CONTENTS

5 Krzysztof Czaplewski, Sambor Guze, Sławomir Świerczyński
THE IMPACT OF RADAR DISTANCE MEASUREMENT ACCURACY ON THE ACCURACY OF POSITION FIXING IN VTS SYSTEMS

14 Shengke Ni, Zhengjiang Liu, Yao Cai, Xin Wang
MODELLING OF SHIP’S TRAJECTORY PLANNING IN COLLISION SITUATIONS BY HYBRID GENETIC ALGORITHM

26 Yong Liu, Renxiang Bu, Xiaori Gao
SHIP TRAJECTORY TRACKING CONTROL SYSTEM DESIGN BASED ON SLIDING MODE CONTROL ALGORITHM

35 Maciej Janecki, Artur Nowicki, Alicja Kańska, Maria Golenko, Lidia Dzierzbicka-Głowacka
NUMERICAL SIMULATIONS OF SEA ICE CONDITIONS IN THE BALTIC SEA FOR 2010–2016 WINTERS USING THE 3D CEMBS MODEL

44 Lucjan Gucma, Julia Raczkowska
AN ANALYSIS OF BASIC PARAMETERS OF RO-PAX FERRIES IN THE BALTIC SEA AS GUIDELINES FOR ITS PRELIMINARY DESIGN

54 Bojan Beškovnik, Marina Zanne
BUSINESS PROCESS RE-ENGINEERING OF A MARITIME CAR TERMINAL: CHANGING FROM INBOUND TO OUTBOUND INTERMODAL NODE IN FINISHED VEHICLE LOGISTICS (FVL)

62 Mateusz Sondej, Chandana Ratnayake, Michał Wójcik
ECONOMICALLY AND SAFE METHOD OF GRANULAR MATERIAL STORAGE IN SILOS IN OFFSHORE PORT TERMINALS

69 Piotr Srokosz, Ireneusz Dyka, Marcin Bujko
DETERMINATION OF SHEAR MODULUS OF SOIL IN THE RC / TS APPARATUS FOR DESIGNING OFFSHORE WIND POWER PLANT FOUNDATIONS

84 Michał Wodtke, Artur Olszewski, Artur Wójcikowski
FEM CALCULATIONS IN ANALYSIS OF STEEL SUBSEA WATER INJECTION FLOWLINES DESIGNING PROCESS

94 N R Ammar, I S Sadek
THERMO-ECONOMIC ANALYSIS AND ENVIRONMENTAL ASPECTS OF ABSORPTION REFRIGERATION UNIT OPERATION ONBOARD MARINE VEHICLES: RO- PAX VESSEL CASE STUDY
Mariusz Bogdan, Marcin Derlatka, Józef Błachnio
CONCEPT OF COMPUTER-AIDED ASSESSMENT OF THE TECHNICAL CONDITION OF OPERATED GAS TURBINE VANES

Adam Góralczyk, Adam Adamkowski
MODEL OF A DUCTED AXIAL-FLOW HYDROKINETIC TURBINE – RESULTS OF EXPERIMENTAL AND NUMERICAL EXAMINATION

Michał Stosiak, Maciej Zawiślak, Bohdan Nishta
STUDIES OF RESISTANCES OF NATURAL LIQUID FLOW IN HELICAL AND CURVED PIPES

Jacek Tomków, Jerzy Łabanowski, Dariusz Fydrych, Grzegorz Rogalski
COLD CRACKING OF S460N STEEL WELDED IN WATER ENVIRONMENT

Eligiusz Mieloszyk, Mariusz Wyroślak
INFLUENCE OF GEOTECHNICAL CONDITIONS ON DAMAGE STATES OF GDANSK BAY COAST CLIFF FORMATIONS
Editorial

POLISH MARITIME RESEARCH is the scientific journal with a worldwide circulation. This journal is published quarterly (four times a year) by Gdansk University of Technology (GUT). On September, 1994, the first issue of POLISH MARITIME RESEARCH was published. The main objective of this journal is to present original research, innovative scientific ideas, and significant findings and application in the field of:

Naval Architecture, Ocean Engineering and Underwater Technology,

The scope of the journal covers selected issues related to all phases of product lifecycle and corresponding technologies for offshore floating and fixed structures and their components.

All researchers are invited to submit their original papers for peer review and publications related to methods of the design; production and manufacturing; maintenance and operational processes of such technical items as:

- all types of vessels and their equipment,
- fixed and floating offshore units and their components,
- autonomous underwater vehicle (AUV) and remotely operated vehicle (ROV).

We welcome submissions from these fields in the following technical topics:

- ship hydrodynamics: buoyancy and stability; ship resistance and propulsion, etc.,
- structural integrity of ship and offshore unit structures: materials; welding; fatigue and fracture, etc.,
- marine equipment: ship and offshore unit power plants: overboarding equipment; etc.

Scientific Board

Chairman : Prof. JERZY GIRTLER - Gdansk University of Technology, Poland
Vice-chairman : Prof. MIROSŁAW L. WYSZYŃSKI - University of Birmingham, United Kingdom

Dr POUL ANDERSEN
Technical University of Denmark
Denmark

Prof. STANISLAW GUCMA
Maritime University of Szczecin
Poland

Prof. EUGEN NEGRUS
University of Bucharest
Romania

Dr MEHMET ATLAR
University of Newcastle
United Kingdom

Prof. ANTONI ISKRA
Poznań University of Technology
Poland

Prof. YASUHIKO OHTA
Nagoya Institute of Technology
Japan

Dr POUL ANDERSEN
Technical University of Denmark
Denmark

Prof. JAN KICIŃSKI
Institute of Fluid-Flow Machinery of PAsci
Poland

Dr YOSHI SATO
National Traffic Safety and Environment Laboratory
Japan

Prof. SERGEY BARSUKOV
Army Institute of Odessa
Ukraine

Prof. ZBIIGNIEW KORCZEWSKI
Gdańsk University of Technology
Poland

Prof. KLAUS SCHIER
University of Applied Sciences
Germany

Prof. MUSTAfA BAYHAN
Süleyman Demirel University
Turkey

Prof. JANUSZ KOZAK
Gdańsk University of Technology
Poland

Prof. FREDERICK STERN
University of Iowa
USA

Prof. VINCENZO CRUPI
University of Messina,
Italy

Prof. ILCEV DIMOV STOJCE
Durban University of Technology
Africa

Prof. MAREK DZIDA
Gdańsk University of Technology
Poland

Prof. NICOS LADOMMATOS
University College London
United Kingdom

Prof. JÓZEF SZALA
University of Technology and Science in Bydgoszcz
Poland

Prof. ODD M. FALTINSEN
Norwegian University of Science and Technology
Norway

Prof. JÓZEF LISOWSKI
Gdynia Maritime University
Poland

Prof. WITALIJ SZCZAGIN
State Technical University of Kaliningrad
Russia

Prof. PATRICK V. FARRELL
University of Wisconsin Madison, WI
USA

Prof. JERZY MATUSIAK
Helsinki University of Technology
Finland

Prof. BORIS TIKHOMIROV
State Marine University of St. Petersburg
Russia

Prof. HASSAN GHASSEMI
Amirkabir University of Technology
Iran

Prof. JERZY MERKISZ
Poznań University of Technology
Poland

Prof. DRACOS VASSALOS
University of Glasgow and Strathclyde
United Kingdom
THE IMPACT OF RADAR DISTANCE MEASUREMENT ACCURACY ON THE ACCURACY OF POSITION FIXING IN VTS SYSTEMS

Krzysztof Czaplewski
Sambor Guze
Sławomir Świerczyński
1 Gdynia Maritime University, Poland
2 Polish Naval Academy, Gdynia, Poland

ABSTRACT

The main source of information on the situation across the sea basins used by operators of shipping monitoring systems is a network of coastal radar stations. Presently, it is possible to gather navigational information from many individual radar stations simultaneously, which may be used for improving the accuracy of vessel position fixing. However, without making other estimates, we obtain an inconsistent image comprising multiple echoes of the same ship, and as such it is impossible to say which echo presents the vessel on the move. Another problem is the method of performing radar observations, which significantly affects the accuracy of position fixing. The estimated radar distance is encumbered with a gross error in the case of large vessels, as the position of a large vessel is not the same as the position of the edge of the radar echo to which the estimation is made. In this paper, the authors present a method to adjust the measured radar distance. The proposed method may be automated easily, which would significantly enhance VTS positioning processes.

Keywords: radar, VTS, navigation, navigation automation, mathematical methods

INTRODUCTION

The main source of information on the situation across the sea basins supporting the work of operators of shipping monitoring systems is a network of coastal radar stations. The coastal radar stations located along the shores of basins under supervision of VTS systems may convey navigational information from one or more radars. VTS systems also use the information from the Automatic Identification System, which uses the information from satellite systems [Urbaniński, Morgaś, Specht 2008; Czaplewski, Goward 2016; Dziewicki Specht 2009]. However, this information area is not considered in this article.

Navigational information may be simultaneously obtained from many individual radar stations [Świerczyński, Czaplewski 2012, 2013], and used to improve the accuracy of estimating vessel positions. However, the information obtained simultaneously from all coastal stations produces an inconsistent image of the ship generated by multiple echoes, in which case it is difficult to determine which echo represents the vessel on the move [Świerczyński 2017]. To eliminate this problem and correct the fixing, modern observation adjustment methods, known in geodesy and presented in [Czaplewski, Świerczyński 2015], among other publications, may be used. Greater ship positioning accuracy translates into improved navigational safety and the possibility of giving an early warning, which is especially important in the case of enclosed waters where the risk of collision is relatively high [Guze et al., 2016]. Another research problem is the adequate use of observations (bearings, range) obtained from coastal radar stations. Good maritime practice requires that in the case of a radar bearing, a watchman performs the bearing to the centre of the radar echo, whilst in the case of radar distance measurements, the observation should be made to the nearest edge of the echo, and not to the echo centre. In the case of vessels of small dimensions (length, width), the measurement error is not considerable, but for large vessels
(such as MAERSK container ships) the observations may be encumbered by gross errors, depending on the position of the vessel in relation to the radar station. The authors of this paper present the results of their studies and propose to automate the process of adjusting distances measured by coastal radar stations. The studies were performed in the basin covered by the Gulf of Gdansk VTS system.

**RADAR SUB-SYSTEM IN THE GULF OF GDANSK VTS SYSTEM**

The radar sub-system provides for early detection and accurate tracking of targets to prevent vessel collisions and environmental threats. The National Maritime Safety System (KSBM) is a monitoring and control system committed to ensuring maritime navigation safety and protection of maritime environment. As part of KSBM’s development (KSBM-I – 2014), the facilities managed by Maritime Offices in Gdynia, Słupsk and Szczecin, and by the Maritime Search and Rescue Service were modernized. The VTS system in Gdynia is contained within the organisational structure of KSBM, providing services for the territory of jurisdiction of Gdynia Maritime Office. The Gulf of Gdansk VTS system was upgraded as part of stage 1 of KSBM’s development. Five radar stations were modernised and other five were put into operation. Presently, there are ten operational radar stations. The location of nine coastal radar stations is presented in Fig. 1. Additionally, one radar station was built on the PetroBaltic platform [UMG, 2016c].

