycle number ranging from  $5 \cdot 10^4$  to  $2 \cdot 10^5$  cycles. For the steel SUJ 2 and the same cycle range, the stress amplitude is higher and approximately equal to 100 MPa.

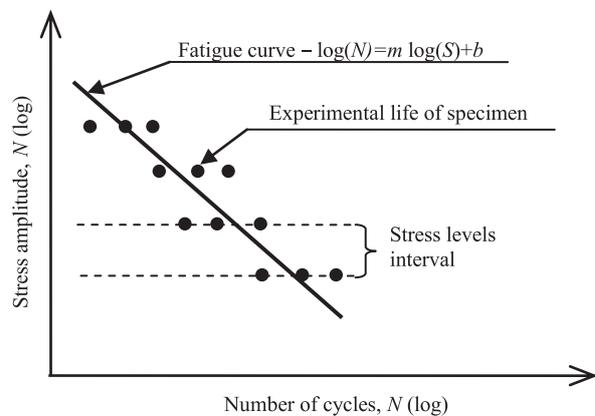


Fig. 3. Scheme of estimating the fatigue curve in high cycle range

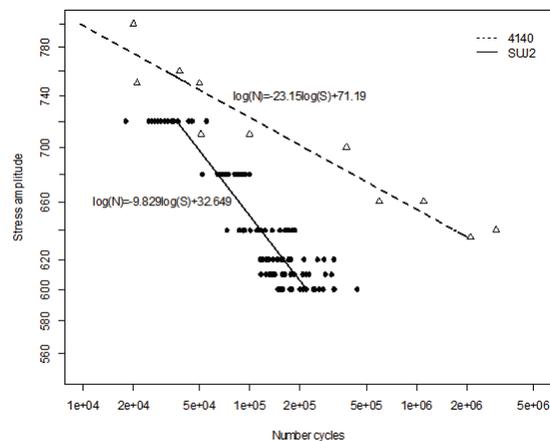


Fig. 4. Fatigue characteristic for steel 4140 [16] and steel SUJ2 [17]

This effect poses a problem when formulating the fatigue test programme, as precise assessment of the next level of load is not possible. To solve this problem, an attempt can be made to assess the most likely value of the coefficient  $m$  which will enable to determine approximately the changeability of the fatigue strength within the given range of fatigue life. Analyses oriented on determining slope coefficient changes are presented in [18,19]. Fig. 5, Fig. 6, Fig. 7 and Fig. 8 show

distributions of values of the slope coefficient  $m$  for structural steel and normal and shear stresses, and for even and notched specimens.

Fig. 5 reveals that the most frequent value of the slope coefficient  $m$  for smooth specimens and normal stresses is 10.9, i.e. the value which is higher than that obtained for the material SUJ2, but lower than that for 4140.

Moreover, it should be noted that in case of notched specimens the dispersion of values of the slope coefficient  $m$  is smaller, which makes assessing intervals between load levels during examination less problematic and more predictable.

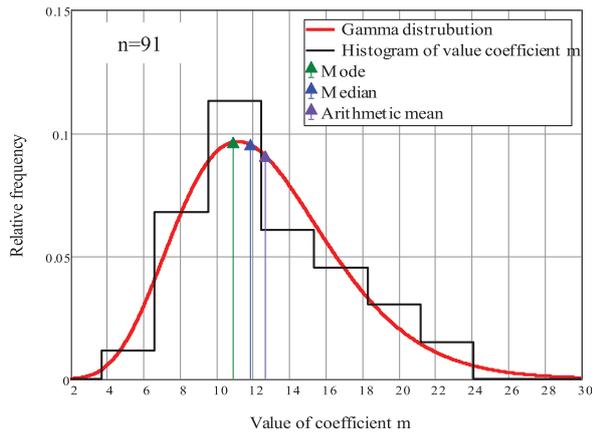


Fig. 5. Distribution of coefficient  $m$  for structural steel, smooth specimen, and normal stress

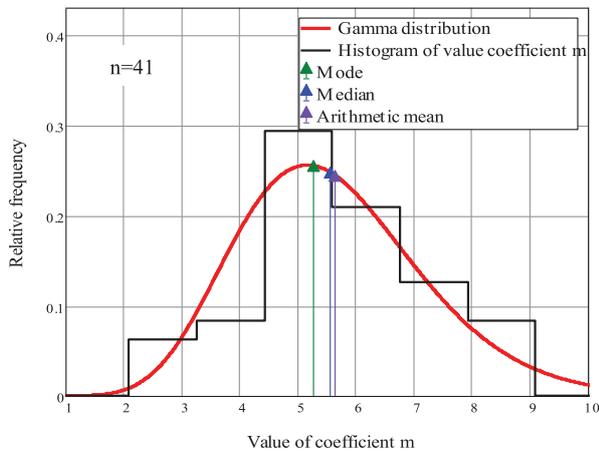


Fig. 6. Distribution of coefficient  $m$  for structural steel, notched specimen, and normal stress

The analysis of values of the slope coefficients  $m$  has made the basis for calculating modal values of this quantity for selected types of loads and specimens. The results of these calculations are given in Tab. 3, which also includes the calculated strength changes for the fatigue life ranging from  $5 \cdot 10^4$  to  $2 \cdot 10^5$  cycles, assuming that the value of the slope coefficient  $m$  is equal to the modal value of the relevant distribution. It is noteworthy that the values obtained for shear stresses and notched specimens have poorer statistical background due to a relatively small amount of data used to construct a histogram (reports on this type of research are under-represented in the literature).

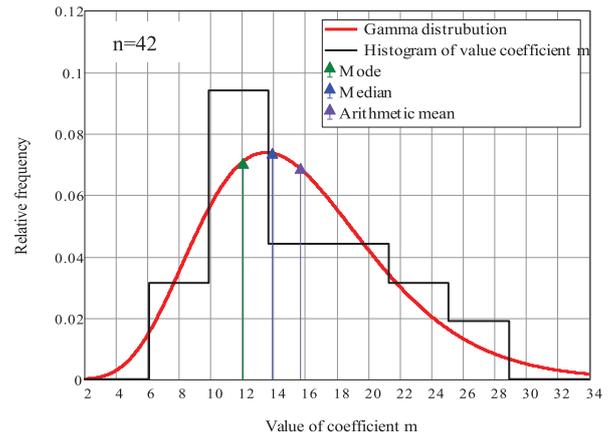


Fig. 7. Distribution of coefficient  $m$  for structural steel, smooth specimen, and shear stress

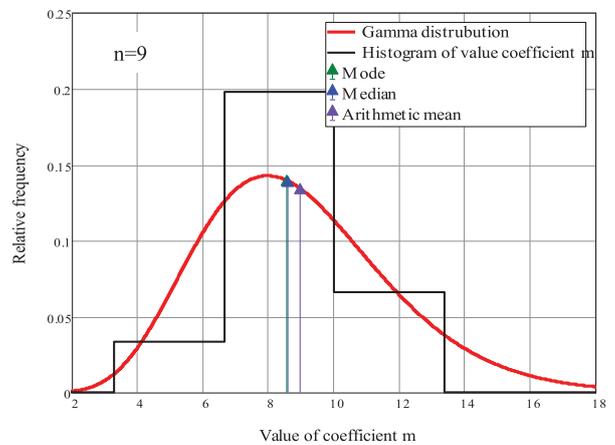


Fig. 8. Distribution of coefficient  $m$  for structural steel, notched specimen, and shear stress

Tab. 3 Modal values of slope coefficient  $m$  for structural steel

Type of load and specimen geometry	Modal value of slope coefficient $m$	Strength change within fatigue life ranging from $5 \cdot 10^4$ to $2 \cdot 10^5$ cycles [MPa]
Normal stress smooth specimens	10.9	38
Normal stress notched specimens	5.3	92
Shear stress smooth specimens	12	25
Shear stress notched specimens	8.5	45

## EXPERIMENTAL VERIFICATION OF THE PRESENTED GUIDELINES

The reported tests aimed at, among other goals, obtaining experimentally verified data constructed in accordance with particular guidelines [5,6,9,10,11] and in comparable test conditions, i.e. for the same specimens (Fig. 9), the same material (Tab. 4), and the same type of load (rotating bending, the same strength testing machine).

The tests were performed for the material 42CrMo4, tempered, and for cylindrical notched specimens. Static properties of the material selected for examination are given in Tab. 4, while the geometry of the specimens is shown in Fig. 9.

Tab. 4 Properties of material 42CrMo4

Ultimate strength $S_u$ [MPa]	1172
Yield stress $S_y$ [MPa]	1095
Hardness HB	308

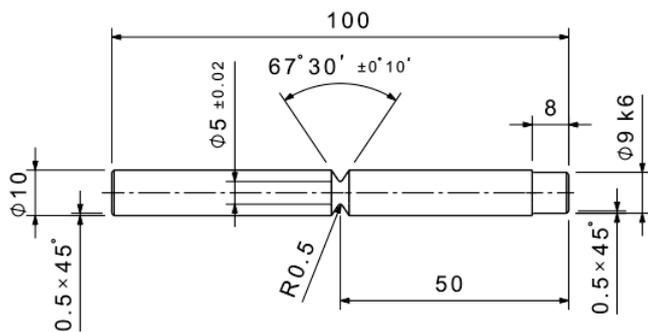


Fig. 9. Notched specimen used for tests

The tests were performed on the research rig used for rotating bending. The strength testing machine and its verification are described in Ref. [20]. The tests were performed at the frequency equal to 28.5 Hz.

The obtained test results are shown in Fig. 10, Fig. 11, Fig. 12 and Fig. 13, along with the estimated S-N curves. The curves were constructed in accordance with the guidelines formulated in the above quoted documents.

In cases when the linear regression was applied to the results shown in Fig. 11 (IIW guidelines [5]), low coefficient of determination and small intervals between load levels were obtained. Moreover, the resultant slope coefficient  $m$  differs much from those obtained for the remaining characteristics. It was equal to 6.197 in tests based on IIW guidelines [5], while its values for the remaining characteristics were, respectively, equal to: ISO [11] – 5.809, EN [9] – 5.563, and PN [10] – 5.768. The value of the intercept term calculated for the characteristic constructed based on the IIW guidelines [5] is also remarkably higher than the remaining estimated values, which most likely results from the fact that the inflection

point of the fatigue characteristic curve is situated close to  $4 \cdot 10^5$  cycles, which made fatigue life assessment in the vicinity of  $10^6$  cycles impossible. It is worth mentioning here that Sonsino [21] suggests that the inflection point cycle number of the S-N curve for high-strength steels is equal to  $5 \cdot 10^5$ . The above facts were the reason why the characteristic obtained in accordance with the IIW guidelines [5] was not taken into account in further analysis.

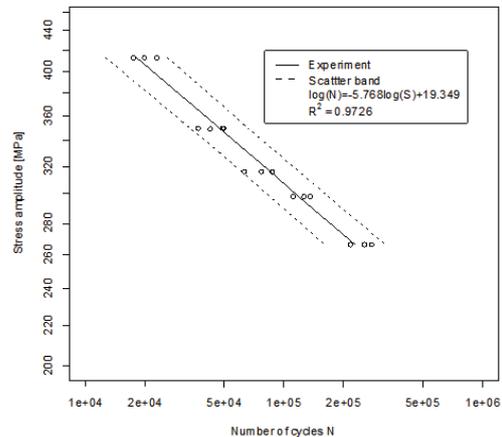


Fig. 10. Fatigue curve of steel 42CrMo4 for notched specimens,  $K_t = 1.99$ , [10]

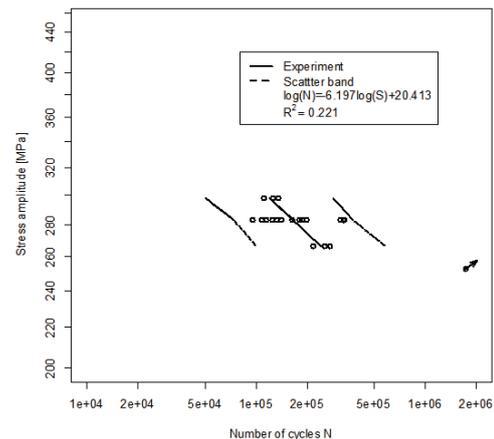


Fig. 11. Fatigue curve of steel 42CrMo4 for notched specimens,  $K_t = 1.99$ , [5]

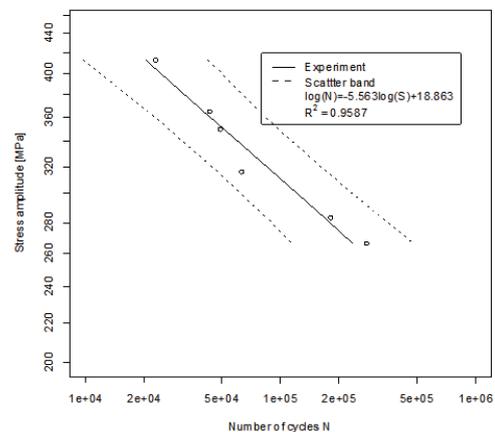


Fig. 12. Fatigue curve of steel 42CrMo4 for notched specimens,  $K_t = 1.99$ , [9]

Fig. 14 compares all constructed characteristics (except that making use of the IIW methodology [5]). The qualitative analysis reveals that the obtained differences between positions of particular lines are small.

The quantitative comparison was done using the following equation:

$$B = \frac{N_1 - N_2}{N_1} \cdot 100\% \quad (4)$$

where:

$B$  – relative difference between the analysed curves,

$N_1, N_2$  – fatigue life values calculated using the analysed curve.

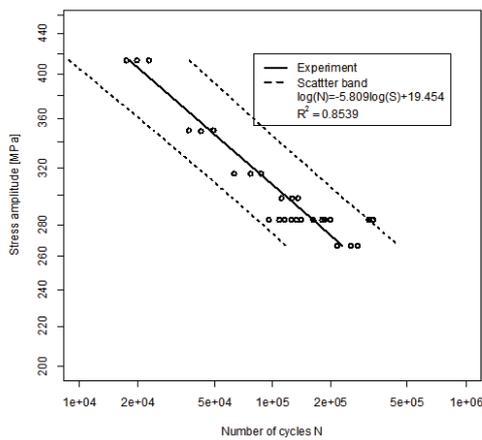


Fig. 13. Fatigue curve of steel 42CrMo4 for notched specimens,  $K_t = 1.99$ , [11]

The results calculated using the equation (4) are shown in Fig. 15. The largest difference amounts to 12.2% and refers to the results obtained based on PN and EN guidelines. The smallest difference is -2.6% and was recorded between the S-N curves constructed based on PN and ISO guidelines. The differences between the linear regression coefficients estimated based on the PN and ISO standards were smaller than those based on the PN and EN standards, and the ISO and EN standards.

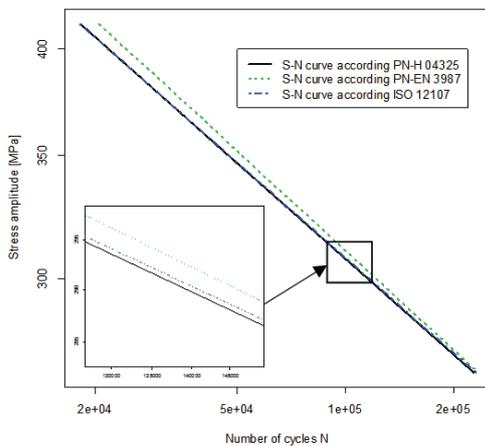


Fig. 14. Comparing the estimated curves

The slope coefficients  $m$  calculated in accordance with the guidelines of ISO [11], EN [9], PN [10] are close to the modal

value of 5.27 obtained from the probability function shown in Fig. 6. At the same time, the stress change amplitude within the fatigue life range from  $5 \cdot 10^4$  to  $2 \cdot 10^5$  cycles was smaller and equal to 74 MPa.

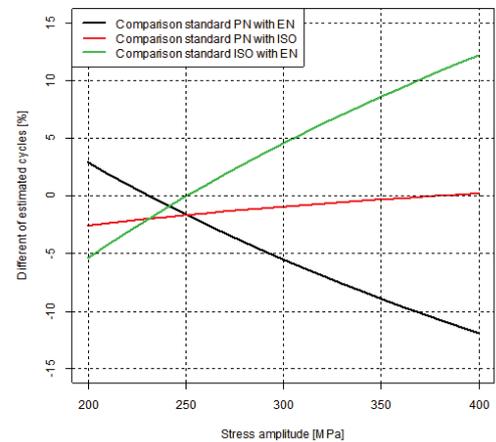


Fig. 15. Differences in cycle estimation based on the obtained fatigue curves for steel 42CrMo4 and notched specimens

## SUMMARY AND CONCLUSIONS

Based on the above analysis of methods to construct the S-N fatigue characteristic within the high-cycle range the following conclusions can be formulated:

- the number of specimens used for tests affects the parameters of the obtained linear regression, which in turn affect the value of the fatigue life calculated using the estimated characteristic,
- small differences between the characteristics obtained with the aid of the methodology recommended in the PN standard (15 tests) and the ISO standard (30 tests) suggest that these curves are more accurate than the characteristic obtained using the EN guidelines (6 tests),
- performing more than 3 measurements at one load level leads to higher dispersion of results (lower value of the coefficient of determination was obtained),
- the largest differences between the obtained characteristics are observed at the ends of the examination ranges,
- the characteristic obtained using the IIW methodology [5] reveals that this methodology was inappropriate for the examined material and specimen geometry,
- guidelines of the EN standard [9] seem to be appropriate for preliminary examination, but they should be verified for other materials and specimen geometries, as well, as for other types of load, twisting for instance,
- although the EN standard [9] is not applicable to notched specimens, a suggestion that it can be used to construct the fatigue curve for such objects seems reasonable, which was confirmed by the experimental results,
- arguments in favour of the methodology contained in the EN document [9] include an unambiguous

procedure of selecting intervals between load levels; this approach is also supported by the distributions of changes of the slope coefficient  $m$  presented in Fig. 5, Fig. 6, Fig. 7 and Fig. 8, as well as the analysis collated in Tab. 3,

- the use of load level intervals given in Tab. 3 when constructing fatigue characteristics for structural steels seems justified.

## BIBLIOGRAPHY

1. Kozak J., Górski Z., *Fatigue strength determination of ship structural joints*, Polish Maritime Research 2(69) Vol 18, 2011.
2. PN-EN 1993-1-9, Eurokod 3: *Designing steel structures. Part 1-9: Fatigue*, 2007 (in Polish).
3. Kocak M., et al., *FITNET Fitness-for-Service PROCEDURE – FINAL DRAFT MK7. tom I: FITNET FFS Procedure. European Fitness-for-Service Thematic Network – FITNET*, 2006.
4. Hobbacher A.F., *The new IIW guidelines for fatigue assessment of welded joints and components- A comprehensive code recently updated*. International Journal of Fatigue, Vol. 31, Is. 1, January 2009.
5. Hobbacher A., et al. *Guidelines for fatigue design of welded joints and components*, IIW document IIW-1823-07ex XIII-2151r4-07/XV-1254r4-07;2008.
6. ASTM E 739-91, *Standard Practice for Statistical Analysis of Linear or Linearized Stress-Life (S-N) and Strain-Life (S-N) Fatigue Data*, 2004.
7. Data Sheet on Fatigue Properties of Rail Steel by Alternating Axial Loading Life Test. [online]. Meiji University Academic Frontier. [access: 07 August 2012]. World Wide Web: <[http://www.isc.meiji.ac.jp/~shinrai/H12/H12%20Alternating%20Axial%20\(Rail%20Steel2\).pdf](http://www.isc.meiji.ac.jp/~shinrai/H12/H12%20Alternating%20Axial%20(Rail%20Steel2).pdf)>.
8. Tomaszewski T., Sempruch J., *Analysis of size effect in high-cycle fatigue for EN AW-6063*, Solid State Phenomena, Vol. 224, 2015.
9. PN-EN 3987, *Aviation and astronautics – Methods to test metal materials – High-cycle fatigue tests with constant amplitude of controlled load*, 2010 (in Polish).
10. PN-H 04325, *Fatigue tests of metals – Basic terms and general guidelines on preparation of specimens and testing procedure*, 1976 (in Polish).
11. ISO 12107, *Metallic materials - fatigue testing - statistical planning and analysis of data*, 2003.
12. Skibicki D., Sempruch J., Pejkowski L., *Steel X2CrNiMo17-12-2 testing for uniaxial, proportional and non-proportional loads as delivered and in the annealed condition*, Material Science Forum, Vol. 726, 2012.
13. Sempruch J., Strzelecki P., *Error of fatigue life determined according to the Fitnet method*, 17th International Conference on Engineering Mechanics, Svatka, Czech Republic, May 2011.
14. Strzelecki P., Sempruch J., *Experimental verification of analytical method for determining the S-N curve for alloy steel*, Key Engineering Materials, Vol. 598, 2014.
15. Strzelecki P., *Analytical method to determine high-cycle fatigue properties of materials and construction elements* (in Polish), Ph.D. thesis, 2014.
16. Kenan G., Mehmet D., Mustafa U., *Effect of cathodic polarisation on corrosion fatigue behaviour of ion nitrided AISI 4140 steel*. International Journal of Fatigue, Vol. 24 Is. 5, May 2002.
17. Data Sheet on Fatigue Properties of SUJ2 Steel by Rotating Bending Test (HRC30), [online], Meiji University Academic Frontier, [access: 07 August 2012], World Wide Web: <http://www.isc.meiji.ac.jp/~shinrai/H01/H01%20Rotating%20Bending%20%28SUJ-HRC30%29.pdf>.
18. Strzelecki P., Sempruch J., *Analysing slope coefficient of the Wöhler curve within finite life range* (in Polish). XXV Conference on Fatigue and Fracture Mechanics, Bydgoszcz-Fojutowo 2012.
19. Strzelecki P., Sempruch J., *Hybrid method for determining fatigue characteristic in high cycle life*. 20th International Conference Engineering Mechanics 2014, Svatka, Czech Republic, May 2014.
20. Strzelecki P., Sempruch J., *Experimental verification of the analytical method for the estimated S-N curve for limited fatigue life*. Materials Science Forum, Fatigue Failure and Fracture Mechanics, Vol. 726, August 2012.
21. Sonsino S.M., 2007. *Course of SN-curves especially in the high-cycle fatigue regime with regard to component design and safety*. International Journal of Fatigue. Vol. 29, Is.12, December 2007.