**RADAR DISTANCE DETERMINATION MODEL ACCOUNTING FOR THE SHAPE OF VESSEL**

In radar navigation, the reciprocal position of ships is characterised by the angle between the true course of the observed vessel and the bearing taken from the observed object on the own vessel [Chrzanowski, et al. 2010]. This angle is called aspect. For the description of a radar cross-section, we use the term “aspect” defined as the angle at which the object is inclined toward the incoming radar wave [Chrzanowski, et al., 2010; Wei, et al. 2010]. In this article, the angle between the course of the vessel observed in the VTS system and the bearing taken on the ship from the coastal radar station (Fig. 2) will be called the aspect ($\alpha_A$).

The coordinates of radar station antennas are presented in Table 1.

While retaining the same functional scope, the modernisation of the Gulf of Gdansk VTS system has brought about increased efficacy. Software and hardware was upgraded, operators’ stands were modernised, new sensors were installed, and data transmission was provided to a central application. The integration of all components is implemented through the Maritime Safety Information Exchange System (SWIBŻ) and high-speed communication networks of KSBM centres in Gdynia, Ustka and Szczecin.

![Fig. 1. Location of coastal stations operating in the system of Gulf of Gdansk VTS](image1)

![Fig. 2. Graphic presentation of the aspect](image2)

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Location</th>
<th>Geographic coordinates</th>
<th>Cartesian coordinates</th>
</tr>
</thead>
<tbody>
<tr>
<td>Władysławowo</td>
<td>Władysławowo Harbour</td>
<td>54°47.809’ N 018°25.281’ E</td>
<td>6075231.55 334234.35</td>
</tr>
<tr>
<td>Hel_L</td>
<td>Hel Lighthouse</td>
<td>54°36.003’ N 018°48.771’ E</td>
<td>6052486.06 358713.75</td>
</tr>
<tr>
<td>HEL_SB</td>
<td>Hel, South Breakwater</td>
<td>54°35.985’ N 018°48.058’ E</td>
<td>6052476.63 357945.55</td>
</tr>
<tr>
<td>Gdynia_WQ</td>
<td>Gdynia, Wenda Quay</td>
<td>54°31.739’ N 018°33.576’ E</td>
<td>6045191.44 342683.22</td>
</tr>
<tr>
<td>Gdynia_HMO</td>
<td>Gdynia, Harbour Master’s Office</td>
<td>54°32.021’ N 018°41.841’ E</td>
<td>6045669.81 341309.47</td>
</tr>
<tr>
<td>Gdańsk_HMO</td>
<td>Harbour Master’s Office, Gdańsk North Harbour</td>
<td>54°23.987’ N 018°41.778’ E</td>
<td>6030447.63 350457.03</td>
</tr>
<tr>
<td>Gdańsk_NH</td>
<td>Gdańsk, New Harbour</td>
<td>54°24.405’ N 018°39.658’ E</td>
<td>6031298.79 348189.74</td>
</tr>
<tr>
<td>GZ_RT</td>
<td>Górki Zachodnie, Radar Tower</td>
<td>54°22.229’ N 018°46.734’ E</td>
<td>6027017.31 355714.79</td>
</tr>
<tr>
<td>KM_L</td>
<td>Krynica Morska, Lighthouse</td>
<td>54°23.123’ N 019°27.046’ E</td>
<td>6027506.85 399392.32</td>
</tr>
</tbody>
</table>
Among all radar observations, bearing and distance measurements are the most common. With regard to radar bearing, good maritime practice recommends that the observation be performed toward the centre of the radar echo for better quality of its final determinations. On the other hand, the distance measurement is to be made to the “closest” boundary of echo created by RCS. For small sailing vessels, such a measurement technique is not encumbered by a big error that could grossly affect final determinations. The situation is dramatically different when a radar band “delivers” echo for large and very large vessels. The site of measurement is significantly different from the centre of the radar echo, which grossly affects the final determination of the coordinates of a ship at sea [Bole et al., 2005].

The factor that heavily affects the accuracy of position fixing is the vessel’s course in relation to the measuring radar. Coastal radar stations are located along the costal line around the basin. Each radar can “see” the sailing vessel at a different angle, at a different bearing. This angle affects the accuracy of fixing the distance or bearing by the radar [Curry, 2004].

The accuracy of position fixing is also affected by the shape of a vessel, which can be different depending on its type and designed use. However, its radar image is usually similar to the shape of an ellipse, and the position of a ship is assumed to be at its centre. In the analyses presented further in this paper, we will assume that the coordinates of the centre of an ellipse are the coordinates of the vessel at sea; the sailing vessel will be represented by a flat figure being a combination of an ellipse and a rectangle (Fig. 2). Such a figure will provide a more precise representation of the actual shape of the hull for the purpose of the model proposed in this article.

Let us assume that the aspect ($\alpha_A$) takes values from the range $\alpha_A = (000°, 180°)$ in the system of horizon half-division. The $\alpha_A$ angle depends on the true bearing taken from the position of the radar antenna on the vessel and the course of the vessel, which may be illustrated by the following formula:

\[
\begin{align*}
\alpha_A &= NR^R - KR + 180° \quad \text{for} \quad (NR^R - KR) < 000° \\
\alpha_A &= 180° \quad \text{for} \quad (NR^R - KR) = 000° \\
\alpha_A &= 000° \quad \text{for} \quad (NR^R - KR) = -180° \\
\alpha_A &= NR^R - KR - 180° \quad \text{for} \quad (NR^R - KR) > 000°
\end{align*}
\] (1)

where:

- $NR^R$ is the true bearing taken from the radar on the ship,
- $KR$ is the true course of the ship.

Assuming that the model of the bow of the ship is a half ellipse and the value of the aspect is within the range of $\alpha_A \in (000°, 090°)$, then the measurement structure of interest will take the shape presented in Fig. 3.

Moreover, the following parameters were introduced: $\alpha$ is a half of the vessel's length (semi-major axis of the ellipse), and $b$ is half of the vessel's width (semi-minor axis of the ellipse). In Fig. 3, the measured radar distance ($d_R$) was marked together with the distance ($d_p$) by which the $d_R$ distance should be adjusted in accordance with the assumption that the centre of the ellipse marks the position of the sailing vessel. The VTS operator is interested in the actual position of the vessel, therefore, the navigational observation obtained from the radar distance measurement should be specified as follows:

\[
d_C = d_R + d_p
\] (2)

where:

$\alpha_A$ is the actual distance between the radar station and the ship position.

To determine the distance $d_p$ by which the radar distance ($d_R$) should be adjusted, the ellipse equation needs to be used. Using a local coordinate system with the centre at the ellipse centre, the equation of the ellipse will take the following form [Yates, 1974]:

\[
\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1 \quad \rightarrow \quad b^2 \cdot x^2 + a^2 \cdot y^2 = a^2 \cdot b^2
\] (3)

Assuming that the radar beam bounces off the hull at point $P'$, then its polar coordinates $(x_p, y_p)$ may be written down as:

\[
\begin{align*}
x_p &= d_p \cdot \cos \alpha_A \\
y_p &= d_p \cdot \sin \alpha_A
\end{align*}
\] (4)

Upon substituting the formula (4) for (3), the following is obtained:

\[
b^2 \cdot d_p^2 \cdot \cos^2 \alpha_A + a^2 \cdot d_p^2 \cdot \sin^2 \alpha_A = a^2 \cdot b^2
\] (5)

And after transformations:

\[
d_p^2 = \frac{a^2 \cdot b^2}{b^2 \cdot \cos^2 \alpha_A + a^2 \cdot \sin^2 \alpha_A}
\] (6)

Using the geometrical formula $\sin^2 \alpha_A + \cos^2 \alpha = 1$, a formula is finally obtained that can be used to determine the adjustment to the measured radar distance:

\[
d_p = \frac{a \cdot b}{\sqrt{b^2 \cdot \cos^2 \alpha_A + a^2 \cdot \sin^2 \alpha_A}}
\] (7)

The formula (7) is valid for the angle’s value $\alpha_A \in (000°, 090°)$.  

**Fig. 3. Measurement structure for the adjustment of radar distance $d_R$ measured up to the bow.**
Assuming that the model for the stern is a rectangle with sides equal to a half of the ship’s length \(a\) and width \(b\), the measurement structure takes the form shown in Fig. 4.

Then, two cases have to be considered: when the radar beam falls on the ship’s broadside (Fig. 4) or on the stern part (Fig. 5).

In the first case (Fig. 4), let the \(\beta\) angle be \(\beta < 180° - \alpha_A\), then, after application of the law of sines, we receive:

\[
\frac{b}{d_p} = \sin(180° - \alpha_A) \rightarrow d_p = \frac{b}{\sin(180° - \alpha_A)} \tag{8}
\]

For the distance measurement performed to the vessel’s stern (Fig. 5), the law of cosines is used. The use of the law of cosines is justified by the determination of the \(\beta_G\) angle value (Fig. 4). The value of this angle can be determined using the formula:

\[
\beta_G = \arctg \frac{a}{b} \tag{9}
\]

Then, it may be assumed that the threshold value of aspect \((\alpha_G)\) above which the law of cosines should be used takes the form:

\[
\alpha_A = \alpha_G = \beta_G + 90° \tag{10}
\]

For this case, the angle \(\beta < 180° - \alpha_A\) (Fig. 5), and the value of the corrected distance is calculated with the use of the law of cosines. Finally:

\[
\frac{a}{d_p} = \cos(180° - \alpha_A) \rightarrow d_p = \frac{a}{\cos(180° - \alpha_A)} \tag{11}
\]

As a result, we obtain a mathematical model of \((d_p)\) distance by which the measured radar distance should be adjusted, as a function of sailing vessel dimensions, determined with the following formulas:

\[
\begin{align*}
\text{for } \alpha_A = 000° & \quad \alpha_A = 180° \\
\text{for } \alpha_A \in (000°, 090°) & \quad d_p = a \\
\text{for } \alpha_A = 090° & \quad d_p = \frac{a \cdot b}{\sqrt{b^2 + (a^2 - b^2) \cdot \sin^2 \alpha_A}} \\
\text{for } \alpha_A \in (090°, \alpha_G) & \quad d_p = b \\
\text{for } \alpha_A \in \alpha_G, 180° & \quad d_p = \frac{a}{\cos(180° - \alpha_A)}
\end{align*}
\tag{12}
\]

Where:
- \(a\) – is half of the vessel’s length
- \(b\) – is half of the vessel’s width

**VALIDATION TEST**

Built in 2007, the Deepsea Container Terminal Gdansk (DTC) has been handling one of the world’s largest container vessels of E Maersk Line class, which regularly calls there every week, since 2011. In their research, the authors have used the actual registrations performed at the Gulf of Gdansk VTS Centre in Gdynia. The test presented below uses the measurement structure that involved four coastal radar stations (Hel_L, Gdynia_HMO, Gdansk_HMO, GZ_RT) and the ship m/s MADISON MAERSK (Fig. 6). In the test the registration of the ship’s passage which took place on 06/18/2017 at 06.20 – 06.35 UTC is used. The container ship sailed to approach
the channel to the Port of Gdansk, at the height of the special mark BY ZS with speeds 12 – 9.2 kn. The vessel was moving through the courses 205° – 227°. Hydro-meteorological conditions prevailing during the registration were as follows:

- visibility: 9 Nm,
- air temperature: 14°C,
- state of the sea: 2,
- wind direction: S.

The coordinates of the coastal radar stations that performed the observations are presented in Table 1. Taking the vessel’s dimensions (\(L = 399\) m, \(B = 59\) m) into account, the formula \(a = L/2 = 199.50\) m, \(b = B/2 = 29.50\) m was applied for further calculations.

Four distances were measured from the coastal radar stations (Table 2) to the vessel positioned at \(\varphi^o = 54^o26.748'N\), \(\lambda^o = 018^o56.154'N\), \((X^o_p = 6035084.50, X^o_p = 366158.25)\). The position was registered in the AIS system.

The test presents the moment when the vessel was on course \(KR = 221^o\).

First, angle \(\alpha_A\) values in the half-division horizon system should be determined. Given the formula (1) and Fig. 7 the conclusion is summarised that:

- the aspect for Hel L radar station is: \(\alpha_A = 155.2^o - 221.0^o + 180^o = 114.2^o\), because \((NR_1^R - KR) < 0\) in the system of half-division horizon \(\alpha_A = 114.2^o\) p.b.
- the aspect for Gdynia_HMO radar station is: \(\alpha_A = 110.8^o - 221.0^o + 180^o = 069.8^o\), because \((NR_2^R - KR) < 0\) in the system of half-division horizon \(\alpha_A = 069.8^o\) p.b.
- the aspect for Gdańsk_HMO radar station is: \(\alpha_A = 072.3^o - 221.0^o + 180^o = 031.3^o\), because \((NR_3^R - KR) < 0\) in the system of half-division horizon \(\alpha_A = 031.3^o\) p.b.
- the aspect for GZ_RT radar station is: \(\alpha_A = 050.7^o - 221.0^o + 180^o = 009.7^o\), because \((NR_4^R - KR) < 0\) in the system of half-division horizon \(\alpha_A = 009.7^o\) p.b.
Then, in accordance with formula (12), adjustments are made to the measured distances for particular values of the aspect:

- $\alpha_A \in (000^\circ, 090^\circ)$

Since aspect $\alpha_A = 009.7^\circ$ p.b. for GZ_RT station, in accordance with (7) the distance adjustment equals as follows:

$$d_p = \frac{a \cdot b}{\sqrt{b^2 + (a^2 - b^2) \cdot \sin^2 \alpha_A}} = \frac{199.50 - 29.50}{\sqrt{(29.50)^2 + \left[(199.50)^2 - (29.50)^2\right] \cdot \sin^2 009.7^\circ}} = 132.41 \text{ [m]}$$

For Gdańsk_HMO station, where the aspect $\alpha_A = 031.3^\circ$ p.b., in accordance with (7) the distance adjustment equals as follows:

$$d_p = \frac{a \cdot b}{\sqrt{b^2 + (a^2 - b^2) \cdot \sin^2 \alpha_A}} = \frac{199.50 - 29.50}{\sqrt{(29.50)^2 + \left[(199.50)^2 - (29.50)^2\right] \cdot \sin^2 031.3^\circ}} = 55.17 \text{ [m]}$$

And for Gdynia_HMO station, where aspect $\alpha_A = 069.8^\circ$ p.b., in accordance with (7) the distance adjustment equals as follows:

$$d_p = \frac{a \cdot b}{\sqrt{b^2 + (a^2 - b^2) \cdot \sin^2 \alpha_A}} = \frac{199.50 - 29.50}{\sqrt{(29.50)^2 + \left[(199.50)^2 - (29.50)^2\right] \cdot \sin^2 069.8^\circ}} = 31.39 \text{ [m]}$$

- $\alpha_A \in (090^\circ, \alpha_G)$

The aspect calculated for Hel_L station is contained within the sector of stern angles; therefore, a threshold value of the aspect angle should be specified in order to determine which formula, (8) or (11), is to be used:

$$\alpha_G = \beta_G + 090^\circ = 081.59^\circ + 090^\circ = 171.59^\circ$$

where:

$$\beta_G = \arctan \frac{a}{b}$$

Then, the value of measured distance adjustment will be calculated with the use of (11).

Since the aspect for Hel_L station is $\alpha_A = 114.2^\circ$ p.b., the value of measured distance adjustment in accordance with (8) is:

$$d_p = \frac{b}{\sin(180^\circ - \alpha_A)} = \frac{29.50}{\sin(180^\circ - 114.2^\circ)} = 32.34 \text{ [m]}$$

The final values of the adjusted radar distances measured by coastal radar stations of the VTS system are presented in Table 4.

**Tab. 4. Measured distances from the coastal radar stations to the vessel**

<table>
<thead>
<tr>
<th>Coastal radar stations</th>
<th>Observations $d_p$ [m]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hel Lighthouse (Hel_L)</td>
<td>$d_1^p = 18885.00$</td>
</tr>
<tr>
<td>Gdynia, Harbour Master's Office (Gdynia_HMO)</td>
<td>$d_2^p = 26975.00$</td>
</tr>
<tr>
<td>Harbour Master's Office, Gdańsk North Harbour (Gdansk_HMO)</td>
<td>$d_3^p = 16310.00$</td>
</tr>
<tr>
<td>Górcz Zachodnie, Radar Tower (GZ_RT)</td>
<td>$d_4^p = 130655.00$</td>
</tr>
</tbody>
</table>

The adjusted distances ($d^C$) are different from the measured values, in extreme cases by 31.39 [m] (for Gdynia_HMO) and by 132.41 [m] (for GZ_RT).

The proposed model for adjusting radar distances measured by VTS system stations provides for a more effective way of using the observations made in the proposed methodology. With the use of the model, the adjustments to the observations may be made more accurately. Eventually, the position of m/s MADISON MAERSK was fixed with the use of adjustment calculus described in [Czaplewski, Świerczyński 2015]:
\[ V = A\hat{d}_x + L \]

\[ C_X = \sigma_0^2 Q_X = \sigma_0^2 P^{-1} \]

\[ \hat{C}_X = \sigma_0^2 \hat{Q}_X = \sigma_0^2 \hat{P}^{-1} \]

\[ \hat{P} = T(\hat{V})P \]

\[ \Psi(\hat{d}_x) = V^T PV = V^T T(\hat{V})P \]

where:

- \( P = T(\hat{V})P \) is the equivalent weight matrix,
- \( \hat{C}_X \) is the equivalent covariance matrix,
- \( \hat{Q}_X \) is the equivalent cofactor matrix,
- \( T(\hat{V}) \) is the diagonal attenuation matrix.

In this paper, the authors have skipped the determination of adjustment calculus (13) as the principles of the surveying methods of robust M-estimation and the Danish damping function are broadly described in literature, see [Wiśniewski 2016; Czaplewski K., Wąż M. 2017] for instance, and were the subject of research described in [Czaplewski, Świerczyński 2015]. It has been assumed that the error was the same for the observations of all distances and amounted to \( m_{d_i} = 10 \) [m] \((i = 1, \ldots, 4)\) [UMG, 2016a, 2016b]. The adjusted distances significantly affected the free term vector \( (L) \):

- before the measured distances were adjusted:

\[
L = \begin{bmatrix}
18927.09 - 18885.00 \\
27009.45 - 26975.00 \\
16371.59 - 16310.00 \\
13196.42 - 13065.00
\end{bmatrix} = \begin{bmatrix} 42.09 \\ 34.45 \\ 61.59 \\ 131.42 \end{bmatrix}_{[m]}
\]

- with the use of the method proposed in this paper:

\[
L = \begin{bmatrix}
18927.09 - 18885.00 - 32.34 \\
27009.45 - 26975.00 - 31.39 \\
16371.59 - 16310.00 - 55.17 \\
13196.42 - 13065.00 - 132.41
\end{bmatrix} = \begin{bmatrix} 9.75 \\ 3.06 \\ 6.42 \\ -0.99 \end{bmatrix}_{[m]}
\]

Finally, this has an influence on the estimated position of the ship:

- before the measured distance was adjusted

\[
\hat{X}_p = X_p + \hat{d}_x = \begin{bmatrix} 6035084.50 \\ 366158.25 \end{bmatrix} + \begin{bmatrix} -29.48 \\ -82.58 \end{bmatrix} = \begin{bmatrix} 6035055.02 \\ 366075.67 \end{bmatrix} = \begin{bmatrix} \hat{X}_p \\ \hat{Y}_p \end{bmatrix}
\]

for which the mean error of the fix was \( m_x = 53.16 \) [m]

- with the use of the measured distance adjustment method:

\[
\hat{X}_p = X_p + \hat{d}_x = \begin{bmatrix} 6035084.50 \\ 366158.25 \end{bmatrix} + \begin{bmatrix} 6.28 \\ -4.80 \end{bmatrix} = \begin{bmatrix} 6035090.78 \\ 366153.45 \end{bmatrix} = \begin{bmatrix} \hat{X}_p \\ \hat{Y}_p \end{bmatrix}
\]

for which the mean error of the position fix was \( m_x = 4.21 \) [m]
SUMMARY

The radar is a fundamental tool used in sea navigation due to its ability to provide visualisation of the surroundings, which is especially indispensable in high-traffic waters monitored by VTS systems. The development of technology has made it possible to obtain information about moving objects from several radar stations simultaneously. This gives an opportunity to improve the safety of navigation in congested waters. However, the use of "unprocessed" radar observations generates ambiguous results. Therefore, there is a need to create analytical solutions that can be applied easily in the process of automation to address and solve this problem of ambiguity.

The studies presented in the article have enabled the creation of an analytic method to improve the measured radar distances. Good maritime practice requires that radar observations be performed up to the closest edge of the echo under observation. In the case of large vessels, the position of the echo's edge is not the same as the position of the vessel, therefore, the measured distances should be adjusted in accordance with the assumptions of the method used to adjust them.

In order to be able to use the proposed method for correcting the measured distances in future sea navigation, further research should be carried out for other ship hull shapes and other configurations of coastal radar stations. Furthermore, an interesting problem is a possibility of correcting radar bearing acquired in VTS systems. Therefore, the authors undertake further research in this area, the effects of which will be the subject of subsequent scientific publications.