## CONTACT WITH THE AUTHORS

Przemysław Strzelecki  
Janusz Sempruch  
Krzysztof Nowicki

UTP University of Science and Technology  
ul. Ks. Kordeckiego 20  
85-225 Bydgoszcz  
**POLAND**

# LASER DOPPLER VIBROMETER BASED EXAMINATION OF THE EFFICIENCY OF INTRODUCING ARTIFICIAL DELAMINATIONS INTO COMPOSITE SHELLS

Kamila Kustron, Ph. D.

Institute of Aeronautics and Applied Mechanics, Poland

## ABSTRACT

*During its operation, the laminate shell of the watercraft hull can be exposed to local stability losses caused by the appearance and development of delaminations. The sources of these delaminations are discontinuities, created both in the production process and as a result of bumps of foreign bodies into the hull in operation. In the environment of fatigue loads acting on the hull, the delaminations propagate and lead to the loss of load capacity of the hull structure. There is a need to improve diagnostic systems used in Structural Health Monitoring (SHM) of laminate hull elements to detect and monitor the development of the delaminations. Effective diagnostic systems used for delamination assessment base on expert systems. Along with other tools, the expert diagnostic advisory systems make use of the non-destructive examination method which consists in generating elastic waves in the hull shell structure and observing their changes by comparing the recorded signal with damage patterns collected in the expert system database. This system requires introducing certain patterns to its knowledge base, based on the results of experimental examinations performed on specimens with implemented artificial delaminations. The article presents the results of the examination oriented on assessing the delaminations artificially generated in the structure of glass- and carbon-epoxy laminates by introducing local non-adhesive layers with the aid of thin polyethylene film, teflon insert, or thin layer of polyvinyl alcohol. The efficiency of each method was assessed using laser vibrometry. The effect of the depth of delamination position in the laminate on the efficiency of the applied method is documented as well.*

**Keywords:** watercraft hull, laminate, artificial delamination, SHM diagnostics, advisory expert system knowledge base

## INTRODUCTION

The durability of the watercraft hull structure affects considerably the voyage safety and is subject to demanding requirements from maritime surveillance institutions.

The developed hull structures are oriented on the use of intelligent materials and structures. Research activities in this area are directed towards gradient and self-repairing materials, new technologies of production of elements and combining them into modules, subassemblies and the final structure, new solutions and methods of system operation administration, new and/or modernised methods to analyse correlations of links/factors in the human being/technology/environment (C-T-O) macrosystem, new applications of mechanics methods and related techniques, modelling methods to simulate operating processes, and/or methods of structure diagnostics to assess its technical state.

The present approach to structural durability of watercraft hulls defines the durability as structure's robustness to the agents which lead to the degradation of material properties or structural integrity of the hull operating in the corrosion/fatigue environment. In normative terms, the need to preserve sufficient watercraft hull durability imposes certain

conditions to ensure material consistency and integrity of structural joints of the hull at the macroscopic level in real operating conditions. The loss of consistency or integrity is caused by damage development to the scale in which part of the construction material which is free from damages, or a weakened joint, is not able any longer to carry the operating load. The damages can be initiated at different levels of matter organization: as (1) the development of discontinuities formed during production processes of materials, elements, subsystems, and their assembly into a final product, or (2) the development of operating damages caused by local transgression of permissible loads due to bumps of foreign bodies. The further course of the discontinuity development depends on the type of construction material, the applied design solutions, and/or other agents acting on the hull structure in the C-T-O macrosystem during the hull operation.

The watercraft hull durability created at the preoperational stage should ensure macroscopic consistency of its structure by preserving a safe level of stresses and deformations generated by external and internal loads acting on the hull during its

operation. Moreover, possible damages should develop only locally within the range which is acceptable from the point of view of safety requirements, and should be able to be detected and monitored using non-destructive methods. The durability is to be ensured by an effective diagnostic system which will detect and assess the degradation or destruction processes.

Modern engineering structures are developed with orientation towards intelligent applications [4]. One of solutions here is intelligent diagnostics. Applying a system of integrated diagnostics based on intelligent solutions enables to obtain a more robust and trustworthy structure, as a result of real assessment of its technical state [5, 6].

Two philosophies are applied when designing ship hulls. One of them consists in designing the hull structure in accordance with the applicable regulations, with further verification, based on real and virtual experiments, that in the assumed lifetime of operation the integrity of joints and material consistency will be preserved at the assumed safety level. The other philosophy takes into account the fact that all materials contain internal discontinuities resulting from technological processes, and these discontinuities can develop during the hull operation only up to a level at which the fatigue crack type damage does not jeopardise the metal structure of the ship hull and therefore can be permitted. Dural structures of ship hulls should ensure their safe use by revealing sufficiently high resistance to fatigue, corrosion, and external damages. When well recognised in experimental examination, the fatigue processes of dural materials used in ship hull building make the basis for preparing structural inspection procedures. Dural materials are constantly improved with respect to their fatigue crack resistance. Designs of modern hulls make frequent use of laminates due to their high fatigue and corrosion resistance, and the progressive increase of their use is observed [10, 16].



Fig. 1. Laminate structures of contemporary watercraft units [15]

Laminates reveal high corrosion and fatigue resistance, but when they are exposed to bumps of foreign bodies they crack. Damages resulting from low-energy bumps, especially those which cannot be observed from outside, can lead to internal integrity loss between composite layers, a process which is recognised as delamination. This type of damage is dangerous for laminates and should be detected in the early phase of existence, as its propagation in the fatigue load environment poses a threat to the safety of the structure.

During its operation, the thin-walled structure of the composite watercraft hull shell is susceptible to local stability losses resulting from the appearance and development of delaminations. These delaminations appear as a consequence of the development of material discontinuities generated during the lamination process, when compression stresses

act on the laminate. They can also be initiated during the hull operation, when the laminate hull shell is exposed to bumps of foreign bodies. Especially dangerous are low-energy bumps, as they do not cause macroscopic damages which can be observed from the outside, and lead only to local loss of adhesive connection between the laminate layers. The appearance and development of these discontinuities can remain undetected. When propagating, the discontinuities initiate delaminations in the layered composite, which poses a threat of stability loss of the composite shell exposed to operating fatigue loads, especially compression loads. Therefore there is a need to improve tools used to diagnose the technical state of the ship hull for the appearance and development of delaminations [1-3, 9, 11-13]. Among other methods, elastic Lambda waves generated in thin-walled structures can be used for this purpose [8, 14].

When designing the ship hull, certain preventive measures should be named which will protect the hull against damages, including delaminations. The effect of damages on ship’s safety is assessed based on observations of other real structures and modelling the observed behaviours in experiments with the newly designed structure. The time of experiments performed during the design stage should be shortened by performing intensive tests. Their only goal is to shorten the time of experiment, and it needs to be examined whether the speeding-up procedures do not affect the structure damage criterion, whether the damage mode (type) is preserved, and whether other modes which would change the nature of damage propagation and structure destruction are absent. Theoretically, the accelerated tests, in which additional modes associated with acceleration are observed, can also be performed. However, for the accelerated tests, the behaviours which are characteristic for the observed mode and for the associated modes should be extracted, which is a difficult process.

Observing the behaviour of a structure exposed to simulated operating loads enables to assess the range of its safe operation. An aspect of highest importance here is to attribute certain symptoms of gradual loss of material consistency and/or integrity of structural joints to successive phases of the experiment. These symptoms are used in diagnostics. The recognised symptoms make a basis for assessing the remaining time of operation defined by the residual durability. Particular phases depend on the adopted time intervals, referred to as observation windows. The lower limit for the adopted time interval is the sampling interval, while the upper limit can be defined by the linearity, or quasi-linearity, of the behaviour of the symptom. The observation of the structure in operation should be oriented on detecting the abovementioned symptoms. This problem can be dealt with in two ways: (1) based on experimental tests the times of successive inspections are determined during which the technical state of the structure is examined using non-destructive tests – NDT, or (2) passive and/or active sensors, introduced to the structure, make it possible to detect the symptoms online during structure operation. The latter approach is more effective, both economically and in safety terms, as it eliminates the problem of access to diagnosing

points, and reduces, or totally eliminates, the effect of human interference. Moreover, it has a wider applicability scope and increases the technical readiness, thus taking care of the quality of service offered for the final user. It is noteworthy, however, that this higher efficiency bases on the following assumptions: (1) the system of sensors has been selected correctly and optimally for the given structure, and (2) its reliability is high and does not affect the reliability of the structure, in particular it does not generate false alarms.

The efficiency of the diagnostic systems used for delamination assessment is improved by implementation of expert systems. The diagnostic advisory expert systems make use of, among other tools, non-destructive test methods which consist in generating elastic waves in the examined structure and comparing the recorded signal with the damage reference pattern. This system requires introduction of reference patterns to the knowledge base of the expert system. The knowledge base is obtained from experimental examinations performed on specimens with implemented artificial delaminations.

The article presents the evaluation of efficiency of diagnostic methods which introduce artificial delaminations into composite shells made of glass- and carbon-epoxy laminates in order to define the reference delamination patterns for the expert system. The performed tests made use of three methods to introduce artificial delamination to the structure, and the modern diagnostic method based on the laser Doppler vibrometer to detect it.

## CHARACTERISTICS OF SPECIMENS USED IN TESTS

The article reports results of tests oriented on evaluating the delamination area and position depth in specimens with delaminations artificially generated between the composite layers. The specimens, prepared from glass- and carbon-epoxy laminates, had the A5 size. The generated artificial delaminations had the form of non-adhesive areas created between the laminate layers during the composite production process. These areas were created using three methods, by introducing: (1) a thin polyethylene film, (2) a teflon insert, and (3) a thin layer of polyvinyl alcohol. The delamination areas had the shapes of isosceles trapeziums with bases of 5 and 20 mm in length, arranged in such a way that the longer base is situated alternately at the left and right side of the specimen (Fig. 1a).

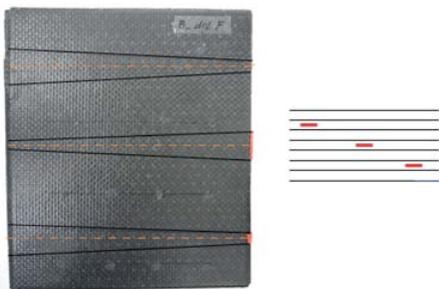


Fig. 1. A5 specimen made of carbon-epoxy laminate with local separation layers alternately introduced to obtain delaminations having the assumed areas and the shapes of isosceles trapeziums with bases of 5 and 20 mm

The above structure of the specimen was selected to obtain the information for the knowledge basis of the advisory expert system, which was the main goal of the research reported in the article. These specimens will also be used in further tests, and the adopted system should facilitate evaluation of the measurement uncertainty at the next stage of construction of the diagnostic system based on piezoelectric sensors permanently integrated with the ship hull structure.

## METHODOLOGY OF EXAMINATION

The tests were performed in the Engineering Institute, Los Alamos National Laboratory, USA, using the measuring set schematically shown in Fig. 2.

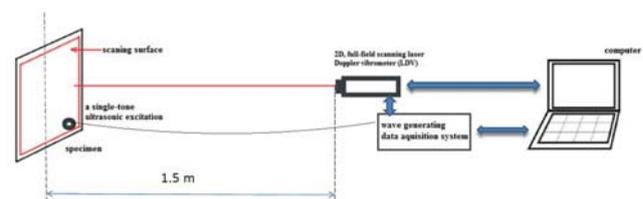


Fig. 2. Scheme of the test rig used for evaluation of the area of artificial delamination in the laminate structure

The examined specimen was mounted vertically at the distance of 1,5 m from the laser Doppler vibrometer. A high-frequency piezo activator which generates constant-frequency waves was wax-glued on the specimen. The assumed vibration frequency was equal to 80 kHz, following expert recommendations by Erick Flynn of Los Alamos National Laboratory. During the test, Lambda waves were generated in the specimen and, simultaneously, the surface of the specimen was scanned using the 2D Doppler laser. The visualisation was obtained via automatic estimation based on the assessment of the wave number, pixel-by-pixel, being the response to the excitation in the wave number domain [7]. The advantage of the applied method is that it does not need reference patterns. Anomalies in the image indicate the presence of deformations, which are larger in places with delaminations due to lower stiffness of these areas caused by the loss of interlayer integrity.

The obtained results are presented in 3D using an additional map of colours. These 3D images can be used to evaluate the depth of delamination, measured from the surface of scanning.

The specimens prepared for the tests are collated in Table 1, which in successive columns shows the number of the specimen, its label, and basic data concerning the reinforcement material (glass, carbon) and type of material used to generate local delamination (teflon, polyvinyl alcohol, polyethylene film).

Table 1. Specimens prepared for testing the efficiency of introduction of artificial

No of specimen	Symbol of specimen	Glass fabric reinforcement	Carbon fabric reinforcement	Teflon T	Polyvinyl alcohol A	Film F
1	A_delA	+	-	-	+	-
2	A_delF	+	-	-	-	+
3	A_delT	+	-	+	-	-
4	B	-	+	-	-	-
5	B_delA	-	+	-	+	-
6	B_delF	-	+	-	-	+
7	B_delT	-	+	+	-	-

Six specimens shown in Fig 3 were used in the tests.

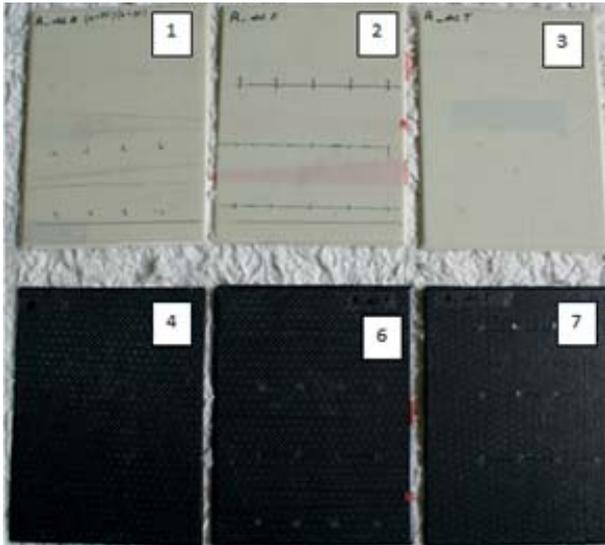


Fig. 3. Photo of the set of specimens used in the examination performed with the aid of the laser Doppler vibrometer to assess the areas of delamination

Specimens 1 to 3 were prepared from glass-epoxy laminate. During the delamination introduction process, all trapezoidal artificial delamination areas of the specimens were filled with: (1) polyvinyl alcohol with a thin copper wire, or (2) thin polyethylene film, or (3) rectangular teflon insert. Specimens 4 to 7 were prepared from carbon-epoxy laminate. Specimen 4 was pure, without artificial delamination, while specimens 5 to 7 had delamination areas filled with polyvinyl alcohol, thin polyethylene film, and rectangular teflon insert, respectively, like in case of the glass-epoxy specimens.

Specimen 5 made of carbon-epoxy laminate with local artificial delaminations introduced using polyvinyl alcohol was damaged during annealing in the production process.

## RESULTS OF TESTS

The analysis of the test results having the form of the recorded images clearly detects the delaminations. The method of their detection is very effective, fast, and efficient. The recorded images have made a basis for evaluating the efficiency of particular delamination introduction methods. Covering Specimen 1, made of glass-epoxy laminate, (Fig. 4), with polyvinyl alcohol did not result in local adhesion loss (Fig. 5), and the only visualised damage was caused by removing the specimen from the mould.

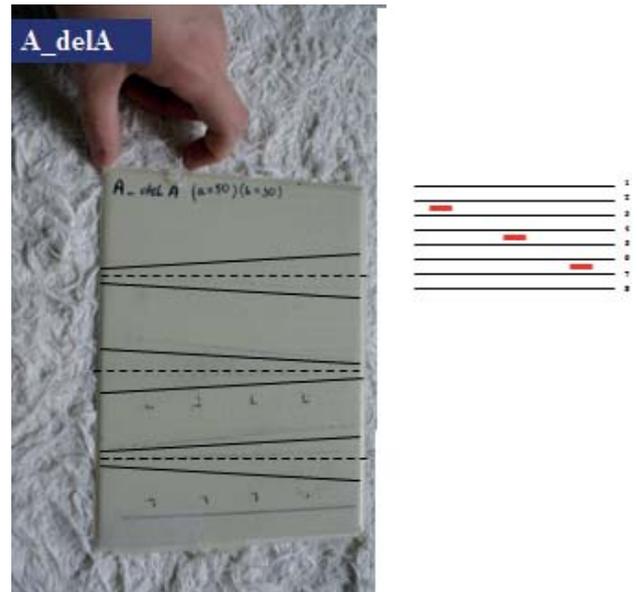


Fig. 4. View of Specimen 1, made of glass-epoxy laminate. To generate delamination, the places marked as trapeziums were covered with polyvinyl alcohol between layers 2 and 3, 4 and 5, and 6 and 7 in the production process

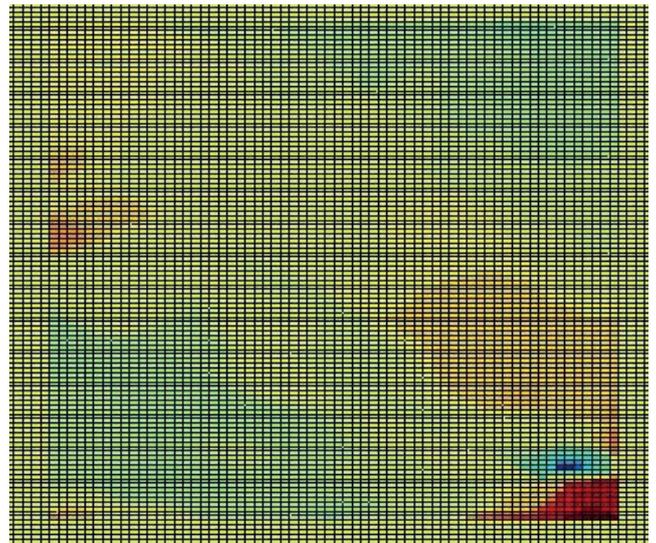


Fig. 5. Visualisation of Specimen 1, made of glass-epoxy laminate with trapeziums covered with polyvinyl alcohol to provoke local adhesion loss and further delamination between layers 2 and 3, 4 and 5, and 6 and 7

During the production process, a copper wire was introduced to the places covered with polyvinyl alcohol to enable the generation of additional high-voltage discharge, and thus provoke the appearance of delamination being the object of further examination.

Specimen 5 made of carbon-epoxy laminate with trapeziums covered with polyvinyl alcohol was damaged during annealing.

Three thin trapezium shaped films were introduced to Specimen 2, made of glass-epoxy laminate, between layers 2 and 3, 4 and 5, and 6 and 7 (Fig. 6).

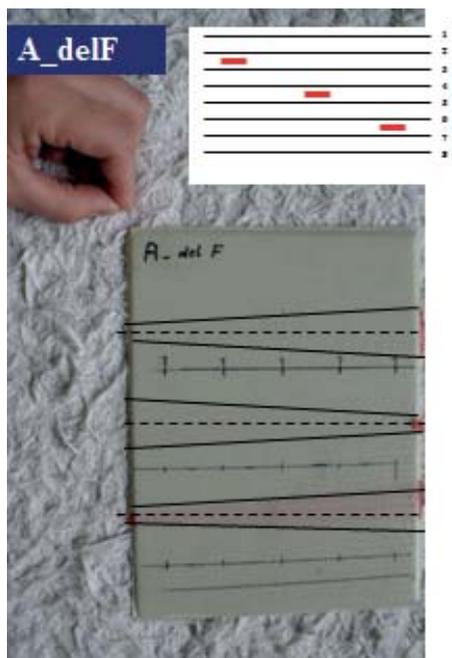


Fig. 6. View of Specimen 2, made of glass-epoxy laminate with places of introduction of trapezium shaped polyethylene films between layers 2 and 3, 4 and 5, and 6 and 7

Figure 7 shows visualisation of elamination of Specimen 2.

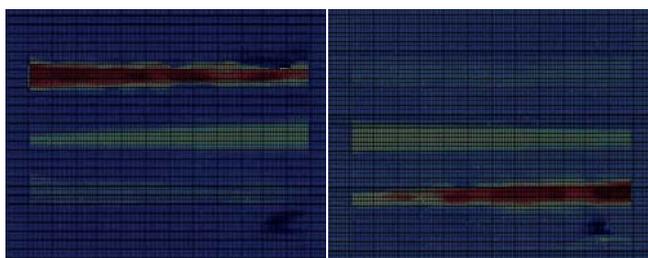


Fig. 7. Visualisation of delamination of Specimen 2. Internal delaminations obtained when the transducer was glued at the bottom left corner on the smooth side of the specimen and the rough surface was scanned (left), and when the transducer was glued at the bottom left corner on the rough side of the specimen and the smooth surface was scanned (right)

On the left-hand side, Fig. 7 shows the visualisation of internal delaminations obtained by exciting the Lamba waves with the aid of the trasducer glued at the bottom left corner on the smooth surface of the specimen and scanning the rough surface, while the visualisation shown on the right-hand side was obtained in the opposite situation, i.e. when the trasducer was moved to the bottom left corner of the rough surface and the smooth surface of the specimen was scanned.

Figure 7 shows clear patterns of internal delaminations, the used colour palette illustrates the depth of delamination position. The patterns on the left-hand and right-hand sides are similar to each other. The right-hand pattern can be obtained after transforming the left-hand pattern by rotation by 180° with respect to the bottom right corner, which reflects the arrangement of the delaminations purposely introduced to the structure of the laminate. The performed examination indicates that the efficiency of generating artificial internal

delaminations by introducing the films during the production process is high; the images of the introduced films are almost identical in both cases. We can also conclude that the side on which the trasducer is glued and the side along which the specimen is scanned do not affect the results of the examination. Further research is to be performed to separate the trasducer position/surface smoothness relation, from that between the surface smoothness and the scanned pattern.

The method to generate delamination by inserting a thin polyethylene film has also turned out effective for carbon-epoxy laminates.

However, unlike the test with the glass-epoxy laminate, here scanning Specimen 6 (Fig. 8), made of carbon-epoxy laminate with thin film introduced to generate delamination, has revealed some effect of trasducer position and scanned surface smoothness on the recorded results. More precise evaluation of this effect requires additional tests. The visualisation patterns on the left-hand and right-hand sides clearly differ between each other, by colours attributed to particular delaminations and depths of their positions in the laminate, rather than by delamination shapes themselves.



Fig. 8. Carbon-epoxy laminate with

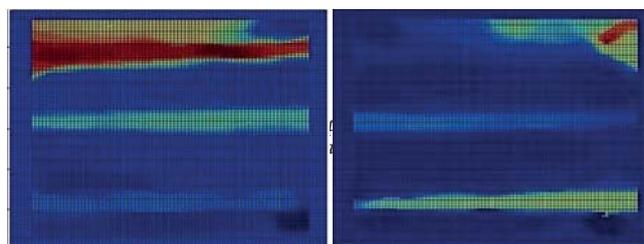


Fig. 9. Visualisation of delamination of Specimen 6 made of carbon-epoxy laminate with artificial delamination generated by introducing a thin film. The trasducer was glued on the smooth side of the specimen and the rough side was scanned (left), and trasducer was glued on the rough side of the specimen and the smooth side was scanned (right).

For Specimen 7 (Fig. 3 and Table 1), made of carbon-epoxy laminate, a rectangular teflon insert with a copper wire was introduced into it, and separately a cooper wire alone, see Fig. 10. If no delamination had been recorded, the specimen was planned to be exposed to high-voltage discharge.

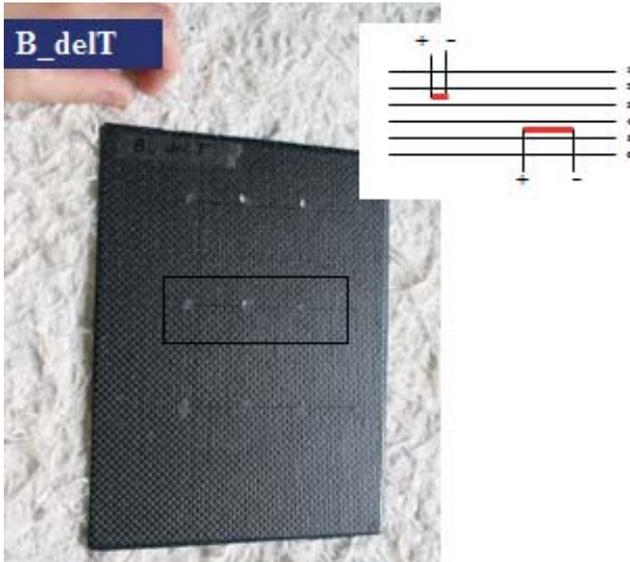


Fig. 10. Specimen 7, made of carbon-epoxy laminate, with (1) rectangular teflon insert and copper wire, and (2) a separate copper wire introduced during the production process

The visualisation did not detect the presence of delamination (Fig. 11). Here, the introduction of teflon turned out ineffective, as far as the goals of the reported research were concerned.

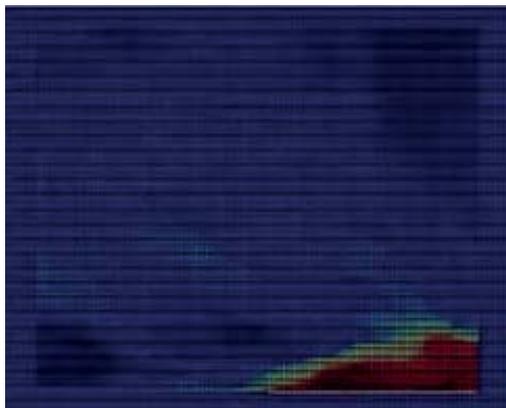


Fig. 11. Visualisation of Specimen 7 made of carbon-epoxy laminate, with locally introduced teflon insert. The transducer was glued on the smooth surface of the specimen and the rough surface was scanned

## CONCLUSIONS AND DIRECTIONS OF FURTHER RESEARCH

The article presents the results of the research oriented on evaluating the efficiency of different methods to introduce delamination into structures of glass- and carbon-epoxy laminates. The delamination was obtained by generating, in the production process, local non-adhesive areas between the laminate layers. A thin layer of polyvinyl alcohol, a thin polyvinyl film, and teflon were used for this purpose. The efficiency of particular methods was evaluated using the Laser Doppler Vibrometer. The highest efficiency was obtained using the thin polyethylene film, which was confirmed by the recorded visualisation images. Further research should

analyse in more detail the effect of surface smoothness on the introduction of Lamb waves and on the results of scanning.

The performed tests have also proved the efficiency of the applied method in assessing the depth in which the delamination takes place, i.e. between which layers. The efficiency of this method decreases with the increasing depth of delamination. The laminate thickness limits for which the method can be applied should be assessed taking into account the type of reinforcement fabric, its thickness, and volumetric percentage of the reinforcement and the matrix (binder) in the laminate.

The research was performed during author's stay in the Los Alamos National Laboratory, within the framework of the grant for study visit financed from the Development Programme of the Faculty of Power and Aeronautical Engineering, Warsaw University of Technology, for years 2011-2015, being part of the EU Operational Programme - Human Capital.

## BIBLIOGRAPHY

1. Bolotin V.V. Delaminations in Composites Structures: Its Origin, Buckling, Growth and Stability. *Composites: Part B Engineering*, 1996; 27(2): 129-145.
2. Bolotin V.V. Mechanics of Delaminations in Laminate Composite Structures. *Mechanics of Composite Materials*, 2001; 3(5/6): 367-380.
3. Clarke M.J., Pavier M.P. Experimental Techniques for the Investigation of the Effects of Impact Damage on Carbon Fiber Composites. *Composite Sciences and Technology*, 1995; 55: 157 - 169.
4. Dilhac J-M. Smarter. CHIST-ERA Project Seminar, 2013.
5. Farrar C. R., Worden K. An Introduction to Structural Health Monitoring. *Philosophical Transactions of the Royal Society A*, 365 (LA-UR-04-8385), 2007: 303-315.
6. Farrar C. R., Worden K. *Structural Health Monitoring: A Machine Learning Perspective*. John Wiley & Sons, 2012.
7. Flynn E.B., Jarmer G.S. High-Speed, Non-Contact, Baseline-Free Imaging of Hidden Defects Using Scanning Laser Measurements of Steady-State Ultrasonic Vibration. *International Workshop on Structural Health Monitoring*, Stanford, USA, DEStech Publications, Inc. 2013.
8. Giurgiutiu, V., (2005) Tuned Lamb Wave Excitation and Detection with Piezoelectric Wafer Active Sensor of Structural Health Monitoring. *Journal of Intelligent Material Systems and Structures*, 2005; 16(4): 291-305.

## CONTACT WITH AUTHOR

Kamila Kustróń

Institute of Aeronautics and Applied Mechanics  
Faculty of Power and Aeronautical Engineering  
Warsaw University of Technology  
Politechniki 1 Sq.  
00-661 Warszawa  
Poland

9. Liu S.F., Hwang G.H. Buckling Behaviour of Composite Laminates with Multiple Delaminations under Uniaxial Compression. *Composite Structures*, 2001; 53:
10. Mouritz A.P., Gellert E., Burchill P., Challis K., Review of advanced composite structures for naval ships and submarines. *Composite Structures*. Vol. 53(1),2001, pp. 21-42
11. Oswald A., Tafreshi T. Global Buckling Behaviour and Local Damage Propagation in Composite Plates with Embedded Delaminations. *International Journal of Pressure Vessels and Piping*, 2003, Vol. 80.
12. Schoeppner N.J., Pagano G.A. Delamination of Polymer Matrix Composites: Problems and Assessment. *Comprehensive Composite Materials*, Elsevier, 2000.
13. Soni C. D., Sahoo P. K., Srinivasan S., Santosh K. Fracture Mechanics Analysis and Strength Prediction of Carbon Fiber Composite Laminate with a Delamination. *JEST-M*, 2012;1(1): 8-11.
14. Wandowski T., Malinowski P., Kudela P., Ostachowicz W. Guided wave-based detection of delamination and matrix cracking in composite laminates. *Proceedings of the Institution of Mechanical Engineers Part C-Journal of Mechanical Engineering Science*, 2011; 225: 123-131.
15. [www.3mb.asia/the-use-of-composite-materials-in-nautical-applications/](http://www.3mb.asia/the-use-of-composite-materials-in-nautical-applications/)(on-line 30/03/2015)
16. [www.compositesworld.com/articles/the-markets-boatbuilding-and-marine-2015](http://www.compositesworld.com/articles/the-markets-boatbuilding-and-marine-2015)(on-line 30/03/2015)

# STUDY OF HARMONIC DETECTION METHODS UNDER NON-IDEAL CONDITIONS IN SHIP POWER NETWORK

Xiao-Yan Xu, Prof. <sup>1</sup>

Janusz Mindykowski, Prof. <sup>2</sup>

Tomasz Tarasiuk, Prof. <sup>2</sup>

Chen Cheng, M.Sc. <sup>1</sup>

1) Shanghai Maritime University, Shanghai, China

2) Gdynia Maritime University, Gdynia, Poland

## ABSTRACT

*An improved harmonic detection method based on average arithmetic is proposed. According to the research results, the designed solution uses an LPF (low-pass-filter) and a mean value module connected in series instead of the conventional mean value module, and simultaneously, a three-phase voltage phase-locked module instead of commonly used PLL (phase lock loop) module is applied in order to reduce the influence caused by three-phase distorted voltage and rapid variation of load. The experimental results show that the application of this solution leads to increase in the accuracy of harmonics detection for distorted three-phase voltage and rapid variation of load.*

**Keywords:** distorted three-phase voltage; rapid-variation load; harmonic detection; average arithmetic

## INTRODUCTION

With the development of power electronic devices in recent decades, more and more power converters are applied onboard, and the ship electric propulsion system is the representative example. But the fact is, power quality in marine electric power networks is becoming worse and worse, especially the level of harmonics increases. For example, in the case of Queen Mary 2 [10], the bulky ship has suffered a failure of harmonic passive filter, which resulted in explosion and damage of the surrounding electrical panels, and finally a total black-out of the vessel. In this case, the only protection against catastrophic failure of the capacitors being a part of harmonic passive filter, was a current imbalance detection system, which after the accident was found to be inoperable [10]. Anyway, a problem how to detect and compensate harmonic components with appropriate accuracy in a real time is one of the most important questions concerning a problem of power quality in ship networks. This question is fundamental from the point of view active or hybrid power filters control [9], [15]. Conventional harmonic detecting method is based on the theories such as instantaneous reactive power theory, Fourier transform or adaptive algorithms. The most commonly used methods are  $i_p-i_q$  algorithm [1] and its advanced method [8]. A method using a mean value module to replace LPF is described in [6] and [9] to improve the accuracy and reduce the response time of harmonic detection. The methods described in [1] and [9] using  $i_p-i_q$  algorithm without PLL are proposed to solve the problem caused by asymmetrical load. In [16] and [17], symmetrical component

method is proposed for positive sequence voltage and current measurement. But in the case of distorted three-phase voltage, these methods sometimes cannot detect positive sequence component accurately. And sometimes their accuracy may be influenced by the variation of the fundamental current. To reduce the influence of distorted three-phase voltage, a new three-phase PLL circuit is proposed in [7]. But using this circuit, the response time is longer. In [14], a PI controller is added to the detection system. In [12], a new PLL circuit is proposed for compensation of error which is caused by system delay. In this paper, a improved harmonic current detection method is discussed under the circumstances of distorted three-phase voltage - with unbalanced initial phase angles and amplitudes in the presence of rapid variation of load. In shorthand, an advanced PLL circuit is proposed for detection of the positive sequence voltage completed by a module consisting of a LPF and a mean module that are connected in series. It is to supersede the conventional LPF.

## HARMONIC DETECTION METHOD WITH MEAN VALUE MODULES

### PRINCIPLE OF HARMONIC DETECTION METHOD WITH MEAN VALUE MODULES

The harmonic detection method with mean value modules is graphically presented in Fig. 1. It applies  $i_p-i_q$  algorithm but LPF module is replaced by mean value modules. The

mean value obtained by simple integration is the DC value corresponding to the fundamental current. In an ideal power network, its detection time is only 1/6 cycle [6]. Its principle analysis is as follows. Assuming that the three-phase load current is,

$$\begin{cases} i_a = \sum_{n=1}^{\infty} I_n \sin(n\omega t + \varphi_n) \\ i_b = \sum_{n=1}^{\infty} I_n \sin[n(\omega t - 2\pi/3) + \varphi_n] \\ i_c = \sum_{n=1}^{\infty} I_n \sin[n(\omega t + 2\pi/3) + \varphi_n] \end{cases} \quad (1)$$

where,  $I_n$  and  $\varphi_n$  are the amplitude and initial phase angle of the  $n$ th harmonic current. The sine and cosine signals corresponding to the voltage of phase A,  $\sin\omega t$  and  $\cos\omega t$  can be detected by the PLL. They form matrix C which is showed in Fig.1,

$$C = \begin{bmatrix} \sin \omega t & -\cos \omega t \\ -\cos \omega t & -\sin \omega t \end{bmatrix} \quad (2)$$

According to the scheme, the formulas of  $i_p$  and  $i_q$  are,

$$\begin{cases} i_p = \sqrt{\frac{2}{3}} \sum_{n=1}^{\infty} I_n \left[ \sqrt{3} \sin \frac{2}{3} n\pi \cos \omega t \cos(n\omega t + \varphi_n) + \left(1 - \cos \frac{2}{3} n\pi\right) \sin \omega t \sin(n\omega t + \varphi_n) \right] \\ i_q = \sqrt{\frac{2}{3}} \sum_{n=1}^{\infty} I_n \left[ \sqrt{3} \sin \frac{2}{3} n\pi \sin \omega t \cos(n\omega t + \varphi_n) + \left(1 - \cos \frac{2}{3} n\pi\right) \cos \omega t \sin(n\omega t + \varphi_n) \right] \end{cases} \quad (3)$$

Here,  $i_p$  is taken for example; when  $n=1$ ,

$$i_p = \sqrt{\frac{3}{2}} I_1 \cos \varphi_1 = \bar{i}_p \quad (4)$$

when  $n = 3k$  ( $k = 1, 2, 3, \dots$ ),

$$i_p = 0 \quad (5)$$

when  $n = 3k + 1$  ( $k = 1, 2, 3, \dots$ ),

$$i_p = \sqrt{\frac{3}{2}} I_n \cos(3k\omega t + \varphi_n) \quad (6)$$

when  $n = 3k - 1$  ( $k = 1, 2, 3, \dots$ ),

$$i_p = -\sqrt{\frac{3}{2}} I_n \cos(3k\omega t + \varphi_n) \quad (7)$$

And  $i_q$  can be obtained in the same way. In three-phase symmetry circuit, even order harmonics do not exist. The system only contains harmonic currents of 5th, 7th, 11th and so on. So it can be stated that only components  $n = 6k \pm 1$  ( $k=1, 2, 3, \dots$ ) would exist. Then  $i_p$  and  $i_q$  become,

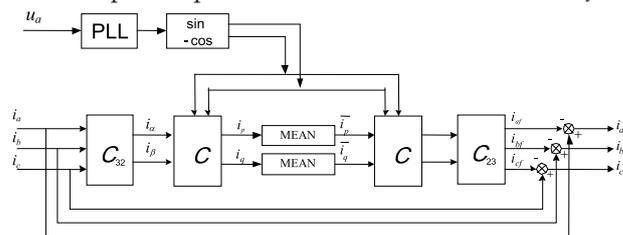


Fig.1 Block diagram of harmonic detection method based on  $i_p$ - $i_q$  concept with mean value modules

$$\begin{cases} i_p = \bar{i}_p \pm \sqrt{\frac{3}{2}} \sum_{n=2}^{\infty} I_n \cos(6k\omega t + \varphi_n) \\ i_q = \bar{i}_q - \sqrt{\frac{3}{2}} \sum_{n=2}^{\infty} I_n \sin(6k\omega t + \varphi_n) \end{cases} \quad (8)$$

After integrating the above formulas for 1/6 cycle, the fundamental component of  $i_p$  and  $i_q$  are obtained.

$$\begin{cases} \bar{i}_p = \frac{6}{T} \int_{t-\frac{T}{6}}^t i_p dt \\ \bar{i}_q = \frac{6}{T} \int_{t-\frac{T}{6}}^t i_q dt \end{cases} \quad (9)$$

After transforming them into three phase, fundamental currents  $i_{af}$ ,  $i_{bf}$  and  $i_{cf}$  are detected. And the harmonic component can be obtained.

## ERROR ANALYSIS

As shown in Fig.1, the phase of sine & cosine signal detected by method with mean value modules is determined by  $u_a$ . When three-phase voltage is asymmetrical and contains harmonic components, it has positive sequence component and negative sequence component. Detected by the PLL circuit, as shown in Fig.1, the phase of sine signal is the same as that of  $u_a$ , but differs by  $\theta$  from its positive sequence component. The exact signal are  $\sin(\omega t + \theta)$  and  $\cos(\omega t + \theta)$ . Calculating with the matrix C formed by them, the detection of fundamental active and reactive current will be influenced. Moreover, in the operation process of the APF (active power filter) from sampling of filtering signal to inverter output (Fig. 2), there exists sampling delay, transformer phase delay, signal processing delay and inverter delay of main circuit [3]. If the time delay is not compensated, the accuracy of detection will be very poor, and even the harmonic problem becomes more serious. That is, phase preprocessing during detection is necessary. But when load varies, the conventional detection method with mean modules may be readily disturbed by changing currents.