Adjustment of navigational observations coupled with the application of modern surveying methods of adjusting observations allows for a more accurate position fix, which is of crucial importance for waters where VTS systems are established. The ways in which surveying methods can be used for observation adjustments were described by the authors in [Czaplewski K., Świerczyński S. 2015; 2015b; Świerczyński S., Czaplewski K. 2012, 2013].

REFERENCES

13. UMG, 2016a. SCANTER 2001i Setting to Work, For VTS & CS Applications (262001-ZW), Gdynia
14. UMG, 2016b. Terma Compliance Matrix, Gdynia
15. UMG, 2016c. Koncepcja zastosowania radarów Terma w projekcie KSBM-I / Concept of Terma radars usage in KSBM-I Project, Gdynia
18. Wiśniewski Z. 2016. *Rachunek wyrównawczy w geodezji (z przykładami)*. Wydawnictwo UWM, Olsztyn


CONTACT WITH THE AUTHOR

Krzysztof Czaplewski
e-mail: krzysztof@czaplewski.pl
Akademia Morska w Gdyni
Morska 81–87, 81–225 Gdynia
Poland

Sambor Guze
e-mail: s.guze@wn.am.gdynia.pl
Akademia Morska w Gdyni
Morska 81–87, 81–225 Gdynia
Poland

Sławomir Świeczynski
e-mail: s.swierczynski@amw.gdynia.pl
Akademia Marynarki Wojennej
ul. inż. J. Śmigowicza 69, 81–127 Gdynia
Poland
MODELLING OF SHIP’S TRAJECTORY PLANNING IN COLLISION SITUATIONS BY HYBRID GENETIC ALGORITHM

Shengke Ni
Zhengjiang Liu
Yao Cai
Xin Wang
Dalian Maritime University, China

ABSTRACT

Ship collision-avoidance trajectory planning aims at searching for a theoretical safe-critical trajectory in accordance with COLREGs and good seamanship. In this paper, a novel optimal trajectory planning based on hybrid genetic algorithm is presented for ship collision avoidance in the open sea. The proposed formulation is established based on the theory of the Multiple Genetic Algorithm (MPGA) and Nonlinear Programming, which not only overcomes the inherent deficiency of the Genetic Algorithm (GA) for premature convergence, but also guarantees the practicality and consistency of the optimal trajectory. Meanwhile, the encounter type as well as the obligation of collision avoidance is determined according to COLREGs, which is then considered as the restricted condition for the operation of population initialization. Finally, this trajectory planning model is evaluated with a set of test cases simulating various traffic scenarios to demonstrate the feasibility and superiority of the optimal trajectory.

Keywords: trajectory planning, Multiple Genetic Algorithm, ship collision avoidance, nonlinear programming, COLREGs

INTRODUCTION

Currently, the navigation of ships is mainly depending on the navigators’ experience and techniques coupled with traditional practice of good seamanship. With the growing overload, safety issues associated with human errors contribute to about 75 percent of maritime accidents and lead to significant human, monetary and environmental loss [12]. Therefore, unmanned surface vehicles have been gaining attention in the past decade. For example, in 2005, e-navigation was an initiative started by IMO to increase the safety of navigation by using modern technology. In 2012, the European Union invested in the MUNIN (Maritime Unmanned Navigation through Intelligence in Networks) project to develop an autonomous dry bulk carrier [3]. In 2015, the Finnish Funding Agency approved the academic research project - Advanced Autonomous Waterborne Applications (AAWA). In 2016, Rolls-Royce published the white paper to announce their plan for the construction of autonomous ship [9].

One of the key points for the autonomous ship formulation is to construct the Decision Support System. And the main challenges are to find the optimum approaches in the aspects of the situational awareness and the trajectory planning. Currently, the technologies, in particular navigational instruments such as Automatic Identification System (AIS), Automatic Radar Plotting Aid (ARPA) and Electronic Chart Display and Information System (ECDIS) etc., have supplied a revolutionary platform for the former, thus the selecting of an appropriate algorithm for determining the safe trajectory becomes an urgent matter. In fact, the development of ship’s trajectory planning will be a gradual and iterative process subject to extensive testing and simulation. It is remarkable that the evasive manoeuvres in the optimal trajectory should
be based on an interpretation of common practice, which performs the process of collision avoidance with fewer course deviations. On topic of ship’s trajectory planning, a number of studies have been performed and reviewed [18, 26]. For example, the existing approaches cover the Sequential Gradient Restoration Algorithm [14], Recursive Algorithm [5], Fuzzy Set Theory [6-7], Knowledge-based System [11], the Interval Programming [2], the Maze Routing Algorithm [21], the Fast Marching Method [13], Cooperative Path Planning Algorithm [29], Distributed Decision Support Formulation [35], the Artificial Potential Field [34], the Fuzzy Logic based method [15-16], Genetic Algorithm [8, 31, 32], Evolutionary Algorithm [17, 22, 23, 24, 25, 28], the Ant Colony Algorithm [10, 30], the Danger Immune Algorithm [33]. Based on the methods, the approaches assume that ship’s trajectory planning can be categorized into two general groups, namely the deterministic and heuristic ones [26]. The deterministic approaches follow a set of rigorously defined steps to converge to the feasible solutions characterized by low computational time, but they might be incapable of solving complex situations. The heuristic approaches possess the ability to deal with complex environment, and it only searches inside a subspace of the search space for an optimal solution that satisfies the design requirements. However, the optimal evasive manoeuvre usually involves multiple course alterations which is not complaint with the common practice. In addition, the consistency of the heuristic approaches’ outputs cannot be guaranteed, which hinders its adoption for application. In recent years, EA and GA have received great concern by scholars in the field of trajectory planning. They are both the subset of evolutionary computation and are mainly different in the evolutionary manners. In fact, EA is a next generation of GA. These techniques, based on the important principle of ‘survival of the fittest’, model some natural phenomena of genetic and phenotypic inheritance and Darwinian strife for survival. It is remarkable that GA is easy to combine with other algorithms to improve its optimal ability. However, the previous work based on GA mainly employed the Standard Genetic Algorithm (SGA) to search for a satisfactory trajectory for ship collision avoidance. Before the optimization, it needs to debug the parameters’ values for different genetic operators by experiment. In the light of the above given comments, this paper modifies the algorithm by combining the Multiple Genetic Algorithm (MPGA) and the Nonlinear Programming methodology. The idea of the MPGA is that several populations with different combinations of parameters’ value for genetic operators evolve simultaneously. Meanwhile, these populations are related to each other by the establishment of Immigrant Operator and Elite-individual Operator, which guarantee the information exchange and the cooperative work among populations. As a result, the MPGA solves the problem of parameters’ setting and is beneficial for improving the ability of local optimization. In order to further enhance the ability of local optimization and ensure the practicality and consistency of the optimal trajectory, the Nonlinear Programming methodology is incorporated into the MPGA at the specific generation. Additionally, the proposed algorithm adopts a real number encoding method which is convenient for handling constraint conditions. All the simulations are based on the assumption that the own ship is the only manoeuvring party, and target ships keep their heading and speed. This article is composed of seven sections. Section 1 briefly introduces the motivation behind the study. Section 2 explains the division of the encounter situation. Section 3 provides an overview of the SGA and MPGA. Section 4 illustrates the local optimal algorithm - Nonlinear Programming. Section 5 describes this trajectory planning model-the hybrid genetic algorithm. Section 6 outlines the test cases used to discuss the feasibility and superiority of this trajectory planning model. This study is then summarized in section 7.

JUDGMENT OF THE ENCOUNTER SITUATION

In fact, only when the encounter type and the obligation of collision avoidance are classified, the calculation of trajectory planning for collision avoidance in marine traffic can be conducted. Each target ship is categorized into a particular encounter type based on its position, heading as well as its relative bearing with respect to the heading of the own ship (OS). The coordinate system XOY is established as shown in Fig. 1. Let’s assume the coordinate, speed and heading of the OS are \((X_O, Y_O)\), \(V_O\) and \(\phi_O\), and the corresponding information of the target ship (TS) are \((X_T, Y_T)\), \(V_T\) and \(\phi_T\), respectively. Therefore, the related parameters are calculated as follows.

The speed components of the OS and TS along the X and Y coordinate axes are determined by Eq. (1).

The relative speed \(V_R\) of the TS with respect to the OS is determined by Eq. (2)-(4).

![Fig. 1. The space-fixed coordinate system XY](image-url)
\[
\begin{align*}
\begin{cases}
V_{OX} = V_o \times \sin(\phi_o) \\
V_{OY} = V_o \times \cos(\phi_o) \
\end{cases} & \quad \begin{cases}
V_{TX} = V_t \times \sin(\phi_t) \\
V_{TY} = V_t \times \cos(\phi_t) \
\end{cases}
\end{align*}
\]

(1)

\[
\begin{align*}
\begin{cases}
V_{RX} = V_{TX} - V_{OX} \\
V_{RY} = V_{TY} - V_{OY} \
\end{cases}
\end{align*}
\]

(2)

\[V_R = \sqrt{V_{RX}^2 + V_{RY}^2}\]

(3)

\[
\phi_h = a \tan \frac{V_{RX}}{V_{RY}} + \Delta \alpha
\]

(4)

\[
\Delta \alpha = \begin{cases}
0^\circ & \text{if } V_{RX} \geq 0, V_{RY} > 0 \\
360^\circ & \text{if } V_{RX} < 0, V_{RY} > 0 \\
180^\circ & \text{others}
\end{cases}
\]

The relative distance \( R_t \) between the OS and the TS is calculated as:

\[R_t = \sqrt{(X_T - X_O)^2 + (Y_T - Y_O)^2}\]

(5)

The true bearing between each other are shown below.

\[\alpha_t = a \tan \frac{X_T - X_O}{Y_T - Y_O} + \Delta \alpha\]

(6)

\[\alpha_o = a \tan \frac{X_O - X_T}{Y_O - Y_T} + \Delta \alpha\]

(7)

The relative bearing of the TS with respect to the OS is displayed below.

\[\theta_t = \alpha_t - \phi_o \pm 360^\circ\]

(8)

The Distance at the Closest Point of Approach between the ships is determined as:

\[\text{DCPA} = R_t \times \cos(\phi_h - \alpha_t - \pi)\]

(9)

As a result, the encounter situation can be deduced by the ship motion parameters, the orientation division of the TS as well as the identification indicators for specific encounter types. For example, first calculate the corresponding parameters in Fig. 1 and determine the region which the TS situates in by the parameter \( \theta_t \) shown in Fig. 2. In fact, the orientation division of the TS is categorized into six sections \( (P_1, P_2, P_3, P_4, P_5, P_6) \) which is based on an interpretation of the COLREGs, navigation habits and good seamanship. The corresponding values of the bearing lines taken clockwise are \([\pi/8, \pi/2, 5\pi/8, 11\pi/8, 3\pi/2, 15\pi/8]\), respectively, which are mainly based on the provision of light’s arc of horizon defined in COLREGs, except that the region of \( P_1 \) is a little larger than the visibility of the masthead light defined in Annex I 9(a) of the COLREGs. This is because the encounter situation is classified from the point of view of coordination collision avoidance as well as the reduction of uncertainties between head-on and other situations [27]. And, the descriptions of every possible encounter type is listed in Tab 1.