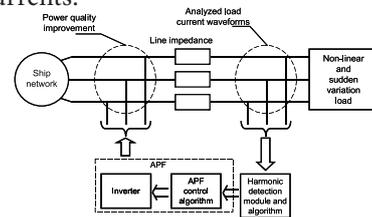


Fig.2 Principle block diagram of the considered ship network, active power filter and its control

## PRINCIPLE OF PROPOSED HARMONIC DETECTION METHOD

According to the above error analysis, an improved harmonic detection method is proposed. Its scheme (block diagram) is shown in Fig. 3.

The proposed solution is able to detect positive sequence component of three-phase voltage when it is distorted and asymmetrical. Its principle analysis is as follows. Assume that the asymmetrical three-phase voltage is,

$$\begin{cases} u_a = \sum_{n=1}^{\infty} U_n^+ \sin(n\omega t + \varphi_{u_n^+}) + \sum_{n=1}^{\infty} U_n^- \sin(n\omega t + \varphi_{u_n^-}) \\ u_b = \sum_{n=1}^{\infty} U_n^+ \sin[n(\omega t - 2\pi/3) + \varphi_{u_n^+}] + \sum_{n=1}^{\infty} U_n^- \sin[n(\omega t + 2\pi/3) + \varphi_{u_n^-}] \\ u_c = \sum_{n=1}^{\infty} U_n^+ \sin[n(\omega t + 2\pi/3) + \varphi_{u_n^+}] + \sum_{n=1}^{\infty} U_n^- \sin[n(\omega t - 2\pi/3) + \varphi_{u_n^-}] \end{cases} \quad (10)$$

where, U represents voltage amplitude, n represents harmonic order,  $\varphi$  represents initial angle, and (+,-) represent positive and negative sequence components.

When voltage signal passes BPF, the output signals are  $u_{af}$ ,  $u_{bf}$  and  $u_{cf}$ . Then the output signals are calculated by symmetrical component theory, all asymmetrical three-phase vectors are divided into three parts - positive sequence component, negative sequence component and zero-phase sequence component. The formulas are,

$$\begin{bmatrix} f_a \\ f_b \\ f_c \end{bmatrix} = \begin{bmatrix} f_a^+ + f_a^- + f_a^0 \\ f_b^+ + f_b^- + f_b^0 \\ f_c^+ + f_c^- + f_c^0 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 \\ a^2 & a & 1 \\ a & a^2 & 1 \end{bmatrix} \begin{bmatrix} f_a^+ \\ f_a^- \\ f_a^0 \end{bmatrix} \quad (11)$$

where,  $a = e^{j120^\circ}$ ,  $a^2 = e^{j240^\circ}$ , and  $f_a, f_b, f_c$  are unsymmetrical three-phase vectors [17]. Then the above formula is transformed into,

$$\begin{bmatrix} f_a^+ \\ f_a^- \\ f_a^0 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 \\ a^2 & a & 1 \\ a & a^2 & 1 \end{bmatrix}^{-1} \begin{bmatrix} f_a \\ f_b \\ f_c \end{bmatrix} = \frac{1}{3} \begin{bmatrix} 1 & a^2 & a \\ 1 & a & a^2 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} f_a \\ f_b \\ f_c \end{bmatrix} \quad (12)$$

Similar method is applied to define the matrix T - in Fig.3 - as,

$$T = \begin{bmatrix} 1 & a^2 & a \\ 1 & a & a^2 \\ 1 & 1 & 1 \end{bmatrix} \quad (13)$$

The positive sequence component of fundamental voltage can be readily detected. Its formula is,

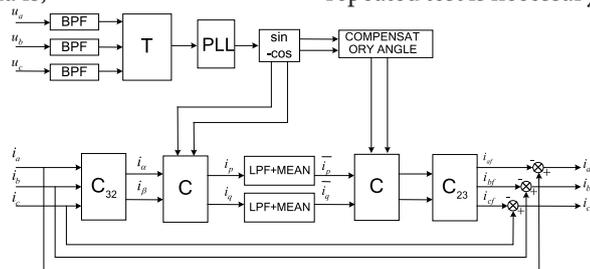


Fig.3 Block diagram of harmonic detection method with a three-phase voltage phase-locked module as well as low-pass filter and mean modules connected in series, BPF - band-pass-filter

$$u_{af}^+ = (u_{af} + a^2 u_{bf} + a u_{cf}) / 3 \quad (14)$$

According to the above theory, the calculation of  $u_{bf}$  need a  $240^\circ$  phase shift, and  $120^\circ$  for  $u_{cf}$ . The time delays at least  $2\pi/3$ . Using Formula (15),

$$\begin{cases} u_{bf} e^{j240^\circ} = -u_{bf} e^{j60^\circ} \\ u_{cf} e^{j120^\circ} = -u_{cf} + u_{cf} e^{j60^\circ} \end{cases} \quad (15)$$

Formula (14) becomes,

$$u_{af}^+ = (u_{af} - u_{bf} e^{j60^\circ} - u_{cf} + u_{cf} e^{j60^\circ}) / 3 \quad (16)$$

So, signal processing method as shown in formula (16) is able to save time. After getting the positive sequence component of fundamental voltage, the sine signal  $\sin(\omega t + \varphi_{u_n^+})$  and the cosine signal  $\cos(\omega t + \varphi_{u_n^+})$  can be easily found. The matrix C as formed is,

$$C = \begin{bmatrix} \sin(\omega t + \varphi_{u_n^+}) & -\cos(\omega t + \varphi_{u_n^+}) \\ -\cos(\omega t + \varphi_{u_n^+}) & -\sin(\omega t + \varphi_{u_n^+}) \end{bmatrix} \quad (17)$$

Moreover, because of the time delay caused by the system itself, as shown in Fig. 2, compensation module as compensatory angle is added to make up for it. Assuming that the operation of APF is with a time delay equal to  $\Delta t$ , there is a delay angle equals to  $\omega \Delta t$ . After the compensation module is applied, the new matrix  $C_{\Delta t}$  is able to guarantee the identity of the phase after transformation. The matrix  $C_{\Delta t}$  is,

$$C_{\Delta t} = \begin{bmatrix} \sin(\omega t + \varphi_{u_n^+} + \omega \Delta t) & -\cos(\omega t + \varphi_{u_n^+} + \omega \Delta t) \\ -\cos(\omega t + \varphi_{u_n^+} + \omega \Delta t) & -\sin(\omega t + \varphi_{u_n^+} + \omega \Delta t) \end{bmatrix} \quad (18)$$

And the fundamental currents after compensation are,

$$\begin{bmatrix} i_{af} \\ i_{bf} \\ i_{cf} \end{bmatrix} = C_{32}^{-1} C_{\Delta t} \begin{bmatrix} i_p \\ i_q \end{bmatrix} \quad (19)$$

The system time delay angle is predicted by theoretical analysis of configuration shown in Fig 2. There is a difference between theoretical result and the result obtained in practical operation. So the fundamental current detected by this method is different from practical one. In practical operation, repeated test is necessary for finding the accurate time delay.

## SIMULATION

### SIMULATION MODEL OF THE SHIP NETWORK UNDER NON-IDEAL CONDITIONS

Simulation was carried out using Matlab / Simulink environment, applied for the configuration of ship electrical power system resulting from Fig.2. A comparative analysis of level of the waveform distortion and time of response of harmonic detection for the active power filtering was carried out by using Mathcad environment for waveform distortion analysis. The parameters of load are  $R = 20 \Omega$  and  $L = 2 \text{ mH}$ . According to the analysis of [12], it is assumed that the time delay of the system is  $140 \mu\text{s}$ . That would be  $2.52^\circ$  if it is expressed in angle (for  $f = 50 \text{ Hz}$ ). The value of positive sequence voltage is  $220 \text{ V}$ . The initial phases of A phase voltage, B phase voltage, and C phase voltage are  $25^\circ$ ,  $-115^\circ$ , and  $120^\circ$  respectively. Then, harmonic voltage of 5th order, 7th order, and 11th order are added to the three phases. In Fig. 4 the waveforms of the distorted three-phase voltage before APF compensation are shown.

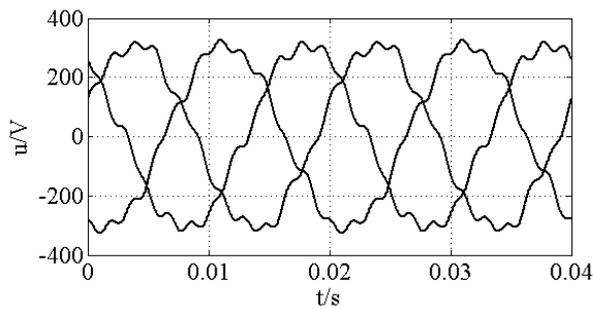


Fig.4 Distorted and asymmetrical three-phase voltage waveforms before APF compensation

### SIMULATION RESULTS UNDER DISTORTED THREE-PHASE VOLTAGE

On the basis of the aforementioned model, comparative study by simulation experiments has been carried out, with usage of the three different detection methods, i.e. with mean modules, with mean modules and a three-phase voltage phase-locked module, and the proposed method. The experimental results are based on the assumption, that in all cases the same APF control method and load are used. An assessment of the distortion level of the analyzed waveform is based on the THD concept, in the meaning [4], [5]. Other distortion factors definitions are sometimes used [11]. Fig. 5 shows the results of harmonics detection by using the mean modules. Fig. 5(a) shows the positive sequence voltage and detected positive fundamental active current of phase A. Fig. 5(b) shows the detected fundamental current of phase A. It can be seen that a quite good waveform of the fundamental current can be detected after slightly above 0.01s. Fig. 5(c) shows the spectrum of detected fundamental current of phase A. Its THD value is 2.80%.

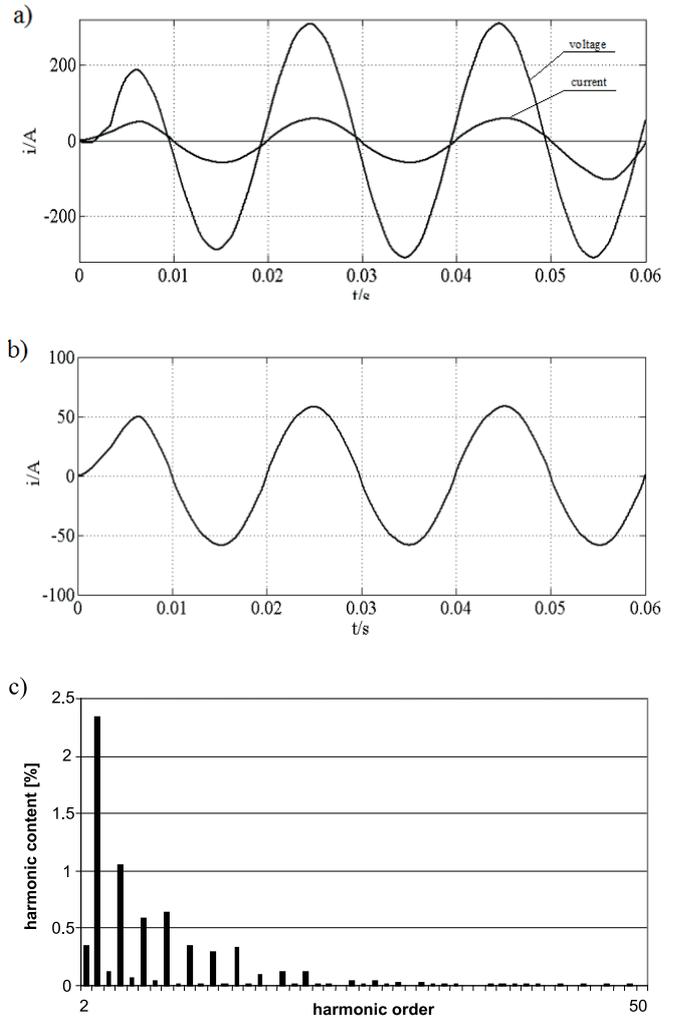


Fig.5 Simulation results of harmonic detection method with mean modules, (a) positive sequence voltage and detected positive fundamental active current of phase A, (b) detected fundamental current of phase A, (c) spectrum of detected fundamental current of phase A

Fig. 6 shows the results of harmonics detection by using the method with mean modules and a three-phase voltage phase-locked module. Fig. 6(a) shows the positive sequence voltage and detected positive fundamental active current of phase A. It is clear that there exists no phase difference. So the reactive compensation can be carried out accurately. Fig. 6(b) shows the detected fundamental current of phase A. It can be seen that a pretty ideal waveform of fundamental current can be detected after slightly above 0.01s. Fig. 6(c) shows the spectrum of detected fundamental current of phase A. Its THD value is 1.16%.

Fig.7 shows the results of harmonics detection by using the proposed method. Fig.7 (a) shows the positive sequence voltage and detected positive fundamental active current of phase A. It's clearly that there exists no phase difference. So the reactive compensation can be carried out accurately. Fig.7 (b) shows the detected fundamental current of phase A. It can be seen that a pretty ideal waveform of fundamental current can be detected after one cycle. Fig.7 (c) shows the spectrum of detected fundamental current of phase A. Its THD value is 0.52%.

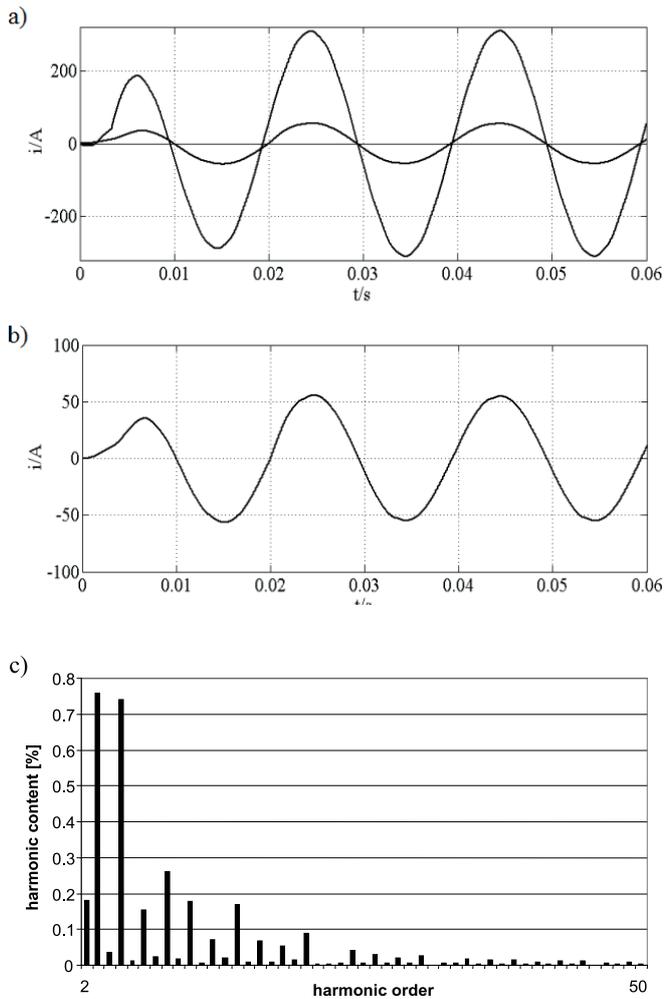


Fig.6 Simulation results of harmonic detection method with mean modules and a three-phase voltage phase-locked module, (a) positive sequence voltage and detected positive fundamental active current of phase A, (b) detected fundamental current of phase A, (c) spectrum of detected fundamental current of phase A

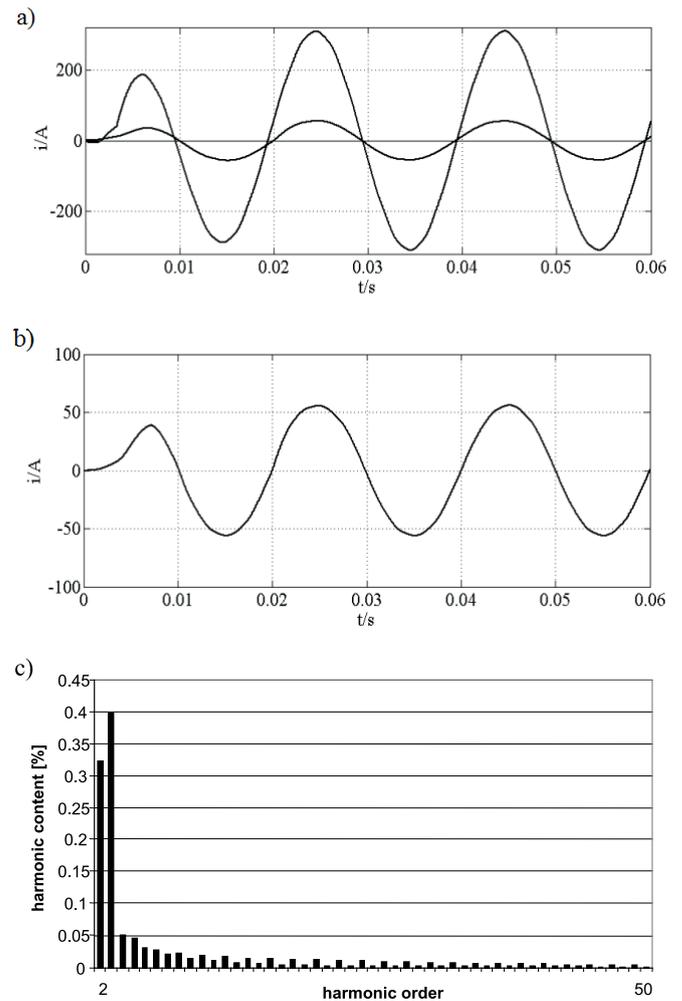


Fig.7 Simulation result of harmonic detection method with an LPF and a mean module connected in series and a three-phase voltage phase-locked module, (a) positive sequence voltage and detected positive fundamental active current of phase A, (b) detected fundamental current of phase A, (c) spectrum of detected fundamental current

Fig.7 shows the results of harmonics detection by using the proposed method. Fig.7 (a) shows the positive sequence voltage and detected positive fundamental active current of phase A. It's clearly that there exists no phase difference. So the reactive compensation can be carried out accurately. Fig.7 (b) shows the detected fundamental current of phase A. It can be seen that a pretty ideal waveform of fundamental current can be detected after one cycle. Fig.7 (c) shows the spectrum of detected fundamental current of phase A. Its THD value is 0.52%.

After the comparison of the obtained results, it can be concluded that the accurate sine & cosine signals can be detected by using the three-phase voltage phase-locked module. It helps with the accuracy of harmonic detection and reactive power compensation. The accuracy of the detection can be further improved by using a LPF and a mean module connected in series. The THD value of the detected fundamental current dropped from 1.16% to 0.31%. Though the THD value is noticeably lower, the response time increases more or less by 0.005s.

## SIMULATION RESULTS UNDER DISTORTED THREE-PHASE VOLTAGE AND SUDDEN-VARIATION RAPID VARIATION OF LOAD

An experiment of rapid variation of load is performed by load step change from 50% to 100% at 0.05 s time instant. That means the resistance drops from 20  $\Omega$  to 10  $\Omega$ , and the inductance drops from 2 mH to 1 mH. The resulting waveform of load current during the process is shown in Fig. 8.

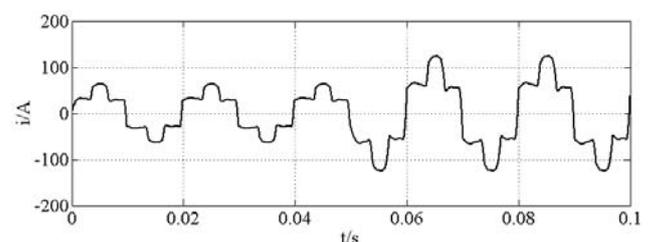


Fig.8 Simulated current during load step change by 50% at 0.05 s time instant

Next, the harmonic detection method with mean modules and the harmonic detection method with the proposed method are applied respectively. And the results are shown below. Fig. 9 shows the results obtained by using the harmonic detection method with mean modules. Fig. 9(a) shows the detected fundamental current of phase A. Fig. 9(b) and Fig. 9(c) are the spectrums of detected fundamental current of phase A before and half cycle after step load change. The THDs are 2.96% and 2.33% respectively. Further, the THD has been calculated for the cycle just after the step change (0.05 s to 0.07 s). This has been equal to 8.37%. Fig. 9(d) shows the fluctuation of detected active component and reactive component respectively. It can be seen that fluctuation is very serious.

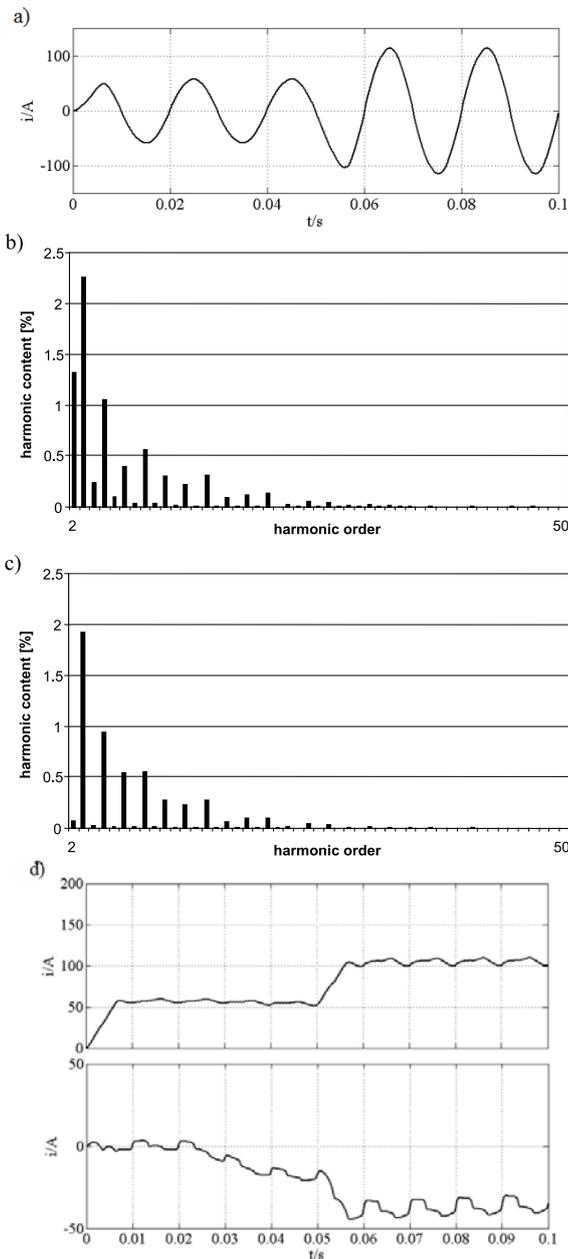


Fig.9 Simulation result of harmonic detection method with mean modules when load varies, (a) detected fundamental current of phase A, (b) spectrum of detected fundamental current for one cycle before step load change (0.03 s to 0.05 s), (c) spectrum of detected fundamental current for one cycle after step load change (0.06 s to 0.08 s), (d) detected active component and reactive component

Fig.10 shows the results obtained by using the proposed method. Fig.10(a) shows the detected fundamental current. Fig.10(b) and Fig.10(c) depict the spectra of detected fundamental current for one cycle before (0.03 s to 0.05 s) and one cycle after step load change (0.07 s to 0.09 s). The THDs of them are 1.31% and 0.5% respectively. Further, the THD has been calculated for the cycle just after the step change (0.05 s to 0.07 s) and half cycle after load step change (0.06 s to 0.08 s). These have been equal to 14.18% and 2.52%. Fig. 10(e) shows the dynamic changing processing of detected active component and reactive component. The fluctuation is a little bit lower than in Fig. 9(d).

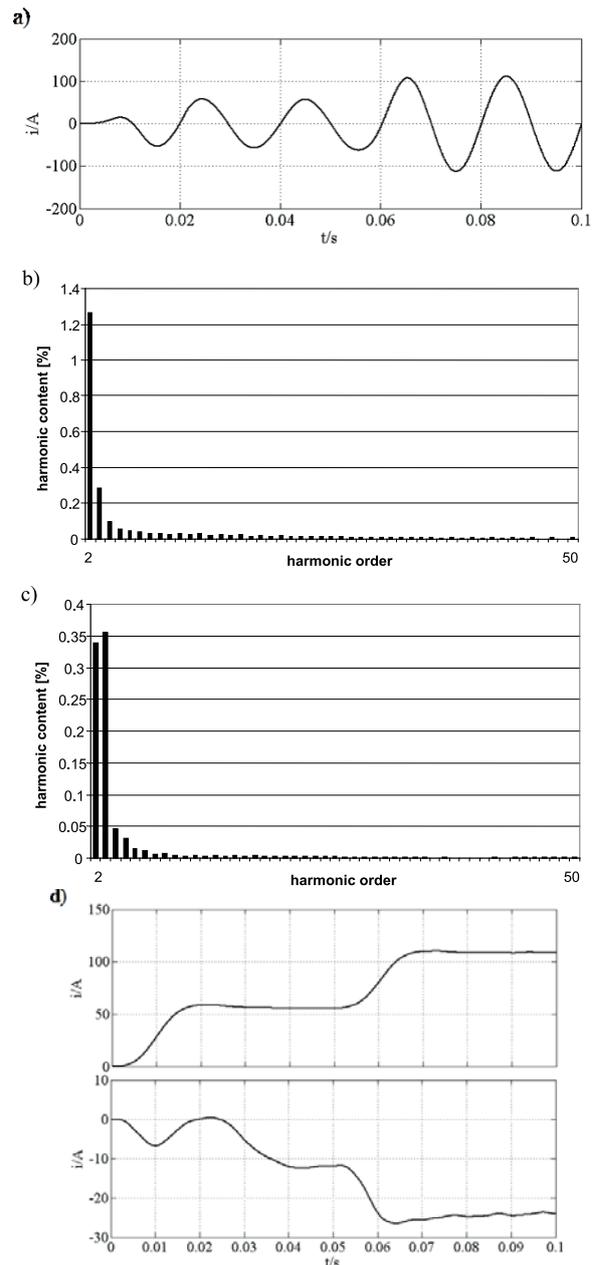


Fig.10 Simulation result of harmonic detection method with an LPF and a mean module connected in series and a three-phase voltage phase-locked module, when load varies, (a) detected fundamental current of phase A, (b) spectrum of detected fundamental current for one cycle before step load change (0.03 s to 0.05 s), (c) spectrum of detected fundamental current for one cycle after step load change (0.07 s to 0.09 s), (d) detected active component and reactive component

The careful comparison of the afore described experimental results leads to the following conclusions:

1. Within the first period immediately after the moment of load variation, THD value of fundamental current of phase A obtained by harmonic detection method with mean modules is 8.37%, and that obtained by harmonic detection method with an LPF and a mean module connected in series and a three-phase voltage phase-locked module is significantly higher, i.e. 14.18%.
2. It is also higher, if one considers its value half cycle after load step change, 2.33% versus 2.52%. But after more than 0.01s, the THD value obtained by the harmonic detection method with an LPF and a mean module connected in series, and a three-phase voltage phase-locked module, started to drop quickly. After one cycle delay (0.02s), it becomes 0.5%. Whereas in the case of the harmonic detection method by only mean modules, THD reaches 2.33% and almost keep in this level. It's obvious that, though the response speed of the harmonic detection method an LPF and a mean module connected in series and a three-phase voltage phase-locked module, is not as quick as the response speed of the harmonic detection method with mean modules only, the former solution is more accurate. The difference of the two THD values of fundamental currents is 1.83%.
3. The waveforms of active component and reactive component detected by the harmonic detection method with an LPF and a mean module connected in series reveals lower ripples of current components in comparison to the harmonic detection method with mean modules only. It shows that the proposed harmonic detection method has better performance.

### FINAL REMARKS

The proposed solution is based on average arithmetic. It combines a LPF and a mean module connected in series with a three-phase voltage phase-locked module. It is a harmonic detection method with good performance under non-ideal conditions. According to the experimental results, four conclusions are given:

1. The conventional detection method based on average arithmetic cannot detect the positive sequence component and initial phase of the distorted voltage. The proposed solution solves this problem. And active power and reactive power compensation can be carried out accurately.
2. The THD value of fundamental current detected by conventional method is above 2.3%. And the THD value of fundamental current detected by the proposed method is below 0.5%. So that the proposed solution is more accurate.
3. The response time of conventional method is about 0.01s. The response time of the proposed solution is below one cycle, but more or less 0.005s slower.

4. When load suddenly changes, the detected by the proposed solution active and reactive current components have less ripples, which shows that the proposed harmonic detection method has better performance.

In shorthand, the proposed solution is able to overcome the shortcoming of the conventional harmonic detection method with mean modules when the two solutions are applied to detect current harmonics under the condition that 3-phase voltage is distorted or load varies suddenly. Simply put, the proposed solution leads to broadly understood higher detection accuracy. The downside of the proposed solution is that its response time is 0.005s longer than that of conventional harmonic detection method based on average arithmetic. Further study will be undertaken to solve this problem.

### ACKNOWLEDGEMENT

International S&T Cooperation Program of China (2012DFG71850) & Poland (No 35-20, 2013-2014 joint research project)

### BIBLIOGRAPHY

1. Akagi H., Kanazawa Y., Nabae A., Instantaneous reactive power compensators comprising switching devices without energy storage components, IEEE Transactions on Industry Application 20 (3) (1984) 625-630
2. Ding Ju-Xia, Zhang Xiu-Feng. A harmonics-detection method without PLL and based on average theory in single-phase circuit [J]. Power System Protection and Control, 2010, 38(13): 26-30
3. Dong Tian-Bing, Cheng Han-Xiang, He Man-run. Time delay analyzing and the setting about the frequency compensation of active power Filter [J]. Electric Drive, 2011, 41(2): 24-27
4. EN Standard 50160. Voltage characteristics of electricity supplied by public distribution systems
5. IEC 61000-4-7:2007 General Guide on Harmonics and Interharmonics Measurements for Power Supply Systems and Equipment Connected Thereto
6. Jiang You-Hua, Cao Yi-long, Tang Zhong. Design and application of active power filter based on average arithmetic [J]. Power Electronics, 2010, 44(2): 42-43
7. Lu Zhen, Liang Guo-Lu, Lu Yu. Improved ip-iq harmonic detection method and simulation [J]. Power Electronics, 2012, 46(5): 81-83
8. Meng Jun-Xia, Zhou Zi-Guan, Li Guang-Hui. Detection of Distorted Current Based on ip-iq Method and Its Physical Realization [J]. Power System Technology, 2012, 36(5): 125-131

9. Mindykowski J., Xu X., Tarasiuk T., A new concept of harmonic current detection for shunt active power filters control, *Measurement* 46 (2013) 4334-4341
10. Report of the investigation of the catastrophic failure of a capacitor in the aft harmonic filter room on board RMS Queen Mary 2 while approaching Barcelona on 23 September 2010, Marine Accident Investigation Branch, December, 2011
11. Tarasiuk T., Mindykowski J., An extended interpretation of THD concept in the wake of ship electric power system research, *Measurement* 45 (2012), 207-212
12. Tu Yong-Chang, Liu Jian-Gong, Wang Wei. Research on improved algorithm for harmonic current detection based on instantaneous reactive power theory [J]. *Modern Electronics Technique*, 2013, 36(3): 145-147
13. Wang Yi-Fei, Li Yun-Wei. Three-Phase Cascaded Delayed Signal Cancellation PLL for Fast Selective Harmonic Detection [J]. *IEEE Transactions on industrial electronics*, 2013, 60(4): 1452-1463
14. Xie Dong, Hui Jing. Application and research on modified ip-iq harmonic detection method [J]. *Power Electronics*, 2013, 47(1): 95-97
15. Xu X, Mindykowski J., Chen P., Study on Hybrid Filtering Solution for Marine Electric Network, *Polish Maritime Research* 2 (65) 2010. vol.17, pp.72-78
16. Yang Sheng-Yue, Wu Min, Huang Shen-Xi. A harmonic current detection method considering frequency shift [J]. *Automation of Electric Power Systems*, 2012, 36(7): 57-61
17. Yuan Xu-Feng, Cheng Shi-Jie, Wen Jin-Wu. An improved method of instantaneous symmetrical components and its detection for positive and negative sequence current [J]. *Proceedings of the CSEE*, 2008, 28(1): 52-58
18. Zhou Qing, Mao Ya-Hui, Zhao Yong-Bin. Improved harmonic detection method without PLL and LPF [J]. *Electrical Measurement & Instrumentation*, 2012, 49(554): 27-31

## CONTACT WITH AUTHOR

Xiao-Yan Xu

e-mail: xuxy@shmtu.edu.cn

Logistics Engineering Faculty,  
Shanghai Maritime University  
Shanghai 201306  
China

Janusz Mindykowski

e-mail: janmind@am.gdynia.pl

Marine Electrical Engineering Faculty  
Gdynia Maritime University  
81-225 Gdynia  
Poland

Tomasz Tarasiuk

e-mail: totar@am.gdynia.pl

Marine Electrical Engineering Faculty  
Gdynia Maritime University  
81-225 Gdynia  
Poland

Chen Cheng

e-mail: losenessc@live.cn

Logistics Engineering Faculty  
Shanghai Maritime University  
Shanghai 201306  
China

# AN EXPERIMENTAL STUDY OF EMISSION AND COMBUSTION CHARACTERISTICS OF MARINE DIESEL ENGINE IN CASE OF CYLINDER VALVES LEAKAGE

Jerzy Kowalski, Ph.D.  
Gdynia Maritime University, Poland

## ABSTRACT

*Presented paper shows the results of the laboratory tests on the relationship between throttling of both air intake duct and exhaust gas duct and a gaseous emission from the marine engine. The object of research is a laboratory, four-stroke, DI diesel engine, operated at loads from 50 kW to 250 kW at a constant speed equal to 750 rpm. During the laboratory tests over 50 parameters of the engine were measured with its technical condition recognized as a „working properly” and with simulated leakage of both air intake valve and exhaust gas valve on the second cylinder. The results of this laboratory research confirm that the leakage of cylinder valves causes no significant changes of the thermodynamic parameters of the engine. Simulated leakages through the inlet and exhaust valve caused a significant increase in fuel consumption of the engine. Valve leakages cause an increase of the exhaust gas temperature behind the cylinder with leakage and behind other cylinders. The exhaust gas temperature increase is relatively small and clearly visible only at low loads of the engine. The increase of the temperature and pressure of the charging air behind the intercooler were observed too. Charging air temperature is significantly higher during the engine operation with inlet valve leakage. The study results show significant increases of the CO, NO<sub>x</sub> and CO<sub>2</sub> emission for all the mentioned malfunctions. The conclusion is that the results of measurements of the composition of the exhaust gas may contain valuable diagnostic information about the technical condition of the air intake duct and the exhaust gas duct of the marine engine.*

**Keywords:** marine diesel engine, exhaust gas composition, emission, exhaust gas valve leakage, inlet valve leakage

## INTRODUCTION

In classic technical solution of a ship engine room a few diesel engines are installed. One or more engines are installed as a main propulsion. Usually there are low-speed, two-stroke diesel engines operating with fixed pitch propeller [1], or medium - speed, 4-stroke diesel engines operating at a constant speed with variable pitch propeller. Moreover, there are two or more power generators and one emergency power generator in the ship engine room. The power generator usually contains medium-speed, 4-stroke diesel engine operating at a constant speed.

One of the crucial tasks of the ship's crew is keeping marine engines in proper technical condition. This task is especially important for the ship safety, but also because of reduction of the operating costs and of the negative impacts on the marine environment. The ship's crew cannot allow the main propulsion engine to emergency stop in case of unfavorable weather conditions. It should also be noted that the engine with nominal power equal to 10 MW consumes about 2 tons of fuel, emitting large quantities of toxic compounds into the atmosphere. Each engine failure causes an increase of fuel consumption and of toxic compounds emissions. For the reasons the ship's crew controls technical

parameters of engines and tries to predict their condition. Due to the specification of the onboard operation, ship engines maintenance must be planned well ahead. The main propulsion engine stop for the maintenance can be made only under certain conditions. For example, according to ISGOTT regulation [2], an oil tanker staying at marine terminal must be prepared to an immediate departure in case of danger from land (e. g. fire, terrorist attack). Planned engines maintenance must also consider the supply of spare parts.

For these reasons the correct operation of main propulsion engine boils down to the continuous diagnostics of engine technical condition. Standard approach to the diagnosis of the technical elements of engine cylinders and therefore the combustion process is the observation of the exhaust gas temperature behind individual cylinders of the engine and observation of the turbocharger operating parameters. In cases of significant temperature variations between cylinders or exceeding the threshold values (which is usually only during significant failures and/or the engine operating at the load close to the nominal) additional measurements of the cylinder combustion pressure are made. This method is ineffective. Even small changes of the engine operating

parameters can lead to a significant increase of emission levels and the fuel consumption. It should be remembered that the analysis of the combustion pressure in marine engine cylinders is considerably impeded. In the presented case the determination of the crankshaft angular position, based on compression pressure characteristics in engine cylinder is usually impossible. The reason of this is the ignition of the fuel before the TDC (Top Dead Centre) engine crankshaft position. According to the research results presented in [3], TDC determination error of one degree of crank angular position leads on average to 7 - 10 % error of mean in-cylinder pressure. The distance between the cylinder chamber and the place of installation of combustion pressure sensor is also a problem. According to the research results presented in [4], this distance can result in the displacement of the characteristics of combustion pressure in the cylinder of up to 6 - 7 degrees of the angular position of the crankshaft. The presented conditions cause the quality and quantity of diagnostic signals to limit significantly during the onboard operation, which in consequence leads to a significant increase of the engine maintenance costs.

International Maritime Organization introduced a legislation imposing on the ship owners to control and reduce nitric oxides emission from the engines with a nominal power over 130kW [5] installed onboard after 2001. These regulations, among others, allow to measure toxic compounds emissions. This provides an opportunity to obtain additional information about the engine condition. The author believes that the composition of the exhaust gases can bring diagnostic signals of the engine condition during onboard operation.

Construction of diesel engines with mechanically controlled fuel systems were intensively studied in 20th century. Unfortunately, most publications were focused on the aspects related to improving the energy efficiency. Less attention was paid to reduction of the toxic compounds emission. Only a few new experimental works focusing on the emission characteristics of marine diesel engines are available.

Sarvi et al. [6], [7], [8] present emission characteristics of a large medium-speed diesel engine with parameters similar to those of marine engines. Desantes et al. [9], [10] show influence of fuel injection characteristics on emission characteristics and parameters of combustion in one cylinder engine. Weiser, in his doctor's thesis [11], presents results of direct measurements on 9- cylinder, medium-speed engine. Works about comparing the emission characteristics of diesel engines with direct and indirect fuel injection [12], different types of fuels [12], injection strategies [13] – [16], and nozzle geometries [17], are also available. The experimental investigations of diesel combustion with water emulsion [18], [19] and different fuel bio-components [20] – [30] and much frequently are presented.

This paper presents the results of the experimental studies on the effects of the cylinder valves leakage (both at air inlet and exhaust outlet) on the composition of the exhaust gas.

## LABORATORY STAND

The object of research is a three-cylinder, four-stroke, turbocharged, laboratory engine with mechanically controlled fuel pumps, inlet and exhaust valves. The fuelling system of the engine consists of Bosch fuel pumps and multi-hole fuel nozzles. The engine was loaded by an electrically-connected, water resistance generator. During the tests the engine was fuelled by diesel oil and operated at a constant speed equal to 750 rpm. This type of engine is commonly used onboard as a power generator or a main propulsion system with variable pitch propeller [1]. The load and speed of the engine, speed of the turbocharger and parameters of cooling, fuelling, lubricating, and air exchange systems, were measured. The composition of exhaust gas was also recorded by using an electrochemical gas analyser with infrared carbon dioxide sensor. Pressure, temperature and humidity of air were also recorded by laboratory equipment. All the mentioned results were recorded with a 1- second sampling time. Injection pressures and pressures of the combustion in all cylinders of the engine were also collected with a resolution of 0,5 degree of the engine shaft rotation angle. The scheme of the laboratory stand is presented in [1] and the engine parameters are given in Tab.1.