![Fig. 2. Orientation division of the TS (the OS is situated at the centre)](image)

Tab. 1. Abbreviations and brief descriptions of the encounter types

<table>
<thead>
<tr>
<th>Abbreviations</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>HO</td>
<td>Head-on encounter when OS and TS give way simultaneously</td>
</tr>
<tr>
<td>OT1</td>
<td>OS is overtaken by TS when TS gives way</td>
</tr>
<tr>
<td>OT2</td>
<td>TS is overtaken by OS when OS gives way</td>
</tr>
<tr>
<td>CR1</td>
<td>Crossing encounter when TS gives way</td>
</tr>
<tr>
<td>CR2</td>
<td>Crossing encounter when OS gives way</td>
</tr>
<tr>
<td>SF</td>
<td>Safe encounter</td>
</tr>
</tbody>
</table>

Then, the heading of the TS, \( \varphi_t \), is roughly divided to judge the possible encounter type. As the target ship can be situated in different regions with respect to the OS, the categorizing parts are different as shown in Fig. 2.

Finally, the encounter type will be determined based on the combination of the mentioned TS’s relative bearing, the TS’s heading, together with the Distance at the Closest Point of Approach (DCPA). The details are listed in Tab 2. As a matter of fact, the collision risk borrows the concept of ship domain by comparing the DCPA with the radius of the ship’s domain, and \( D_s \) is the radius of ship domain which is a statistical result in the open sea [19]. Usually, the applicable distance for a specific encounter is mainly affected by the visibility distance of lights measured by sight. By the application of modern navigational aids and the principles of good seamanship, the applicable distance is larger.
Tab. 2. The identification indicators for specific encounter type (DCPA<Ds)

<table>
<thead>
<tr>
<th>P</th>
<th>HO CR</th>
<th>CR</th>
<th>OT</th>
<th>OT</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>(\theta_e \in [15\pi/8,17\pi/8])</td>
<td>((\varphi_e - \varphi_o) \in [7\pi/8, 9\pi/8])</td>
<td>((\varphi_e - \varphi_o) \in [3\pi/8, 7\pi/8])</td>
<td>((\varphi_e - \varphi_o) \in [9\pi/8, 13\pi/8])</td>
</tr>
<tr>
<td>2</td>
<td>(\theta_e \in [\pi/8, \pi/2])</td>
<td>(\varnothing)</td>
<td>(\varnothing)</td>
<td>(\varnothing)</td>
</tr>
<tr>
<td>3</td>
<td>(\theta_e \in [\pi/2, 5\pi/8])</td>
<td>(\varnothing)</td>
<td>(\varnothing)</td>
<td>(\varnothing)</td>
</tr>
<tr>
<td>4</td>
<td>(\theta_e \in [5\pi/8, 11\pi/8])</td>
<td>(\varnothing)</td>
<td>(\varnothing)</td>
<td>(\varnothing)</td>
</tr>
<tr>
<td>5</td>
<td>(\theta_e \in [11\pi/8, 3\pi/2])</td>
<td>((\varphi_e - \varphi_o) \in [0, \pi/2])</td>
<td>(\varnothing)</td>
<td>(\varnothing)</td>
</tr>
<tr>
<td>6</td>
<td>(\theta_e \in [3\pi/2, 15\pi/8])</td>
<td>((\varphi_e - \varphi_o) \in [0, \pi])</td>
<td>(\alpha_0 &lt; \varphi_e + 5/8\pi)</td>
<td>(\varnothing)</td>
</tr>
</tbody>
</table>

**STRUCTURE OF THE GENETIC ALGORITHM**

Genetic algorithm is a series of search algorithms and optimization technologies in terms of the principles of natural selection inspired by Darwin’s theory of evolution (the survival of the fittest). In GA-based approaches, the trajectories are represented as chromosomes and GA features a group of candidate solutions (chromosomes) on the response surface. Through genetic operations such as natural selection, recombination and mutation etc., the trajectories with better fitness are found.

**STANDARD GENETIC ALGORITHM**

The previous work using GA in the field of trajectory planning is mainly based on the Standard Genetic Algorithm (SGA), which was first presented by Goldberg in 1989 [4]. And the optimization procedure is mainly handled by three kinds of genetic operators (selection, crossover, mutation) within a single population. The flowchart is given in Fig. 3 together with some descriptions shown below. SGA can be defined by a formula with eight parameters:

\[
SGA = (C, P_0, M, E, \phi, p_c, p_m, T)
\]

Where \(C\) is the individual encoding method, \(P_0\) – the initial population, \(M\) – the population size, \(E\) – the fitness evaluation, \(\phi\) – the selection operator, \(p_c\) – the crossover operator, \(p_m\) – the mutation operator, and \(T\) – the termination condition.

**MULTIPLE POPULATION GENETIC ALGORITHM**

One of the important configurations in SGA is to determine the parameters’ values for genetic operators including the crossover rate \(p_c\) and the mutation rate \(p_m\), which both have a great influence on GA’s performance. In fact, the previous study usually adopts experiments to determine the parameters for genetic operators when applying SGA for ship’s trajectory planning. As a result, it is necessary to dynamically adjust the values of genetic operators to ensure its optimal ability for different traffic scenarios.

Based on the above analysis, the MPGA is adopted and its structure is presented in Fig. 4. The MPGA breaks through the simple population searching frame of SGA, and introduces multiple populations with different combinations of parameters’ values to evolve simultaneously. The communication within these populations are closely interwoven by the establishment of an Immigrant Operator, and the main idea is that the worst chromosome in any
population is replaced by the best chromosome in the adjacent population to achieve co-evolution. Besides, by the establishment of an Elite-individual Operator, the individuals with higher fitness will be chosen to enter the elite population so that the superior individuals will not be lost and make it possible to find the global optimal solution. Furthermore, the elite population does not do selection, crossover or mutation operation for ensuring the integrity of the best individuals.

Fig. 4. The flowchart of MPGA

NONLINEAR OPTIMIZATION METHOD

Nonlinear Programming is a mathematical programming method whose objective function is nonlinear or some of its constraints have nonlinear relationship. It is an important branch of operations research which has wide applications in the field of economics, engineering optimization, management science etc. [1]. As Nonlinear Programming has strong compatibility and robustness in the aspects of local optimization, this work integrates the Nonlinear Programming with MPGA to maintain a balance in the ability of global and local optimization. In fact, the software ‘Matlab’ provides the multivariable binding optimization function ‘fmincon’, which attempts to find a constrained minimum of a scalar function of several variables starting at an initial estimate shown below. The basic thought is to search the optimal solution from initial feasible solution, which finds another better one by iteration method based on specific constraints until the optimal one is fond.

\[
\begin{align*}
[x, fval] &= fmincon(f(x), x_0, A, b, Aeq, beq, lb, ub, nonlcon ops) \\
\min f(x) &
\begin{cases}
Ax \leq b \\
Aeq x = beq \\
lb < x < ub \\
c(x) \leq 0 \\
c_{eq}(x) = 0
\end{cases}
\end{align*}
\]

Where \(x\) is the solution returned as a real vector or array, and ‘fval’ is the objective function value at solution. As for the input parameters, \(f(x)\) is the objective function, \(x_0\) is the initial point for \(x\). Linear constraint matrices \(A\) and \(A_{eq}\) and their corresponding vectors \(b\) and \(b_{eq}\), can be sparse or dense. The boundary of \(x\) is \(lb\) to \(ub\). And ‘nonlcon’ contains the nonlinear equality constraints \(c_{eq}(x) = 0\) and the nonlinear inequality constraints \(c(x) \leq 0\), which the objective function attempts to satisfy. Finally, ‘ops’ is the optimization options, there are four different optimal algorithms including ‘Interior-point algorithm’, ‘Active-set algorithm’, ‘Sqp algorithm’ and ‘Turst-region-reflective algorithm’. It is remarkable that ‘Interior-point algorithm’ handles large, sparse problems, as well as small dense problems; ‘Sqp’ one satisfies bounds at all iterations and the algorithm can recover from NaN or Inf results; ‘Active-set’ one can take large steps, which adds speed; Trust-region-reflective one requires to provide a gradient, and allows for only bounds or linear equality constraints, but not both.

MODELLING OF HYBRID GENETIC ALGORITHM IN SHIP’S TRAJECTORY PLANNING

In this paper, a novel hybrid genetic algorithm is proposed for ship’s trajectory planning. First, the encounter type is determined according to Section 2. Based on the result, a heuristic method is constructed for the population initialization. Then, by the establishment of the Immigrant Operator and Elite-individual Operator, several populations with different combinations of the parameters’ value for genetic operators evolve simultaneously (MPGA). Meanwhile, the adoption of Nonlinear Programming helps enhance the ability of GA in local optimization, and guarantee the optimal solutions satisfy the requirement of practicality with fewer course deviations as well as the consistency. This technique is illustrated in Fig. 5.