Tab.1. Parameters of the test engine

Parameter	Value	Unit
Max. electric power	250	kW
Rotational speed	750	rpm
Number of cylinders	3	–
Cylinder diameter	250	mm
Stroke	300	mm
Compression ratio	12,7	–
Number of inlet valves per cylinder	1	–
Number of exhaust valves per cylinder	1	–
Valves diameter	98	mm
Valves shift	24	mm

The experimental study consists of 4 stages of 3 observations each with simulations of different malfunctions of both air inlet valve and exhaust valve in second cylinder of the engine. During each start of the observation, the engine was loaded to a maximum load equal to 250 kW, measured as the electric power of the generator. After stabilizing the temperature of the exhaust gas behind the turbine, the engine operating parameters were recorded for 3 to 5 minutes. After that, the load of the engine was decreased by 10 kW and, after stabilizing the temperature of the exhaust gas behind the turbine, the engine operating parameters were recorded again. Observation was continued with loads up to 50 kW. The engine did not work with a load of 190 kW due to resonance vibrations.

The stages of the experiment were set as follows:

- 1st stage - during the operation of the engine assumed as “working properly”,

- 2nd stage - during the operation of the engine with a small simulated leakage of the air inlet valve,
- 3rd stage - during the operation of the engine with a small simulated leakage of the exhaust valve,
- 4th stage - during the operation of the engine with a large simulated leakage of the exhaust valve.

The mentioned simulations of valves leakage were obtained by drilling holes in the valve, according to the scheme presented in Fig.2. A small leakage was simulated by drilling one hole of 1mm diameter and a large leakage was simulated by drilling four holes of 2 mm diameter.

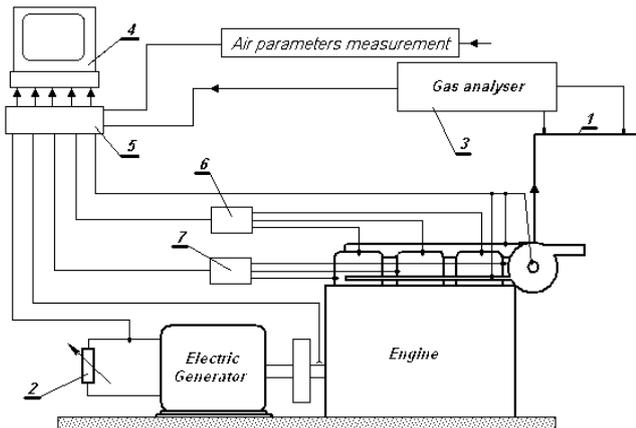


Fig.1. The laboratory stand scheme. 1 – exhaust duct, 2 – water resistance, 3 – gas analyser, 4 – computer, 5 – A/C converter, 6 – combustion pressure indicator, 7 – injection pressure indicator

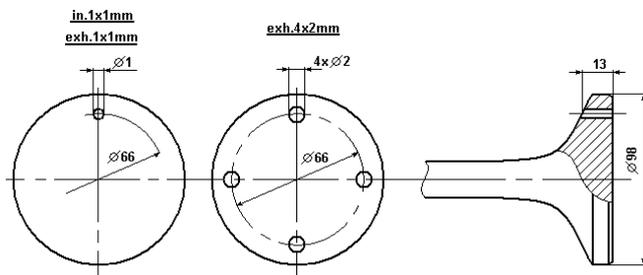


Fig.2. Scheme of the valve leakage simulation

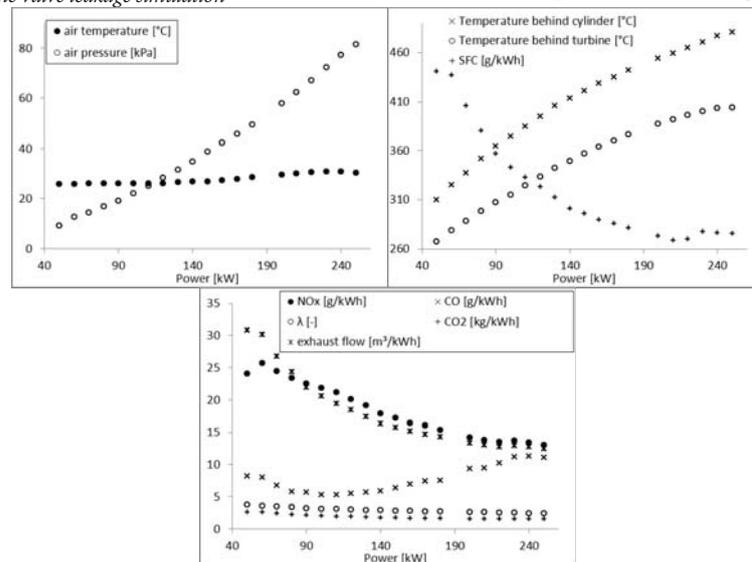


Fig.3. Parameters of the engine assumed “working properly”

## RESULTS AND DISCUSSION

Figures from 4. to 6. present average changes of measured and calculated parameters of the engine with simulated malfunctions in relation to the parameters of the operation of the engine recognized as “working properly” (Fig.3).

### INLET VALVE LEAKAGE

The air intake valve leakage was simulated by drilling one hole in the valve with a diameter of 1mm. The location of the hole was selected in a way to separate the surfaces of the valve seat from leaking gas into cylinder.

The air intake valve leakage in the second engine cylinder causes penetration of the gas mixture to the air intake duct. The mixture of gases in the cylinder has a higher temperature, and therefore, according to the results presented in Fig.4a and Fig.4b, the mentioned malfunction causes an increase of the air pressure and temperature behind the intercooler. The results of the measurements for the engine assumed as “working properly”, presented in Fig.3a, show that the increase of the engine load leads to an increase of the air pressure behind the intercooler at a stable temperature. The air intake valve leakage causes an increase of the air temperature and the air pressure at low loads of engine operation. The result of the increase of thermodynamic parameters of air in the air intake duct is an increase of the temperature and pressure of air entering also other engine cylinders. These changes, however, do not change the amount of supplied air in relation to fuel consumption in all considered engine loads. This conclusion results from the lack of oxygen content changes in the exhaust gas for all considered loads of the engine (the result were not presented in this paper). Therefore, the increase of air pressure is compensated in the object of research by the temperature rise, and thus the amount of air supplied to the cylinders increases proportionally to fuel consumption increase. The change of temperature and pressure of the air behind the intercooler results in a deterioration of the combustion process in all engine cylinders. Fig.5c shows

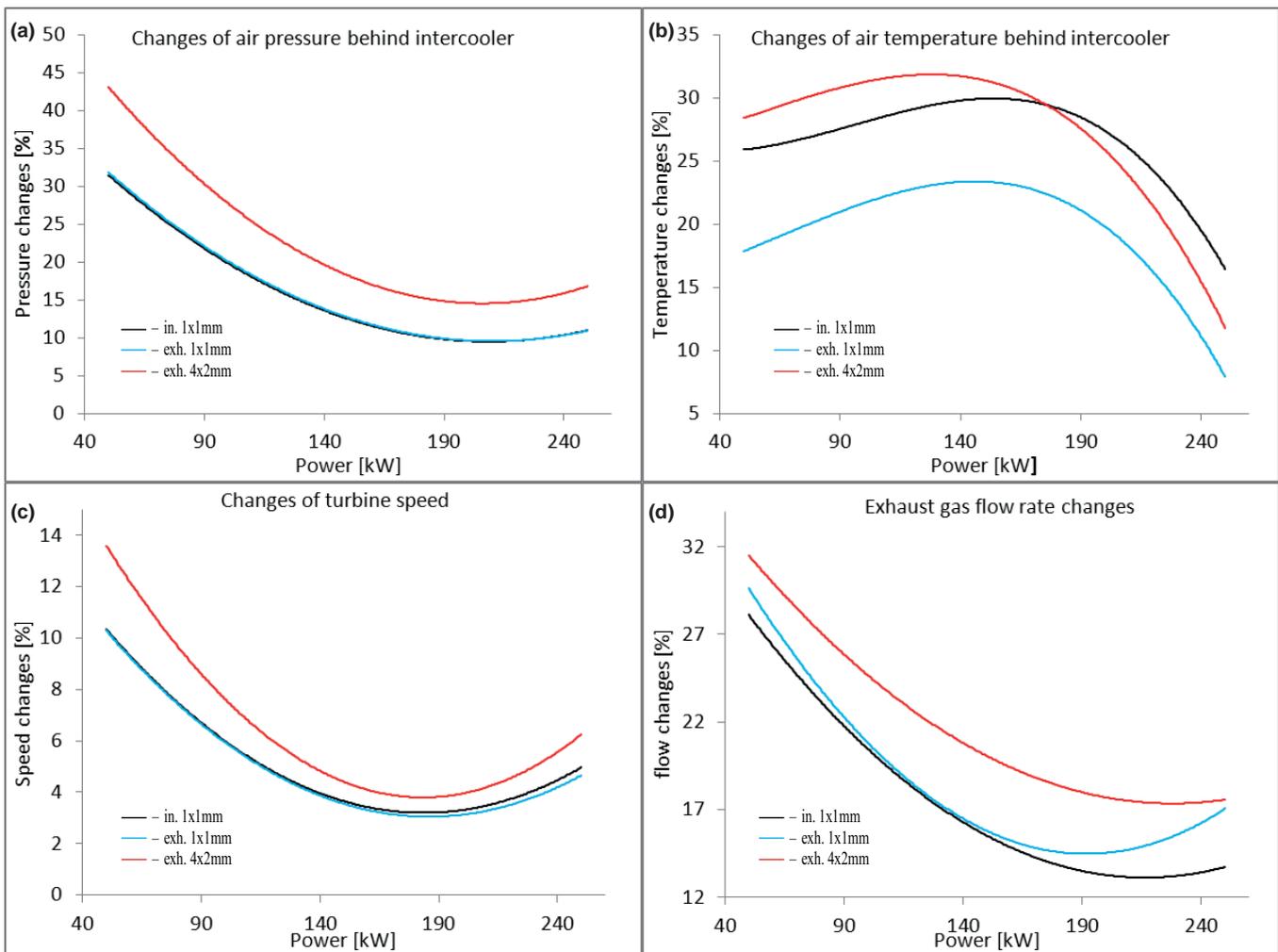


Fig.4. Changes of the air parameters in front of the cylinders, changes of the turbine speed and the exhaust gas flow rate

changes of the specific fuel consumption (SFC) with changes of the engine loads. According to the presented results, the 10-27% increase in fuel consumption follows, clearly visible especially at low loads engine operation. The increase of the fuel consumption also increases amounts of the carbon dioxide ( $\text{CO}_2$ ) emission, which was presented in Fig.6d. The characteristics of carbon monoxide (CO) emission indicate the deterioration of the combustion process. According to Fig.6c, the air intake valve leakage causes about 60 % increase in amounts of CO emission in relation to those from the engine assumed “working properly”, presented in Fig.3c. The deterioration of the combustion process causes the increase of the exhaust gas temperature behind the engine cylinders. According to the results presented in Fig.5a and Fig.5b, changes of the exhaust gas temperature behind cylinders are similar for all engine cylinders. Fig.5d shows the change in the exhaust gas temperature measured at the exhaust gas duct behind the turbocharger. According to the presented results the air intake valve leakage causes a relatively small increase of temperature.

As a result, an increase in the turbocharger speed occurs. According to the results in Fig.4c, the increase of the rotational speed of the turbocharger is so small that it may go unnoticed during the onboard operation. The change in

the turbocharger and the engine cooperation as well as the mentioned conditions cause an increase of the volumetric exhaust gas flow rate. Fig.4d presents changes of the exhaust gas flow rate according to the engine load conditions. A volumetric exhaust gas flow rate was obtained by calculation with the use of the carbon balance method, in accordance with ISO 8178 standard regulation [31]. According to the presented results, the 10-28 % increase of volumetric exhaust gas flow rate followed. The largest changes of the presented parameters were observed for low engine loads. The reason for this is the constant size of the simulated valve leakage. The air inlet valve leakage causes a leakage of the gaseous mixture from a cylinder, which, in relation to combusted mixture mass in the cylinder, increases at low engine load operation.

Fig.6a and Fig.6b present changes in values of both fraction and emission of nitric oxides ( $\text{NO}_x$ ) in relation to the operation of the engine assumed “working properly” for all considered loads of the engine. According to the presented results, the air intake valve leakage causes a noticeable decrease of the  $\text{NO}_x$  fraction in exhaust gas. The probable reason for this is a decrease of the combustion temperature in the leaky cylinder. However, this thesis has not been proved in this research. It should be noted that the deterioration of the combustion process in the leaky cylinder and the changes

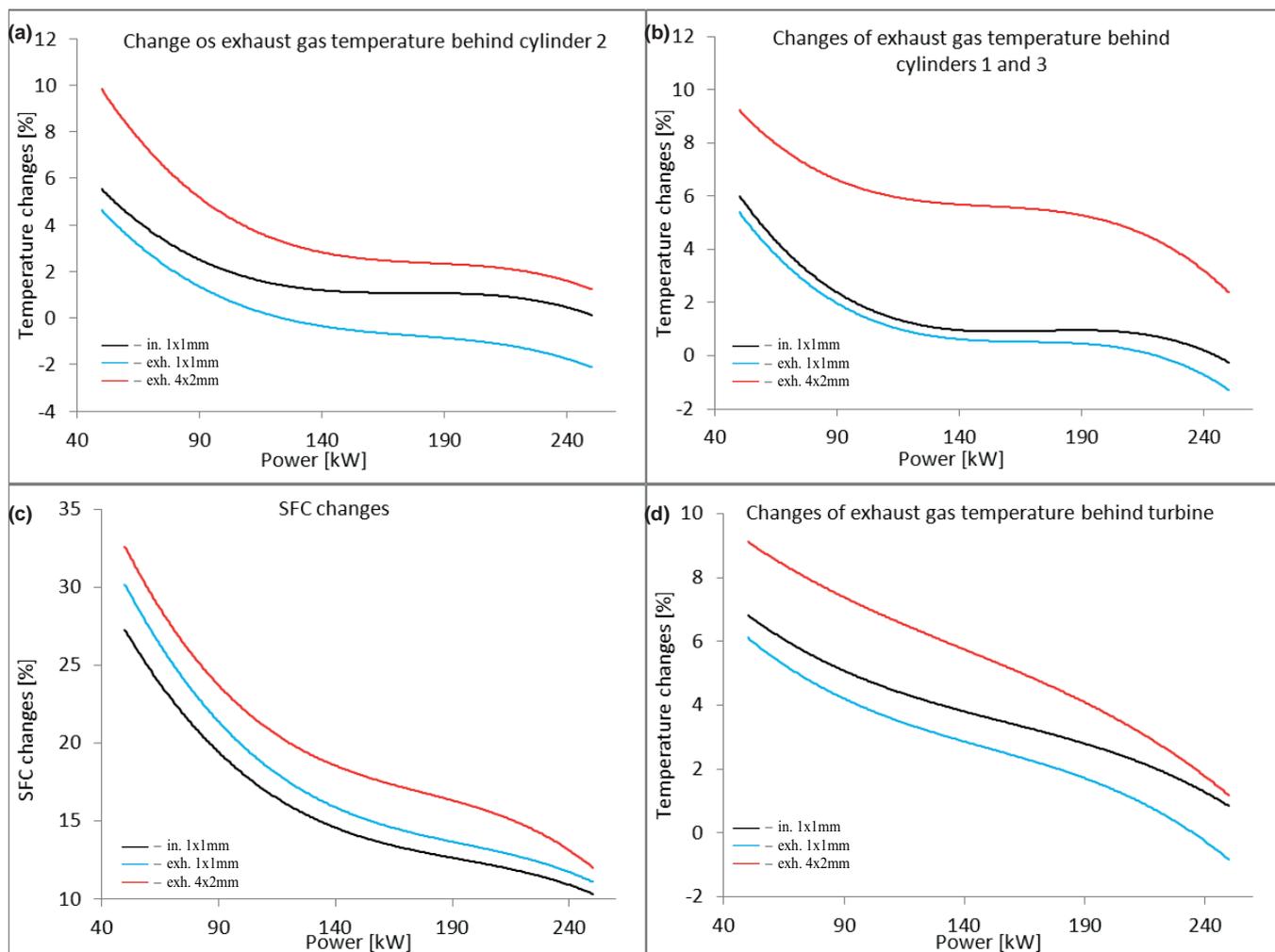


Fig.5. Changes of exhaust gas parameters and specific fuel consumption in relation to engine assumed "working properly"

in the engine and the turbocharger cooperation result in the fuel consumption increase and the increase of the exhaust gas volume. This implies an increase of  $\text{NO}_x$  emission, despite a decrease of  $\text{NO}_x$  fraction in exhaust gas. Changes of  $\text{NO}_x$  emission in relation to  $\text{NO}_x$  emission from the engine assumed "working properly" are presented in Fig.6b.

According to the presented results, a relatively small leakage of air intake valve can cause significant changes in the emission of  $\text{CO}_2$ ,  $\text{CO}$  and  $\text{NO}_x$  and the fuel consumption increase.

### EXHAUST VALVE LEAKAGE

Exhaust valve leakage is a very common malfunction. It is caused by the damage of the valve seat due to hot and corrosive exhaust gases.

In this experiment the exhaust valve leakage was simulated by drilling holes in the valve as shown in Fig.2. Two leakages were simulated:

- a small leakage, which consists of one hole with a diameter of 1 mm, its cross - section area corresponding to leakage of 0,1 mm thickness of valve seat with angle equal to 10 degrees,

- a large leakage, which consists of four holes with diameters of 2 mm, which corresponds to leakage of 0.3 mm thickness of the valve seat with angle equal to 55 degrees.

The exhaust valve leakage causes the penetration of the gas mixture to the exhaust gas duct during the combustion process in the cylinder. The result is a reduction of the maximum combustion pressure in the cylinder and the increase of the exhaust gas temperature behind the cylinder. According to Fig.5a and Fig.5b, the increase of the exhaust gas leakage causes the increase of the exhaust gas temperature behind a damaged cylinder, but also for cylinders without leaks. The presented increase of the exhaust gas temperature is relatively low and clearly visible only at low loads of the engine. Accordingly, the observation of this parameter allows to detect a large leakage of the exhaust valves only. It should be noted that simulated leakage will not initiate the automatic alarm systems. A leakage of the combustible mixture to the exhaust gas duct causes a deterioration of the engine efficiency. According to Fig.5c, the simulated malfunctions producing even small changes of the exhaust gas temperature behind cylinders, cause a large increase of SFC, reaching up to 30 % at low engine loads operation. Similarly to the case

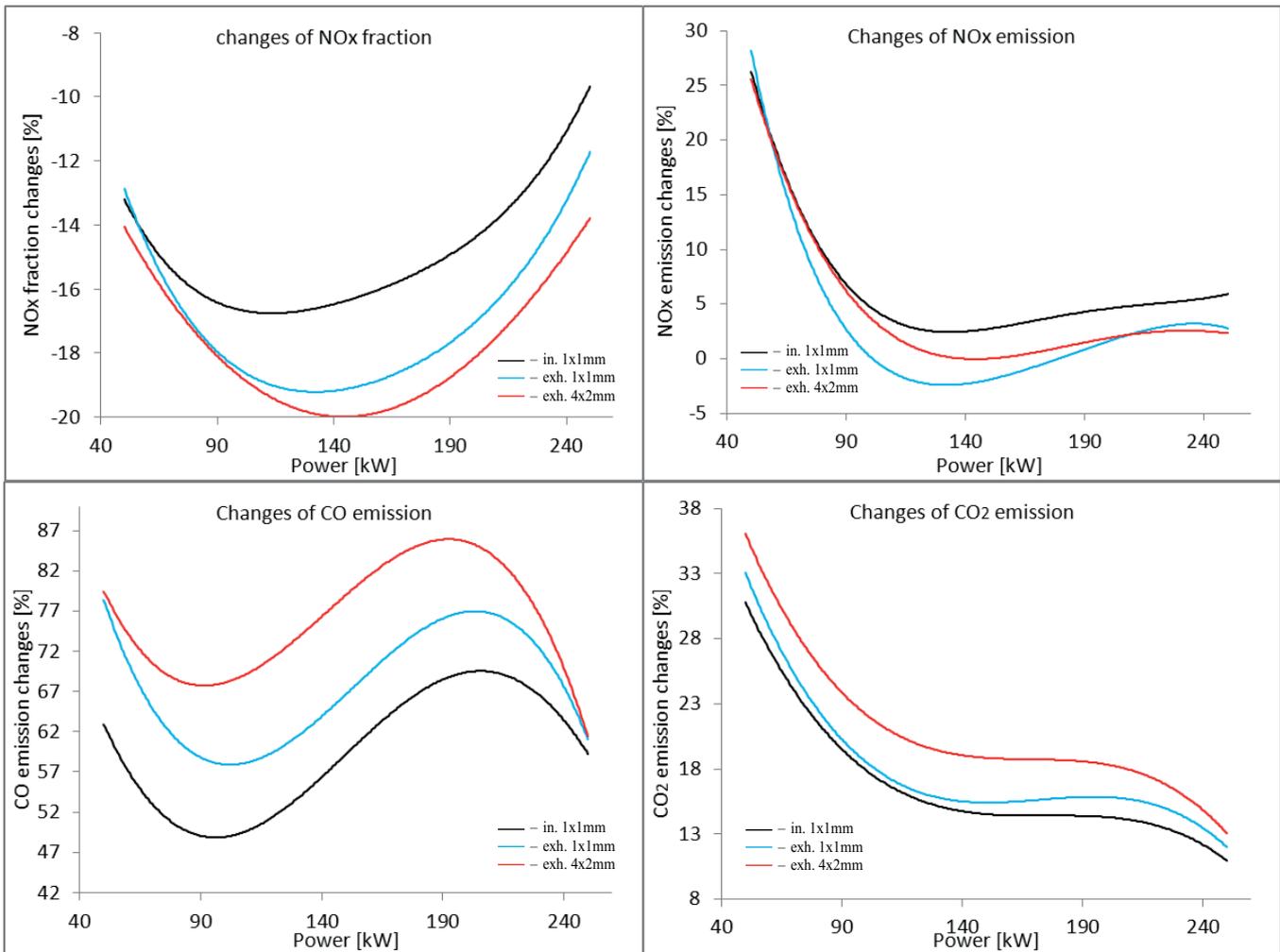


Fig.6. Exhaust gas composition changes

of the air intake valve leakage, the amount of air in relation to the amount of burned fuel undergoes no visible change (this result is not presented in this paper). Leakages of the gaseous mixture into the exhaust gas duct and the increase of the exhaust gas temperature behind all cylinders cause the increase of the rotational speed of the turbocharger. Fig.4c presents the turbocharger speed changes caused by the simulated malfunctions. The mentioned results are compared to the parameters of the operation of the engine assumed as “working properly”, presented in Fig.3b. According to the presented results, the increase of the leakage causes the increase of the turbocharger speed. Volumetric flow rate of the exhaust gas, calculated according to the mentioned algorithm, rises with the increase of the valve leakage. As reported by the results presented in Fig.4d, the exhaust valve leakage in one cylinder causes even a 30% increase of the volumetric exhaust gas flow rate at low engine loads. This result is also the effect of the exhaust gas temperature increase behind the turbine, which is presented in Fig.5d.