Fig. 5. The flowchart of hybrid genetic algorithm
THE ENCODING TECHNIQUE

The encoding technique, the length of code and search space have an influential effect on the optimal speed of the algorithm. As a consequence, the paper adopts the real number encoding method on the basis of previous research, which decreases the dimension of solution space and sets the stage for handling constraint conditions in the Nonlinear Programming. Each chromosome of the population constructed by a series of genes represents an intact trajectory shown in Fig. 6, and each gene represents the heading variation angle between the initial heading and the heading at this turning point.

\[
\begin{array}{cccccccc}
\varphi_1 & \varphi_2 & \varphi_3 & \ldots & \varphi_n \\
\end{array}
\]

Fig. 6. The heading variation angle at different turning points

POPULATION INITIALIZATION

Compared with random population initialization in the SGA for ship’s trajectory planning, this paper adopts a heuristic method to generate initial populations. The idea is to set up constraints in the function of population initialization in accordance with the encounter type and good seamanship. First, to determine the obligation of collision avoidance between ships so that the direction of course changing is decided. If \( \varphi_1 > 0 \), the obligated ship turns right, otherwise, the obligated ship turns left. Then, considering the intention of evasive manoeuvre should be obvious according to COLREGs, the value of first gene should exceed 15° [20]. And the deviation of adjacent genes should be less than \( \pi/3 \). It’s remarkable that the original gene does not do genetic operation in order to satisfy the above-mentioned constraints. In fact, this heuristic method lowers the computational cost in the optimal process. The constraints are calculated as follows:

\[
\varphi_i \rightarrow \begin{cases} 
\pi / 12 < |\varphi_i| < \pi / 6 \\
\|\varphi_i + \varphi_{i+1}\| = \pi / 3 (i = 2, 3, \ldots, N)
\end{cases}
\]

FITNESS FUNCTION MODEL

The fitness evaluation of a chromosome measures the cost of a trajectory, which mainly accommodates three different optimal goals: 1) the safety requirements; 2) less distance traveled; 3) the restoration of navigational state. Therefore, we have selected a linear combination of these factors.

\[
f = \begin{cases} 
\beta \cdot |D_e| + \gamma \cdot D_t & D_s < D_e \\
\beta \cdot |D_e| + \alpha \cdot (D_t - D_s) & D_s > D_e
\end{cases}
\]

\[
D_e = \sqrt{(R_s(n,1) - R_t(n,1))^2 + (R_s(n,2) - R_t(n,2))^2}
\]

\[
D_t = \min(\sqrt{(R_s(t,1) - R_t(t,1))^2 + (R_s(t,2) - R_t(t,2))^2})
\]

\( (t = 0, T, 2 \times T, 3 \times T, \ldots, n \times T) \)

Where \( R_s, R_t, R_i \) are two-dimensional coordinate arrays that record the original trajectory of the OS, the trajectory of the TS and the optimal trajectory of the OS. \( D_e \) is the destination deviation between the optimal trajectory and the original trajectory of the OS. \( D_t \) is the minimum distance during the movement between the OS and TS. T is the time step. The weighting coefficients of \( \alpha \) and \( \beta \) indicate the risk of collision and the deviation between the optimal trajectory and the original trajectory of the OS. And \( y \) is the penalty factor.

GENETIC OPERATION

Selection operator: this paper adopts roulette wheel selection for real number encoding method.

\[
P_i = \frac{F_i}{\sum_{j=1}^{N} F_j}
\]

Where \( F_i \) is the fitness value of chromosome \( i \) and \( N \) is the total number of chromosomes for each population.

Crossover operator: this paper applies one-point crossover for real number encoding shown as follows:

\[
a_{ij} = a_{ij} (1 - b) + a_{ib} b
\]

\[
a_{ij} = a_{ij} (1 - b) + a_{ib} b
\]

Where \( a_{ij} \) is the value of \( j'-\text{th} \) gene in the chromosome \( k \); \( a_{ij} \)- the value of \( j'-\text{th} \) gene in the chromosome \( i \); \( b \) - a random number between 0 and 1.

Mutation operator: this paper adopts Gauss mutation operator which generates a random number conforming to normal distribution to replace the value of specific gene.

\[
a_{ij} = \text{normrnd}(0.5 \times (\text{bound}(i,1) + \text{bound}(i,2)), (\text{bound}(i,2) - \text{bound}(i,1))/6)
\]

Where \( \text{bound}(i,1) \) is the lower limit of \( i'-\text{th} \) gene and \( \text{bound}(i,2) \) is the upper limit of \( i'-\text{th} \) gene; the order of ‘normrnd’ in the Matlab generates random numbers from the normal distribution with mean parameter and standard deviation.
NONLINEAR OPTIMIZATION

The application of Nonlinear Programming is used to further tap the potential in the ability of local optimization and improve the superiority of the final solution. It is applied every 50 generations and the detailed constraints are described as follows:

\[
\begin{align*}
\min f(\varphi) &= D_x \\
\varphi_1 + \varphi_2 + \cdots + \varphi_n &= 0 \\
\varphi &\in [-30^\circ, 30^\circ] \\
c(x) &= D_x - D_t \leq 0 \\
c.(\varphi) &\in \emptyset
\end{align*}
\] (21)

SIMULATION AND ANALYSIS

Several typical traffic scenarios are adopted for simulation. The encounter type and the obligation are determined based on Fig. 2 and corresponding indicators given in Tab. 2. The simulation results are primarily intended to evaluate the optimal solutions in the aspects of the consistency and practicality. It is noticeable that the set of test cases are mainly based on target ships approaching the OS from a range of directions. And Tab. 3 shows the initial configuration of the traffic scenarios for all test cases.

Tab. 3. Ship encounter information for different traffic scenarios

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Head-on</th>
<th>Crossing</th>
<th>Overtaking</th>
<th>Multi-ship</th>
</tr>
</thead>
<tbody>
<tr>
<td>Position of OS(km)</td>
<td>(0,0)</td>
<td>(0,0)</td>
<td>(0,0)</td>
<td>(0,0)</td>
</tr>
<tr>
<td>Speed of OS(m/s)</td>
<td>7.5</td>
<td>7.5</td>
<td>11</td>
<td>7.5</td>
</tr>
<tr>
<td>Course of OS(°)</td>
<td>45</td>
<td>45</td>
<td>45</td>
<td>45</td>
</tr>
<tr>
<td>Position of TS(km)</td>
<td>(10,10)</td>
<td>(14,6)</td>
<td>(3,3)</td>
<td>(10,10)/(14,6)</td>
</tr>
<tr>
<td>Speed of TS(m/s)</td>
<td>5</td>
<td>7.5</td>
<td>5</td>
<td>5/7.5</td>
</tr>
<tr>
<td>Course of TS(°)</td>
<td>225</td>
<td>270</td>
<td>45</td>
<td>225</td>
</tr>
</tbody>
</table>

Let’s assume that the own ship’s speed is constant. And the OS only relies on changing its course to perform the process of collision avoidance, which is complaint with Rule 8 of the COLREGs for action to avoid a close-quarter situation. After a comprehensive consideration of the traffic scenarios for collision avoidance, the relevant parameters of the algorithm is set as follows: a population size of 3, a group size of 100, the termination condition is 100, and the parameters \( \alpha \) and \( \beta \) are equal to 0.6 and 0.4, respectively. In order to contrast the performance of the optimal trajectory, the simulation tests respectively performed by the hybrid genetic algorithm and SGA, adopt the same number of chromosomes. The hybrid genetic algorithm is ran five times in every traffic scenario. The optimal solutions for different test cases are respectively recorded in Tab. 4-7, which remain unchanged for the same input data. The corresponding curves of course variation, optimal trajectory and distance variation are shown in Fig. 7-9 (head-on), Fig. 10-12 (crossing), Fig. 13-15 (overtaking) and Fig. 16-18 (multi-target), accordingly. The evolutionary process is described in Fig. 19. In the trajectory curves, black line and black dash dot line respectively represent the optimal trajectory of the OS obtained by using the proposed algorithm and SGA, blue line indicates the trajectory of the TS, red dotted line indicates the original trajectory of the OS. The information about DCPA and the corresponding time (TCPA) are clearly marked in the distance variation curves between OS and TS.

HEAD-ON SITUATION

In this encounter situation, the TS comes towards the OS from the region \( P_1 \) and the corresponding parameters meets the requirement for HO given in Tab. 1. Meanwhile, the ship domain will be violated with the approaching of the TS. As a consequence, this encounter type can be interpreted by Rule 14 of the COLREGs, which indicates that the ships should pass each other port to port. It is remarkable that the evasive manoeuvre is conducted from the perspective of the OS. And the action taken by the OS guarantees the safety though the TS does not take any action. The course variation, the optimal trajectory and distance variation are presented in Fig. 7-9.

Tab.4. Optimal solutions in head-on situation

<table>
<thead>
<tr>
<th>Sequence</th>
<th>( \varphi_1(°) )</th>
<th>( \varphi_2(°) )</th>
<th>( \varphi_3(°) )</th>
<th>( \varphi_4(°) )</th>
<th>( \varphi_5(°) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Trajectory1</td>
<td>20</td>
<td>14</td>
<td>14</td>
<td>14</td>
<td>-10</td>
</tr>
<tr>
<td>Trajectory2</td>
<td>20</td>
<td>14</td>
<td>14</td>
<td>14</td>
<td>-10</td>
</tr>
<tr>
<td>Trajectory3</td>
<td>20</td>
<td>14</td>
<td>14</td>
<td>14</td>
<td>-10</td>
</tr>
<tr>
<td>Trajectory4</td>
<td>20</td>
<td>14</td>
<td>14</td>
<td>14</td>
<td>-10</td>
</tr>
<tr>
<td>Trajectory5</td>
<td>20</td>
<td>14</td>
<td>14</td>
<td>14</td>
<td>-10</td>
</tr>
</tbody>
</table>

Continuation of Tab. 4: Optimal solutions in head-on situation

<table>
<thead>
<tr>
<th>Sequence</th>
<th>( \varphi_6(°) )</th>
<th>( \varphi_7(°) )</th>
<th>( \varphi_8(°) )</th>
<th>( \varphi_9(°) )</th>
<th>( \varphi_{10}(°) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Trajectory1</td>
<td>-10</td>
<td>-10</td>
<td>-10</td>
<td>-10</td>
<td>-10</td>
</tr>
<tr>
<td>Trajectory2</td>
<td>-10</td>
<td>-10</td>
<td>-10</td>
<td>-10</td>
<td>-9</td>
</tr>
<tr>
<td>Trajectory3</td>
<td>-10</td>
<td>-10</td>
<td>-10</td>
<td>-10</td>
<td>-10</td>
</tr>
<tr>
<td>Trajectory4</td>
<td>-10</td>
<td>-10</td>
<td>-10</td>
<td>-10</td>
<td>-10</td>
</tr>
<tr>
<td>Trajectory5</td>
<td>-10</td>
<td>-10</td>
<td>-10</td>
<td>-10</td>
<td>-9</td>
</tr>
</tbody>
</table>

Fig. 7. The course variation of the optimal trajectory
CROSSING SITUATION

In this encounter situation, the TS comes towards OS from region $P_2$. Similarly, this encounter type allows to evaluate the collision risk interpreting Rule 15 of the COLREGs for a crossing situation. It is deemed practical to pass TS on its stern. The corresponding course variation, optimal trajectory and distance variation are shown in Fig. 10-Fig. 12.