Exhaust valve leakage in one of the engine cylinders causes an increase of both temperature and pressure of the charging air in front of the cylinders. According to Fig.4a and Fig.4b, the characteristics of the charging air temperature and pressure

are qualitatively similar to the previously discussed inlet valve leakage. Although the effect of both the discussed valve malfunctions is similar, the genesis of these phenomena is different. In the case of the exhaust valve leakage the increase of the charging air pressure and temperature in front of the engine cylinders is a result of the increased compressor efficiency, caused by the increase of turbine speed. The inlet valve leakage causes, apart from the mentioned increase of the compressor efficiency, the additional increase of the charging air temperature as a result of the mixing of air and leaking combustible mixture from the cylinder. The result of this is that the same size of a leakage at the inlet and exhaust valve causes the same increase of the charging air pressure (Fig.4a), but the charging air temperature is significantly higher during the engine operation with inlet valve leakage (Fig.4b).

Increased fuel consumption of the engine is the direct cause of the CO<sub>2</sub> emission increase. According to the measurement results presented in Fig.6d, the increase of the exhaust valve leakage causes an increase of the CO<sub>2</sub> emission, especially at low engine loads operation. As mentioned earlier, the exhaust valve leakage causes the decrease of the maximum combustion pressure and the deterioration of the combustion process. The result of this is a large increase of

the CO emission, which rises with an increase of the exhaust valve leakage. Changes of the CO emission in all considered loads of the engine are presented in Fig.6c. As for the case of the inlet valve leakage, the described malfunction reduces the NO<sub>x</sub> fraction in the exhaust gas. In Fig.6a a drop of the NO<sub>x</sub> fraction together with the increase of the leakage is also observed. This is caused by the combustion pressure decrease. There is a probability that the temperature of the combustion process also decreased (thus implicitly indicated also the increase in the CO emission), but it was not clearly confirmed by the measurement results. The correlation of changes in the NO<sub>x</sub> fraction and the exhaust gas flow causes, at low loads of the engine, a significant increase of the NO<sub>x</sub> emission. The increase of load of the engine operating with the exhaust valve leakage in one cylinder, results in a reduction of the NO<sub>x</sub> emissions in comparison to the values observed for the engine assumed as “working properly”. It should be noted that the increase of the engine load rises the amount of combusted mixture in the cylinder, which, at the same cross-section area of valve leakage, provides a relative decrease of leakage through the leaking valve. The mentioned leakage depends also on the pressure difference in front and behind the leaking valve. According to the obtained results the simulated leakages are too small to increase the combusted mixture amount (consequently the load) to such an extent as to cause a proportional increase of the leakage.

## CONCLUSIONS

This paper presents the results of laboratory tests carried out on the four-stroke diesel engine for marine applications. The study consisted of determining the impact of the leakage of both the air intake valve and the exhaust gas valve on the engine operating parameters, including the composition of exhaust gas.

The obtained results of the measurements allow to formulate the following conclusions:

1. The simulated leakages in the inlet and the exhaust valve caused a significant increase in the fuel consumption of the engine. It was also observed that the amount of air delivered to the engine cylinders increases in proportion to the increase of the fuel consumption.
2. The valve leakages cause an increase of the exhaust gas temperature behind the cylinder with a leakage and behind the other cylinders. The presented exhaust gas temperature increase is relatively small and clearly visible only at low loads of the engine. Therefore, the observation of this parameter allows to detect only large leakages of the exhaust valves.
3. The valve leakages taken into account in this experiment cause an increase of temperature and pressure of charging air behind the intercooler. This increase is caused by the risen efficiency of the turbocharger. In case of the inlet valve leakage, an additional increase of temperature was observed, caused by a gas leakage into the air intake duct. Charging air temperature is significantly higher during the engine operation with

inlet valve leakage in one cylinder.

4. In all the simulated malfunctions the deterioration of the combustion process in leaky cylinder and the change of the co-operation between the engine and the turbocharger were observed. Relatively small valve leakages caused significant changes in CO<sub>2</sub>, CO and NO<sub>x</sub> emissions.
5. The largest changes of the mechanical and thermodynamic parameters of the engine and gaseous emissions were observed for low engine loads. The reason for this is the constant size of the simulated valve leakage in all the considered engine loads. The simulated leakages caused an amount of gaseous mixture to flow out from the cylinder, which, in relation to the mass of the mixture in the cylinder, is larger at low loads of engine operation.

## NOMENCLATURE

CO – carbon monoxide

CO<sub>2</sub> – carbon dioxide

ISGOTT - International Safety Guide for Oil Tankers and Terminals

NO<sub>x</sub> – nitric oxides

SFC – Specific fuel consumption [g/kWh]

TDC – top dead centre of crankshaft position

## ACKNOWLEDGMENTS

The project was supported by the National Science Centre in Poland, granted on the basis of decision No. DEC-2011/01/D/ST8/07142.

## BIBLIOGRAPHY

1. J. Carlton: Marine Propellers and Propulsion. Third Ed., Elsevier Ltd., 2012.
2. International Safety Guide for Oil Tankers and Terminals. ICS, OCIMF, IAPH, 2006.
3. R. Pawletko, S. Polanowski: Research of the influence of marine diesel engine Sulzer AL 25/30 load on the TDC position on the indication graph, Journal of KONES , Powertrain and Transport 17 ,2010, pp. 361 – 368.
4. R. Pawletko, S. Polanowski : Influence of gas channels of medium speed marine engines on the accuracy of determination of diagnostic parameters based on the indicator diagrams. Journal of Polish CIMAC 7,2012, pp. 139 – 146.
5. Marine Engine IMO Tier II Programme 2013

6. A. Sarvi, C. J. Fogelholm, R. Zevenhoven : Emissions from large-scale medium-speed diesel engines: 1. Influence of engine operation mode and turbocharger. *Fuel processing technology* 89 , 2008, pp. 510 – 519.
7. A. Sarvi, C. J. Fogelholm, R. Zevenhoven: Emissions from large-scale medium-speed diesel engines: 2. Influence of fuel type and operating mode. *Fuel processing technology* 89, 2008, pp. 520 – 527.
8. A. Sarvi, R. Zevenhoven : Large-scale diesel engine emission control parameters. *Energy* 35 , 2010, pp. 1139–1145.
9. J.M. Desantes, J. Benajes, S. Molina, C.A. Gonzalez : The modification of the fuel injection rate in heavy-duty diesel engines. Part 1: Effects on engine performance and emissions. *Applied Thermal Engineering* 24 , 2004, pp. 2701–2714.
10. J.M. Desantes, J. Benajes, S. Molina, C.A. Gonzalez: The modification of the fuel injection rate in heavy-duty diesel engines. Part 2: Effects on combustion. *Applied Thermal Engineering* 24 ,2004, pp. 2715–2726.
11. G.A. Weiser: Modeling of Combustion and Nitric Oxide Formation for Medium-Speed DI Diesel Engines: A Comparative Evaluation of Zero- and Three-Dimensional Approaches. Doctor's Thesis, Swiss Federal Institute of Technology , Zurich, 2001.
12. J. Huang, L. Lin, Y. Wang, J. Qin, A. P. Roskilly, L. Li, T. Ouyang, Y. Yu : Experimental study of the performance and emission characteristics of diesel engine using direct and indirect injection systems and different fuels. *Fuel Processing Technology* 92 , 2011, pp. 1380–1386.
13. K. Verbezen, A.J. Donkerbroek, R.J.H. Klein-Douwel, A.P. van Vliet, P.J.M. Frijters, X.L.J. Seykens, R.S.G. Baert, W.L. Meerts, N.J. Dam, J.J. ter Meulen: Diesel combustion: In-cylinder NO concentrations in relation to injection timing. *Combustion and Flame* 151 , 2007, pp. 333–346.
14. E. Buyukkaya, M. Cerit: Experimental study of NOx emissions and injection timing of a low heat rejection diesel engine. *International Journal of Thermal Sciences* 47 , 2008, pp.1096–1106.
15. T. Thurnheer, D. Edenhauser, P. Soltic, D. Schreiber, P. Kirchen, A. Sankowski : Experimental investigation on different injection strategies in a heavy-duty diesel engine: Emissions and loss analysis. *Energy Conversion and Management* 52 , 2011, pp. 457–467.
16. A. Parlak, H. Yasar, C. Hasimoglu, A. Kolip: The effects of injection timing on NOx emissions of a low heat rejection indirect diesel injection engine. *Applied Thermal Engineering* 25 , 2005, pp. 3042–3052.
17. R. Payri, F.J. Salvador, J. Gimeno, L.D. Zapata: Diesel nozzle geometry influence on spray liquid-phase fuel penetration in evaporative conditions. *Fuel* 87 , 2008, pp. 1165–1176.
18. M. Ebna Alam Fahd, Yang Wenming, P.S. Lee, S.K. Chou, Christopher R. Yap: Experimental investigation of the performance and emission characteristics of direct injection diesel engine by water emulsion diesel under varying engine load condition. *Applied Energy* 102 , 2013, pp. 1042–1049.
19. W.M. Yang, H. An, S.K. Chou, S. Vedharaji, R. Vallinagam, M. Balaji, F.E.A. Mohammad, K.J.E. Chua: Emulsion fuel with novel nano-organic additives for diesel engine application. *Fuel* 104 ,2013, pp. 726–731.
20. A. Kumar Agarwal, A. Dhar: Experimental investigations of performance, emission and combustion characteristics of Karanja oil blends fuelled DIC I engine. *Renewable Energy* 52 , 2013, pp. 283 – 291.
21. A.K. Hossain, M. Ouadi, S.U. Siddiqui, Y. Yang, J. Brammer, A. Hornung, M. Kay, P.A. Davies: Experimental investigation of performance, emission and combustion characteristics of an indirect injection multi-cylinder CI engine fuelled by blends of de-inking sludge pyrolysis oil with biodiesel. *Fuel* 105, 2013, pp. 135–142.
22. D.C. Rakopoulos : Combustion and emissions of cottonseed oil and its bio-diesel in blends with either n-butanol or diethyl ether in HSDI diesel engine. *Fuel* 105, 2013, pp. 603–613.
23. S. Saravanan, G. Nagarajan, S. Sampath : Combined effect of injection timing, EGR and injection pressure in NOx control of a stationary diesel engine fuelled with crude rice bran oil methyl ester. *Fuel* 104 , 2013, pp. 409–416.
24. P. Behera, S. Murugan: Combustion, performance and emission parameters of used transformer oil and its diesel blends in a DI diesel engine. *Fuel* 104 , 2013, pp. 147–154.
25. B. İşcan, H. Aydın : Improving the usability of vegetable oils as a fuel in a low heat rejection diesel engine. *Fuel Processing Technology* 98 , 2012, pp. 59–64.
26. S. Altun, H. Bulut, C. Oner: The comparison of engine performance and exhaust emission characteristics of sesame oil–diesel fuel mixture with diesel fuel in a direct injection diesel engine. *Renewable Energy* 33 , 2008, pp. 1791–1795.
27. Ch. Yao, C.S. Cheung, Ch. Cheng, Y. Wang, T.L. Chan, S.C. Lee: Effect of Diesel/methanol compound combustion on Diesel engine combustion and emissions. *Energy Conversion and Management* 49, 2008, pp. 1696–1704.

28. H. Aydin, C. Ilkılıç : Effect of ethanol blending with biodiesel on engine performance and exhaust emissions in a CI engine. *Applied Thermal Engineering* 30 , 2010, pp. 1199–1204.
29. J. Huang, Y. Wang, J. Qin, A. P. Roskilly: Comparative study of performance and emissions of a diesel engine using Chinese pistache and jatropha biodiesel. *Fuel Processing Technology* 91 ,2010, pp. 1761–1767.
30. Y.D. Wang, T. Al-Shemmeri, P. Eames, J. McMullan, N. Hewitt, Y. Huang, S. Rezvani: An experimental investigation of the performance and gaseous exhaust emissions of a diesel engine using blends of a vegetable oil. *Applied Thermal Engineering* 26, 2006, pp. 1684–1691.
31. International Organization of Standardization : ISO 8178 regulation.

## CONTACT WIHT AUTHOR

Jerzy Kowalski

Department of Engineering Sciences  
Gdynia Maritime University  
Morska 81-87 Str.  
81-225 Gdynia  
Poland

tel.: +48 58 6901331

fax: +48 58 6901399

e-mail: [jerzy95@am.gdynia.pl](mailto:jerzy95@am.gdynia.pl)

# TRYTON SUPERCOMPUTER CAPABILITIES FOR ANALYSIS OF MASSIVE DATA STREAMS

Henryk Krawczyk, Prof.

Michał Nykiel, Ms. C.

Jerzy Proficz, Ph. D.

Gdańsk University of Technology, Poland

## ABSTRACT

*The recently deployed supercomputer Tryton, located in the Academic Computer Center of Gdansk University of Technology, provides great means for massive parallel processing. Moreover, the status of the Center as one of the main network nodes in the PIONIER network enables the fast and reliable transfer of data produced by miscellaneous devices scattered in the area of the whole country. The typical examples of such data are streams containing radio-telescope and satellite observations. Their analysis, especially with real-time constraints, can be challenging and requires the usage of dedicated software components. We propose a solution for such parallel analysis using the supercomputer, supervised by the KASKADA platform, which with the conjunction with immerse 3D visualization techniques can be used to solve problems such as pulsar detection and chronometric or oil-spill simulation on the sea surface.*

**Keywords:** supercomputer, data streams, distributed systems, radio-telescope, satellite

## INTRODUCTION

The emerging need to analyze radio-telescope data means that scientists are required to use the most sophisticated HPC equipment in their research. The amount of data produced by such devices can only be processed by extremely efficient computers, usually large computer clusters. Thus, HPC solutions seem to be the most appropriate for radio-telescope observation data analysis. Because of the (soft) real-time constraints, dedicated software solutions need to be used.

Moreover, due to the possible interference with the space observations, such powerful computation equipment cannot be located next to the radio telescope, so fast and reliable data transfer is yet another problem.

Currently used at Nicolaus Copernicus University in Torun, radio telescope RT-4 has a 32 m diameter and a number of receivers for various frequency bands, however the proximity to urban areas decreases its usability. Thus, a new planned radio-telescope RT-90+ is going to have a 90 m (or even larger) diameter and will be located in Tuchola Forest, where the interferences are significantly lower (even 1 to 1000 ratio). Such a sophisticated device gives a great research opportunity, but it also introduces new challenges. One of the most important ones is the amount of data gathered during the observations. The data streams produced by the radio-telescope need to be transferred, decoded and analyzed by a powerful computer system, and due to the usual interference produced by digital devices, located at some distance from the radio-telescope.

Considering the above constraints, we perceive the Academic Computer Center of Gdansk University of Technology to be an excellent choice for the gathered data analysis. It is the nearest data center to the proposed radio-telescope's location and the connection to the PIONIER network enables fast data transfer. The Center is also equipped with the most powerful computer in Poland (1.37 PFLOPS) and finally, thanks to the Center of Excellence in Scientific Application Development Infrastructure (CD NIWA), it provides a suitable base for proper software solutions designated for the described data analysis.

The rest of the paper provides the information about the potential of the Academic Computer Center including the Tryton supercomputer. Two following sections present more information about the processed data, the next one describes the supercomputer including its internal characteristics and networking capabilities. Section 5 presents the KASKADA platform as a universal solution for data stream processing, section 6 describes immerse 3D visualization, and finally in the last section, some conclusions and future works are provided.

## RADIO-TELESCOPE DATA STREAM

One of the most popular techniques in the field of radio astronomy is called astronomical interferometry. It involves using multiple connected radio-telescopes and superimposing

the signal waves to significantly increase resolution of the observations. The maximum resolution is determined by a baseline, i.e. the distance between the furthest radio-telescopes. The most known examples of using this method are the Very Large Array (VLA) observatory in USA with a maximum baseline of 36 km and Multi-Element Radio Linked Interferometer Network (MERLIN) in England with the longest baseline of 217 km.

It is possible to use interferometry with even more distant radio-telescopes, for example located in various countries across Europe. This method is known as Very Long Baseline Interferometry (VLBI) and has been used since the 1960s [9]. Until recently, sending data over such distances was possible only by physically transporting magnetic tapes or hard drive arrays to a central location. However, with modern computer networks based on fiber-optic cables it is possible to transfer large amounts of data in real-time, significantly speeding up and simplifying the observation process.

In order to transmit data from a radio-telescope through the Ethernet, it must be first converted from an analog to a digital signal. The Digital Baseband Converter (DBBC) performs this transformation with sampling frequency up to 1 GHz. Digital samples are then transmitted to a VLBI data system such as Mark 5. The data system bundles samples into data frames and aggregates them into streams and threads which can be sent over the Ethernet using TCP or UDT protocol. This sequence is shown in Figure 1.

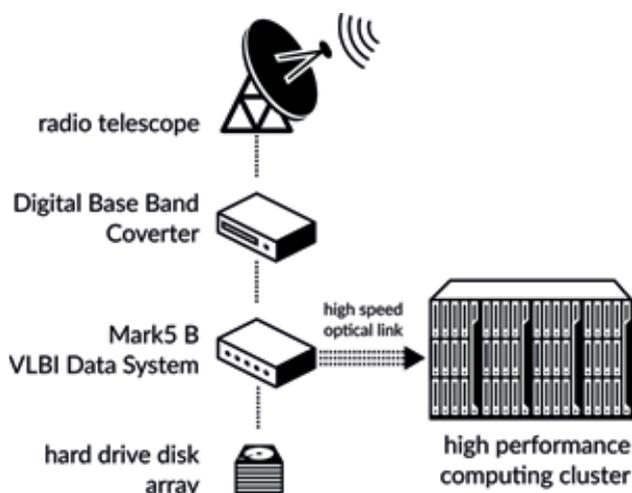


Fig. 1. Data stream flow from a radio-telescope to a computer cluster

Using a supercomputer for data processing is beneficial in many ways. First, the data stream can be analyzed in real-time and the output is almost immediately visible for the observer. Besides the obvious advantage of obtaining results of the observation faster, it allows the researchers to, for example, recalibrate the radio-telescope if some interesting event is detected. Additionally, the supercomputer could be placed at a remote location, many kilometers from the observation site, which results in reduction of radio noise produced by CPUs and other computer equipment.