Tab. 5. Optimal solutions in crossing situation

<table>
<thead>
<tr>
<th>Sequence</th>
<th>$\phi_1$ (°)</th>
<th>$\phi_2$ (°)</th>
<th>$\phi_3$ (°)</th>
<th>$\phi_4$ (°)</th>
<th>$\phi_5$ (°)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Trajectory1</td>
<td>19</td>
<td>15</td>
<td>15</td>
<td>15</td>
<td>-10</td>
</tr>
<tr>
<td>Trajectory2</td>
<td>19</td>
<td>15</td>
<td>15</td>
<td>15</td>
<td>-10</td>
</tr>
<tr>
<td>Trajectory3</td>
<td>19</td>
<td>15</td>
<td>15</td>
<td>15</td>
<td>-10</td>
</tr>
<tr>
<td>Trajectory4</td>
<td>19</td>
<td>15</td>
<td>15</td>
<td>15</td>
<td>-10</td>
</tr>
<tr>
<td>Trajectory5</td>
<td>19</td>
<td>15</td>
<td>15</td>
<td>15</td>
<td>-10</td>
</tr>
</tbody>
</table>

Continuation of Tab. 5: Optimal solutions in crossing situation

<table>
<thead>
<tr>
<th>Sequence</th>
<th>$\phi_6$ (°)</th>
<th>$\phi_7$ (°)</th>
<th>$\phi_8$ (°)</th>
<th>$\phi_9$ (°)</th>
<th>$\phi_{10}$ (°)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Trajectory1</td>
<td>-10</td>
<td>-10</td>
<td>-10</td>
<td>-10</td>
<td>-10</td>
</tr>
<tr>
<td>Trajectory2</td>
<td>-10</td>
<td>-10</td>
<td>-10</td>
<td>-10</td>
<td>-10</td>
</tr>
<tr>
<td>Trajectory3</td>
<td>-10</td>
<td>-10</td>
<td>-10</td>
<td>-10</td>
<td>-10</td>
</tr>
<tr>
<td>Trajectory4</td>
<td>-10</td>
<td>-10</td>
<td>-10</td>
<td>-10</td>
<td>-10</td>
</tr>
<tr>
<td>Trajectory5</td>
<td>-10</td>
<td>-10</td>
<td>-10</td>
<td>-10</td>
<td>-10</td>
</tr>
</tbody>
</table>

OVERTAKING

In this situation, the OS overtakes the TS from the stern of the target when the TS is situated in region $P_1$. The aim of this test is to evaluate the method’s interpretation of Rule 13 of the COLREGs that the OS should give way to the overtaken
ship. Similarly to head-on and crossing simulation, the key variable curves are described in Fig. 13-Fig. 15.

Table 6. Optimal solutions in overtaking situation

<table>
<thead>
<tr>
<th>Sequence</th>
<th>$\phi_1(^{\circ})$</th>
<th>$\phi_2(^{\circ})$</th>
<th>$\phi_3(^{\circ})$</th>
<th>$\phi_4(^{\circ})$</th>
<th>$\phi_5(^{\circ})$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Trajectory 1</td>
<td>22</td>
<td>18</td>
<td>12</td>
<td>-7</td>
<td>-7</td>
</tr>
<tr>
<td>Trajectory 2</td>
<td>22</td>
<td>18</td>
<td>12</td>
<td>-7</td>
<td>-7</td>
</tr>
<tr>
<td>Trajectory 3</td>
<td>22</td>
<td>18</td>
<td>12</td>
<td>-7</td>
<td>-7</td>
</tr>
<tr>
<td>Trajectory 4</td>
<td>22</td>
<td>18</td>
<td>12</td>
<td>-7</td>
<td>-7</td>
</tr>
<tr>
<td>Trajectory 5</td>
<td>22</td>
<td>18</td>
<td>12</td>
<td>-7</td>
<td>-7</td>
</tr>
</tbody>
</table>

Continuation of Tab. 6: Optimal solutions in overtaking situation

<table>
<thead>
<tr>
<th>Sequence</th>
<th>$\phi_6(^{\circ})$</th>
<th>$\phi_7(^{\circ})$</th>
<th>$\phi_8(^{\circ})$</th>
<th>$\phi_9(^{\circ})$</th>
<th>$\phi_{10}(^\circ)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Trajectory 1</td>
<td>-7</td>
<td>-7</td>
<td>-7</td>
<td>-7</td>
<td>-7</td>
</tr>
<tr>
<td>Trajectory 2</td>
<td>-7</td>
<td>-7</td>
<td>-7</td>
<td>-7</td>
<td>-7</td>
</tr>
<tr>
<td>Trajectory 3</td>
<td>-7</td>
<td>-7</td>
<td>-7</td>
<td>-7</td>
<td>-8</td>
</tr>
<tr>
<td>Trajectory 4</td>
<td>-7</td>
<td>-7</td>
<td>-7</td>
<td>-7</td>
<td>-8</td>
</tr>
<tr>
<td>Trajectory 5</td>
<td>-7</td>
<td>-7</td>
<td>-7</td>
<td>-7</td>
<td>-7</td>
</tr>
</tbody>
</table>

MULTI-SHIP ENCOUNTER SITUATION

In this situation, the OS is in a complex condition which faces the encounter types of head-on and crossing at the same time. As there is no risk between target ships shown in Fig. 13, the OS is the only manoeuvring party to perform evasive actions. Therefore, the overall output for the multi-ship encounter situation is illustrated in Fig. 16 - Fig. 18.

Table 7. Optimal solutions in multi-ship encounter situation

<table>
<thead>
<tr>
<th>Sequence</th>
<th>$\phi_1(^{\circ})$</th>
<th>$\phi_2(^{\circ})$</th>
<th>$\phi_3(^{\circ})$</th>
<th>$\phi_4(^{\circ})$</th>
<th>$\phi_5(^{\circ})$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Trajectory 1</td>
<td>21</td>
<td>14</td>
<td>14</td>
<td>14</td>
<td>-10</td>
</tr>
<tr>
<td>Trajectory 2</td>
<td>21</td>
<td>14</td>
<td>14</td>
<td>14</td>
<td>-10</td>
</tr>
<tr>
<td>Trajectory 3</td>
<td>21</td>
<td>14</td>
<td>14</td>
<td>14</td>
<td>-10</td>
</tr>
<tr>
<td>Trajectory 4</td>
<td>21</td>
<td>14</td>
<td>14</td>
<td>14</td>
<td>-10</td>
</tr>
<tr>
<td>Trajectory 5</td>
<td>21</td>
<td>14</td>
<td>14</td>
<td>14</td>
<td>-10</td>
</tr>
</tbody>
</table>

Continuation of Tab. 7: Optimal solutions in multi-ship encounter situation

<table>
<thead>
<tr>
<th>Sequence</th>
<th>$\phi_6(^{\circ})$</th>
<th>$\phi_7(^{\circ})$</th>
<th>$\phi_8(^{\circ})$</th>
<th>$\phi_9(^{\circ})$</th>
<th>$\phi_{10}(^\circ)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Trajectory 1</td>
<td>-10</td>
<td>-10</td>
<td>-10</td>
<td>-10</td>
<td>-10</td>
</tr>
<tr>
<td>Trajectory 2</td>
<td>-10</td>
<td>-10</td>
<td>-10</td>
<td>-10</td>
<td>-10</td>
</tr>
<tr>
<td>Trajectory 3</td>
<td>-10</td>
<td>-10</td>
<td>-10</td>
<td>-10</td>
<td>-10</td>
</tr>
<tr>
<td>Trajectory 4</td>
<td>-10</td>
<td>-10</td>
<td>-10</td>
<td>-10</td>
<td>-10</td>
</tr>
<tr>
<td>Trajectory 5</td>
<td>-10</td>
<td>-10</td>
<td>-10</td>
<td>-10</td>
<td>-10</td>
</tr>
</tbody>
</table>
DISCUSSION

From the simulation results, some conclusions can be drawn. Firstly, the minimum distance between ships (DCPA) in different traffic scenarios exceeds the safety criterion-radius of ship domain (D), which meets the safety requirement in the fitness function model. Furthermore, the uniform variation in the relative distance as well as the closeness between DCPA and the radius of ship domain, reflected in the total length of optimal trajectory, is relatively short, which guarantees the distance requirement. Meanwhile, the optimal trajectories in different traffic scenarios contain intact evasive manoeuvres, including the action of collision avoidance and the restoration to the original trajectory. Additionally, the consistency of the evasive manoeuvres is mainly reflected in two aspects. One is that the obligation is certain as the encounter type is derived from COLREGs whether from the position of the OS or the TS, the other is that the optimal solutions presented in each Table (Tab. 4-7) are nearly the same, and the corresponding optimal trajectories and distance variation curves almost overlap. More importantly, it is clear in the course variation curves that the evasive manoeuvres simulated by the hybrid genetic algorithm is performed with fewer course deviations even when multiple ships are involved, compared with the manoeuvres by SGA. This makes the optimal solution more suitable for the common practice of navigators. As for the searching ability, the fitness values in Fig. 19 reflects the better convergence effect by the hybrid genetic algorithm. Meanwhile, there is an obvious leap in the 50th generation which reflects the influence of the Nonlinear Programming on the ability of local optimization. As a result, we may draw the conclusion that the output of the hybrid genetic algorithm is improved in the aspects of consistency and practicality than that of SGA, i.e. demonstrates the feasibility and superiority of the hybrid genetic algorithm.

EVOLUTIONARY PROCESS

In order to verify the superiority, we run the hybrid genetic algorithm and SGA with the same number of chromosomes to compare as to fitness value. The convergence process in different traffic scenarios is shown in Fig. 19.

CONCLUSION

A novel trajectory planning based on hybrid genetic algorithm has been presented. And several traffic scenarios have been set up for verifying the feasibility and superiority of this proposed algorithm. By comparing the results of the hybrid genetic algorithm and SGA in the aspect of the optimal trajectory, course variation, distance variation as well as the evolutionary process, we can conclude that the hybrid genetic algorithm this hybrid genetic algorithm greatly improves the searching ability and guarantees the optimal solution unchanged for the same input data. Moreover, the final solution with fewer course deviations is more suitable for the practical application.

However, this paper does not consider the influence of ship manoeuvrability as it only discuss the easy multi-ship encounter situation. Therefore, further research plan is aimed at solving the above mentioned issues and establishing the feedback mechanism for monitoring the ship’s operational state, which helps to form an intact decision support system.
Meanwhile, we should try to apply Pareto-based multi-objective optimization method and other selection operators for improving the efficiency of the algorithm.

ACKNOWLEDGEMENTS

This work is supported by National Natural Science Foundation of China under Grant No. 51309041.