A typical single e-VLBI data stream contains multiple threads and has a total bandwidth ranging from 1 Gbit/s (Mark 5B) to 16 Gbit/s (Mark 6). For example, the RT-30 radio-telescope, which is used by Torun Centre for Astronomy at Nicolaus Copernicus University, produces a data stream with 2 polarizations, 4 frequency bands, each with 2 sub-bands, sampled at 32 MHz and using 2 bit digitization. As mentioned before, the interferometry technique requires data from multiple streams to achieve high resolution of the observation, which results in even higher data throughput.

Perhaps one of the greatest challenges would be processing and analyzing the data acquired by RT-90+ radio-telescope, which will be built in Tuchola Forest in Poland. A radio-telescope with a diameter of over 90m and state-of-the-art equipment could produce even several Tbit/s of observation data samples.

## SATELLITE DATA STREAM PROCESSING

Satellite images are commonly used in Geographical Information Systems (GIS) to support decision-making in urban development and infrastructure protection. In general, it is realized by integrating numerical models of hazard scenario simulations with existing information from geodatabases and earth observation data streams from satellites.

SafeCity GIS [6] is an innovative system that allows monitoring threats and analyzing risks for the municipal areas developed by the Geoinformatics Department of Gdansk University of Technology. The system can integrate data from multiple sources, including data from a 1.5 m-wide HRPT/MetOp-A/B satellite ground station and information from Service Support Environment (SSE), and then perform various hazard simulations like blast and chemical attacks, floods or oil-spill.

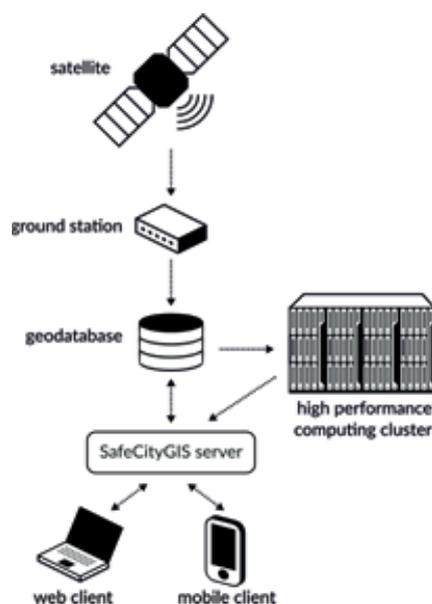


Fig. 2. Architecture of a GIS system integrated with computing cluster

Such simulations often require significant computing power, especially when risks must be analyzed in real-time. For example, oil-spill simulation requires parallel modeling of sea currents and winds, analyzing diffusion of oil spill droplets, processing vaporization and dispersion, and forecasting the size and shape of oil-spill for several hours ahead. The GIS system that handles these types of simulations must offload computation to a high performance computing cluster. Figure 2 demonstrates this concept.

SafeCity GIS system includes two client applications: the complete web client and mobile client. Due to the nature of mobile devices and their restrictions, the mobile client does not offer all functionalities that are available in the complete web client: for example launching a hazard scenario simulation is not possible on a smartphone or tablet. These limitations could be bypassed by offloading some tasks from the mobile device to the cloud [8]. However, the Internet connection required for communication with the cloud services may not be always available and developers have to use adaptive application architecture to overcome this problem, i.e. enable or disable features based on the context of a mobile device, or move the execution between the mobile device and the cloud dynamically.

One of the solutions to this problem is the Mobile Offloading Framework (MOFF) [3] developed as a part of the Center of Excellence in Scientific Application Development Infrastructure “NIWA” project. It provides mechanisms for optimizing the execution of mobile applications in run-time, based on various parameters, such as network availability, network speed, CPU power or battery level. The framework is constantly monitoring parameters of the device and execution time of mobile application components to optimize user experience. The process of offloading execution to the cloud is transparent for the user, as well as for the application developer. The concept of integrating the mobile devices with the cloud is presented in Figure 3.

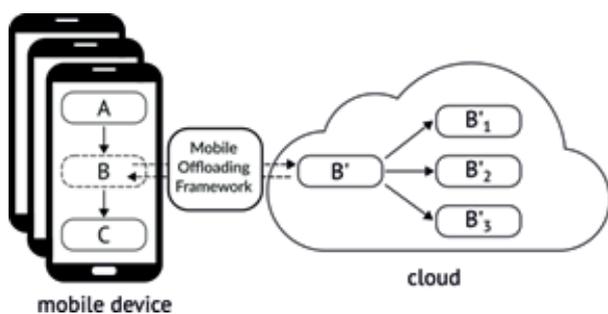


Fig. 3. The concept of Mobile Offloading Framework

### TRYTON SUPERCOMPUTER DESCRIPTION

Tryton supercomputer [10] is the first computer cluster in Poland which exceeded 1 PFLOPS computing power. It is located in the Academic Computer Centre in Gdansk University of Technology and was built to support the newly created Center of Excellence in Scientific Application Development Infrastructure “NIWA”. Its primary purpose

is to provide computing capabilities for the scientific community, including radio-telescope and satellite data streams processing.

Every node of the cluster contains two processors (Intel Xeon Processor E5 v3, 2.3GHz, Haswell), with 12 physical cores (24 logical ones, due to Hyperthreading technology) and 128-256GB RAM memory. Some of the nodes have additional accelerators, such as nVidia Tesla, Intel Xeon Phi or AMD FirePro.

In total the supercomputer consists of 40 racks with 1483 servers (nodes), 2966 processors, 35592 compute cores, 48 accelerators, 202 TB RAM memory. It uses fast FDR 56Gb InfiniBand in fat tree topology and 1Gb Ethernet for networking. The total computing power is 1.37 PFLOPS. The cluster weighs over 20 metric tons.

The computer center also has extensive storage capabilities. Within a few projects, such as PL-Grid, Platon (U4) or Mayday Euro 2012, storage servers can store up to 5PB of data, using modern hard disk and tape archives. All this storage space is easily accessible for the applications and services working on the supercomputer, via InfiniBand or other optic fiber connections.

The supercomputer is placed in one of the most important nodes of the wide area network connecting the Gdansk metropolis and the whole country. The MAN (metropolitan area network) is based on 10 Gigabit Ethernet & ATM technologies. It creates a ring structure connecting academic nodes in Tricity, consisting of 260 km of the fibers in total.

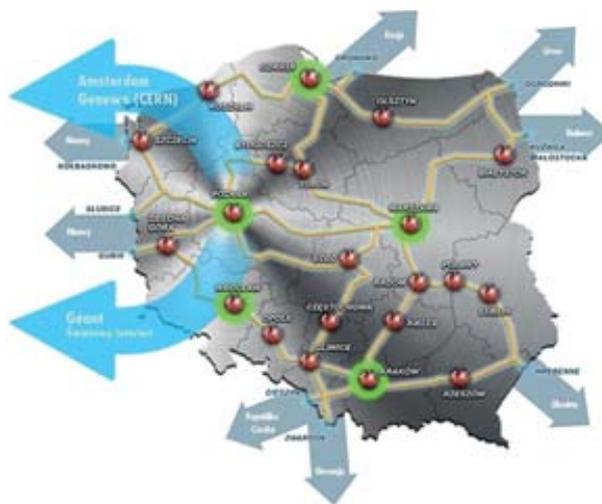


Fig. 4 The PIONIER network structure. Source: <http://www.pionier.net.pl>

The Internet outside connection is facilitated by the PIONIER network (the computer center is one of the shareholders), providing fast access to the current radio-telescope in Torun, and it can be easily extended to the possible location of RT-90+ in Tuchola Forest. The foreign connection is provided by the access to Geant and CERN (through Amsterdam) networks, see Fig. 4.

## KASKADA PLATFORM – FRAMEWORK FOR DATA STREAM PROCESSING

The KASKADA platform is a solution dedicated to processing of continuous streams of data. Initially, it was designed for video and audio processing, however with time it was extended for other data types, like endoscopy examination results or text documents. Moreover, the platform provides the support for high numbers of processed streams and is capable of doing so satisfying (soft) real-time constraints.

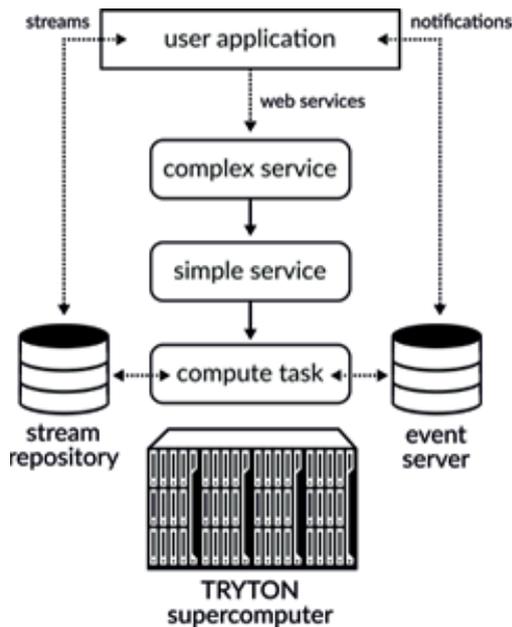


Fig. 5. KASKADA platform architecture

Fig. 5 presents the layered architecture of the KASKADA platform. It is based on the computer cluster equipped with multiple compute nodes connected by a fast network (see section 3 for more details). The second layer uses compute tasks to perform processing dependent on the implemented stream processing algorithm. The next layer provides the support for web services, the simple and the complex ones, respectively. Finally the top layer: user applications enable direct interactions using the proper GUI mechanisms.

The additional components: stream repository and event server are used to support the data stream online computations. The repository uses the mass storage for archiving and managing the incoming streams to facilitate the compute and to provide the data to the user through the user application. The event server provides the means to store and transfer the messages about the events detected during the online processing, and deliver them to the user with the various protocols. See [2] for more details about the architecture and task/service management in general, and [4],[7] specifically for the KASKADA platform.

It is important to emphasize the support for service scenarios implemented in the complex service layer. They enable usage of the pipeline processing including parallel execution of the underlying algorithms. This approach is

crucial for processing of massive data streams because they must be distributed between multiple computing nodes to satisfy real-time constraints. When the user chooses to start some data processing scenario, the application sends a request to the KASKADA platform. The corresponding complex service is selected from the repository, required simple services are assembled, and one or more compute tasks are executed for every simple service. These tasks communicate directly with the stream repository to send or receive data streams and, with the event server, to send notifications to the user application.

Fig. 6 presents an example of such a scenario described in MSP-ML (Multimedia Stream Processing Modeling Language) [1]. The first service (Decoder) is responsible for decoding the data, the second one (SpectralDensity) is used for converting the data from a time to a frequency domain, the two others analyse the pulsar timing (PulsarTiming) and produce the data used for plotting the spectrum function and time series data for the user (GraphPlot). The pulsar timing is only one example of a service scenario that could be used for analyzing data streams from a radio-telescope. Other scenarios that could be easily executed on the KASKADA platform are pulsar discovery, polarimetry or detecting unknown astronomical objects.

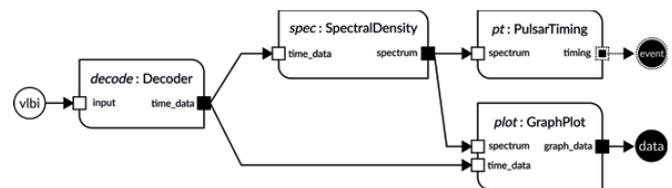


Fig. 6. An example of a service scenario in the KASKADA platform

On the other hand, the KASKADA platform can be also used for non-stream related processing, including such services like simulations, numerical computations or implementation of other scientific algorithms. The platform supports the master-slave processing model (see Fig. 7), where the problem can be split into parts in the master task, and then sent to the slaves; who can be placed on different cluster nodes, increasing the speedup of computations.

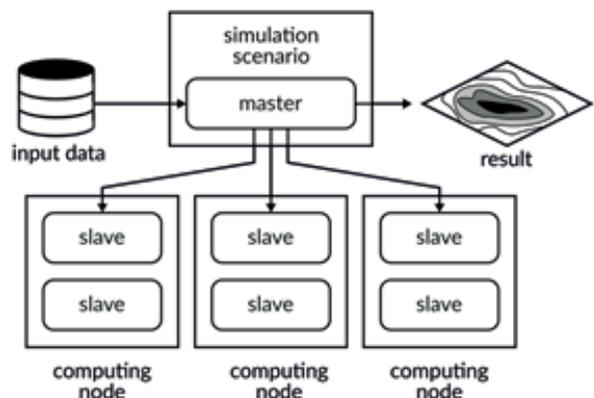


Fig. 7. Master-slave processing model

## IMMERSE 3D VISUALIZATION LABORATORY

The KASKADA platform is also connected with the Immerse 3D Visualization Laboratory [5] that could be used to visualize the results of data processing. The aim of I3DVL is to ensure the highest possible degree of immersion together with the least amount of equipment worn by the user. To achieve this, the laboratory is equipped with a Spherical Walk Simulator placed in a cubic CAVE (Cave Automatic Virtual Environment). Visualization of the laboratory is shown in Figure 8.



Fig. 8. Visualization of the rotary sphere inside the CAVE. Source: [5]

The Spherical Walk Simulator is a transparent sphere rotating on rollers. Therefore, the user can move freely in any direction without actually changing location. Transparent panels allow the user to see the CAVE screen-walls through the sphere. The image is rear-projected on all six faces of the cube structure using stereoscopic projectors. Every wall displays an image with a resolution of 1920x1920 pixels and stereoscopy is achieved either, using an active shutter, or spectrum channels separation glasses.

The I3DVL uses 12 projectors which are driven by a visualization cluster consisting of 12 computing nodes with Nvidia GPUs, connected in an all-to-all communication model. The visualization cluster is controlled by two management nodes, which are responsible for analyzing the context of the simulation, i.e. user position, speed and direction of movement. Architecture of the I3DVL is presented in Figure 9.

This computation model is suitable for visualization of authentic scenarios modeled by maritime applications, including oil-spill simulation, algae bloom monitoring or physical simulation of a ship in rough seas. The results of data processing from the radio-telescope and satellites could be easily transferred from the KASKADA platform to I3DVL. 3D models of planets, stars, galaxies and other astronomical objects could be generated from the observation data and used to better study and analyze results of data processing. The concept of this solution is presented in Figure 10.

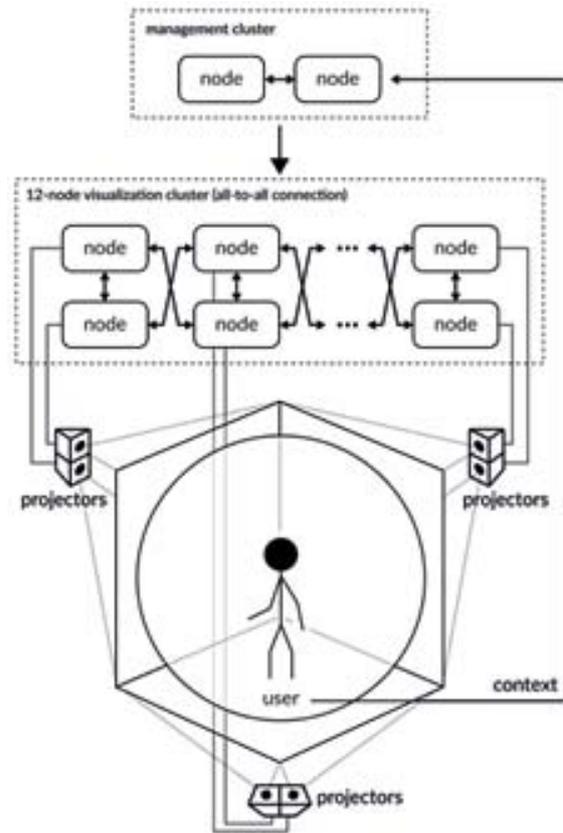


Fig. 9. Model of the I3DVL

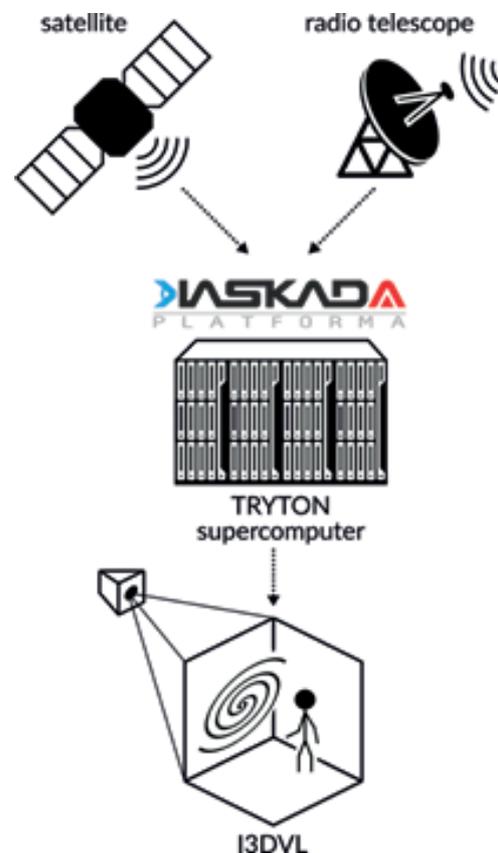


Fig. 10. Immerse 3D Visualization Laboratory

## CONCLUSION

The paper presented the potential of the ACC in GUT for radio-telescope and satellite data processing. Three different computational solutions are suggested (The KASKADA pipeline, master/slave, and I3DVL). The hardware capabilities of the Tryton supercomputer (the highest computational power in Poland), as well as the software provided in the center (KASKADA platform) enables dependable and fast adaptation to the current computational requirements following the different radio-telescope and satellite analysis scenarios, and are ready for the challenges of the new, planned, equipment (e.g. RT-90+).

Although the platform is ready to use for any data stream processing applications devoted to sea and space analysis, we plan to introduce additional features especially designed for the radio-telescope data, such as dedicated input data archiving, or output stream generation. Also, additional effort is required to provide the proper GUI interface and frontend application, first of all for CAVE presentations.

## ACKNOWLEDGEMENT

This work was carried out as a part of the Center of Excellence in Scientific Application Development Infrastructure "NIWA" project, Operational Program Innovative Economy 2007–2013, Priority 2 "Infrastructure area R&D".

## REFERENCES

1. Bobkowska, A., Nykiel, M., Proficz, J., A Model-Driven Solution for Development of Multimedia Stream Processing Applications, Proceedings of the 13th International Conference on Software Engineering, Parallel and Distributed Systems, 2014
2. Hennessy, J. L., Patterson, D. A., Computer architecture: a quantitative approach, Elsevier, 2012
3. Krawczyk, H., Nykiel, M., Proficz, J., Mobile Offloading Framework: Solution for Optimization of Mobile Applications using Cloud Computing, Communications in Computer and Information Science, 2015
4. Krawczyk, H., Proficz, J., Real-time multimedia stream data processing in a supercomputer environment, Interactive Multimedia, InTech, 2012
5. Lebiedź, J., Mazikowski, A., Innovative Solutions for Immersive 3D Visualization Laboratory, WSCG2014 Conference on Computer Graphics, Visualization and Computer Vision, 2014
6. Moszyński, M., Kulawiak, M., Chybicki, A., Bruniecki, K., Bieliński, T., Łubniewski, Z., Stepnowski, A., Innovative Web-based Geographic Information System for Municipal Areas and Coastal Zone Security and Threat Monitoring Using EO Satellite Data, Marine Geodesy, 2014
7. Proficz, J., Krawczyk, H., Task allocation and scalability evaluation for real-time multimedia processing in a cluster environment, proceeding of 6th International Conference on Pattern Recognition and Machine Intelligence, Warsaw, Poland, 2015 (accepted)
8. Rittinghouse, J.W., Ransome, J.F., Cloud Computing: Implementation, Management, and Security, CRC Press, Taylor and Francis Group, 2010
9. Thompson, A., Moran, J., Swenson Jr, G., Interferometry and synthesis in radio astronomy, John Wiley & Sons, ch. 9, pp. 304-325, 2008
10. Tryton Supercomputer, Academic Computer Centre in Gdansk, <http://task.gda.pl/kdm/sprzet/tryton/>

## CONTACT WITH AUTHOR

Henryk Krawczyk  
Michał Nykiel

Faculty of Electronics, Telecommunications  
and Informatics,  
Gdańsk University of Technology,  
G. Narutowicza 11/12 Str.  
80-233 Gdansk  
POLAND

hkrawk@eti.pg.gda.pl  
mnykiel@task.gda.pl

Jerzy Proficz

Academic Computer Centre in Gdansk  
Gdańsk University of Technology  
ul. G. Narutowicza 11/12 Str.  
80-233 Gdansk  
POLAND

jerp@task.gda.pl