BIBLIOGRAPHY


CONTACT WITH THE AUTHORS

Shengke Ni
e-mail: 863369021@qq.com
Dalian Maritime University
Linghai road
116026 Dalian
CHINA

Zhengjiang Liu
e-mail: 2522402892@qq.com
Dalian Maritime University
Linghai road
116026 Dalian
CHINA

Yao Cai
e-mail: 3492105068@qq.com
Dalian Maritime University
Linghai road
116026 Dalian
CHINA

Xin Wang
e-mail: 875924999@qq.com
Dalian Maritime University
Linghai road
116026 Dalian
CHINA
SHIP TRAJECTORY TRACKING CONTROL SYSTEM DESIGN BASED ON SLIDING MODE CONTROL ALGORITHM

Yong Liu
Renxiang Bu
Xiaori Gao
Dalian Maritime University, China

ABSTRACT

The paper reports the design and tests of the planar autopilot navigation system in the three-degree-of-freedom (3-DOF) plane (surge, sway and yaw) for a ship. The aim of the tests was to check the improved maneuverability of the ship in open waters using the improved nonlinear control algorithm, developed based on the sliding mode control theory for the ship-trajectory tracking problem of under-actuated ships with static constraints, actuator saturation, and parametric uncertainties. With the integration of the simple increment feedback control law, the dynamic control strategy was developed to fulfill the under-actuated tracking and stabilization objectives. In addition, the LOS (line of sight) guidance system was applied to control the motion path, whereas the sliding mode controller was used to emulate the rudder angle and propeller rotational speed control. Firstly, simulation tests were performed to verify the validity of the basic model and the tracking control algorithm. Subsequently, full scale maneuverability tests were done with a novel container ship, equipped with trajectory tracking control and sliding mode controller algorithm, to check the dynamic stability performance of the ship. The results of the theoretical and numerical simulation on a training ship verify the invariability and excellent robustness of the proposed controller, which: effectively eliminates system chattering, solves the problem of lateral drift of the ship, and maintains the following of the trajectory while simultaneously achieving global stability and robustness.

Keywords: LOS; ship trajectory; tracking control system; sliding mode control

INTRODUCTION

Unmanned surface vessels (USVs) are attracting a great deal of attention from researchers all over the world because of their extensive applications in military reconnaissance, homeland security, shallow water surveys, environmental monitoring, and operational coordination with autonomous underwater vehicles (AUVs) [1-3]. Three types of control technologies: set-point control, trajectory tracking control, and path following control, play a crucial role in allowing USVs to achieve their specified tasks automatically [4]. Set-point control [5–7] is important for dynamic positioning of a vessel in fixed target operations, such as autonomous docking. Trajectory tracking control [8–12] enables the ship to track the desired time-referenced trajectory. In the path-following control scenario [13–14], a USV is required to follow the path at a certain speed without specified temporal constraints. The path is described by curve parameters, which are usually not time relevant. There are neither lateral nor vertical thrusters on most AUVs, and only longitudinal, yaw and pitch angle speeds are controlled directly. Therefore, the AUV is a typical under-actuated system, which makes trajectory tracking more difficult [15].

Trajectory tracking control of ships is especially difficult due to the effects of underactuativity [16-17] and nonholonomic constraint characteristics [18] of common surface ships, with the addition of the impact of nonlinearity [19], uncertainty [20],
as well as wind and current susceptibility[21-23]. Normally these ships are not equipped with side thrusters, which might only be used when approaching or departing port. Reference [24] integrated the output redefinition method and the sliding-mode control technology, and applied them to the nonlinear non-minimum phase system design to acquire an asymptotically stable effect of ship straight-path tracking control. References [25] and [26] put forward a control algorithm taking into account current interference on the fourth order ship model, which estimates the uncertainty of the current interference. Reference [27] designed an output feedback controller on the basis of the nonlinear observer, and conducted the prototype model simulation experiment. The authors managed to perform straight-path tracking control globally and steadily in the case of external disturbances of wind, wave and current. Reference [28] integrated the Nussbaum technique into the backstepping method, and put forward a robust-adaptive fuzzy control algorithm, which enabled the ship cross track error to reach asymptotically a small neighborhood of the origin at exponential rate. To achieve trajectory tracking or positioning, various control strategies have been proposed, including robust adaptive control [29], sliding mode control [30-32], decoupling control [33], and adaptive neural control [34-37].

The maneuvering motion model proposed by the Japanese Ship Maneuvering Mathematical Model Group and called the MMG model is frequently used to simulate the motion of a ship. The MMG model divides the ship into ship hull, propeller, and rudder, and examines their performance. It can consider different values of fluid power that acts effectively on the ship. The MMG Group further improved the MMG model [38] by incorporating a method to correct the hydrodynamic derivatives in shallow water. As for ship motion guidance, LOS (line of sight) guidance is adopted, which can control the course between consecutive waypoints through correcting the cross-track error (XTE), and can also be used for tracking a moving target on the sea.

The ship trajectory tracking control system is a closed loop feedback system, the aim of which is to reduce the tracking error. Once the tracking error meets the accuracy requirements, the program executes the next step until its completion. The trajectory tracking control system presented in the paper has been established for a container ship. The construction idea of this system is shown in Figure 1.

**MATHEMATICAL MODELING OF SHIP TRAJECTORY TRACKING**

**FIXED COORDINATE SYSTEM AND KINETIC COORDINATE SYSTEM**

Diagrammatic drawing of ship plane coordinates is shown in Figure 2. In the diagram, $\mathbf{x}_E \mathbf{y}_E$ is the fixed coordinate system in which the axis $\alpha_x \mathbf{y}_E$ points to true north, and the axis $\alpha_y \mathbf{y}_E$ points to due east. $G$ is the ship’s center of gravity, and its position is represented by $(x, y)$; $\phi$ is the course angle, and clockwise is positive; $V$ is the speed in tangent direction of ship’s center of gravity over the ground in the plane, and $Vx$, $Vy$ are the projections in the body-fitted coordinate system. $U$ is the speed through water, and $\beta$ is the drift, defined as the intersection angle between the tangential direction of speed through water and the direction of ship fore and aft centerline, and port side is positive. $\phi_c$ and $\beta_c$ represent the angle and rate of homogeneous flow, respectively.

**DEFINITION OF SHIP TRAJECTORY TRACKING ERROR**

By simple coordinate transformation while maintaining generality, the ship lateral tracking error $y$ can be defined as the distance on the rhumb line from the ship to the line connecting two consecutive waypoints, as shown in Figure 3.
MATHEMATICAL MODEL OF SHIP TRAJECTORY TRACKING CONTROL

When the ship sails at fixed speed at sea without controlling the longitudinal motion, the mathematical model of ship trajectory tracking control with the interference of wind and current can be expressed as:

\[
\begin{align*}
\dot{y} &= u \sin \varphi + c(t) \\
\dot{\varphi} &= r \\
\dot{r} &= f(u,v,r) + g(u,v,r, \sigma) + d(t) \\
\dot{\delta} &= K \frac{(\delta - \delta)}{T} \\
c(t) &= v \cos \varphi + u \sin \varphi
\end{align*}
\]

(1)

In this formula, \(y\) is the transverse cross-track error, \(u\) is the ship's speed through water, \(v\) is the transverse speed through water, \(r\) is the heading rate, \(\delta\) is the rudder angle, \(\delta\) is the commanded rudder angle, \(T\) is the time constant of the steering gear, \(f(u,v,r)\) and \(g(u,v,r,\sigma)\) are the nonlinear functions containing \(r\) and \(\delta\), \(d(t)\) is the unknown external disturbance torque, and \(c(t)\) is the drift caused by the unknown current disturbance and ship transverse motion.

SLIDING MODE CONTROLLER DESIGN AND STABILITY ANALYSIS

CONTROL OBJECTIVE AND ASSUMPTION TERMS

The objectives of ship trajectory tracking control are: taking advantage of appropriate methods to determine \(\delta\), and making the cross-track error \(y\) approach zero. In other words, its goal is to make the ship sail on the scheduled track line. In the case of transverse wind and current disturbances, the ship needs to sail at a certain leeway and drift angle. In this case, the yaw rate approaches zero, while the course deviation \(\varphi\) should not be zero.

According to practical conditions of navigation, let us make the following assumptions for system (1):

Assumption 1: \(u \gg u_1, v \gg v_1\), which means that the longitudinal speed is far higher than the flow rate and the transverse speed;

Assumption 2: External interference is bounded;

Assumption 3: Ship course is controllable;

Assumption 4: \(u,v,u, r\) is bounded, and \(u_{\text{max}}, v_{\text{max}}, u_{\text{r max}}, r_{\text{max}}\) are known;

Assumption 5: Lipschitz constants \(L_1, L_2, L_3, L_4\) exist and satisfy the following inequation:

\[
\begin{align*}
|u_{t2} - u_{t1}| &\leq L_1 \Delta t \\
|c_{t2} - c_{t1}| &\leq L_2 \Delta t \\
|f(t_2) - f(t_1)| &\leq L_3 \Delta t \\
|d_{t2} - d_{t1}| &\leq L_4 \Delta t
\end{align*}
\]

(2)

Assumption 1 is easy to satisfy in normal weather and sea state, but when the transverse speed is higher than the longitudinal speed, the ship cannot keep the track. Assumptions 2 and 3 are the prerequisites for a controllable closed-loop system. Assumptions 4 and 5 are the inevitable results of action of the bounded ship control energy and the external interference energy.

REDEFINITION OF OUTPUT VARIABLES AND ANALYSIS OF TRAJECTORY TRACKING ERROR STABILITY

Considering practical conditions of wind and current disturbances, let us redefine the output variables:

\[
Z = \varphi + k_1 \arctan(k_2 y) + k_2 \int_{t_0}^{t} \left( \dot{y} + k_2 \theta(h(y)) \right) dt
\]

(3)

where \(\theta(h)\) denotes the hyperbolic tangent and \(k_1, k_2, k_3, k_4, k_5 \in R^+\) are the design parameters. The main reasons for adopting (3) are: 1) to control the ship track error by finding the nonlinear relationship between the heading \(\varphi\) and the track error \(y\), 2) to ensure that the system is in a controllable state by using arc tangent and hyperbolic tangent functions with strict bounds in the range \((-1,1)\), 3) to ensure that the heading \(\varphi\) and the track error \(y\) are at the same balance point and to ensure the stability of the system, 4) to control the system convergence rate with \(k_1, k_2, k_3, k_4\) and to adjust the slope of the curve with \(k_5, k_6\).

When \(Z\) approaches zero:

\[
\varphi = -k_1 \arctan(k_2 y) - k_2 \theta(h(y))
\]

(4)
Making \( \sigma = u \sin \varphi \), according to equations (1) and (4), we get:

\[
\begin{align*}
\dot{y} &= \sigma + c(t) \\
\sigma &= -C\sigma - Dk,th(k_2y) - C \cdot c(t) 
\end{align*}
\]  

In order to make the following formulas more clear, we use \( C \) and \( D \) to represent the complex variables in formula (5):

\[
C = \frac{u \cos \varphi \cdot k_2}{1 + (k_0y)^2} \frac{u \cos \varphi \cdot k_3}{ch^2(\int (\dot{y} + k,th(k_2y))dy)}, \quad D = \frac{u \cos \varphi \cdot k_4}{ch^2(\int (\dot{y} + k,th(k_2y))dy)}. 
\]

In practical navigation, the leeway and drift angle is smaller than 90\(^\circ\), that is \( \varphi \in (-\pi/2, \pi/2) \), then \( \cos \varphi > 0 \) and, consequently, \( C > 0, D > 0 \).

Applying the Lyapunov function:

\[
V = (\sigma + Cy)(\dot{\sigma} + Cy)
\]

\[
\begin{align*}
\dot{V} &= \sigma\dot{\sigma} + C\dot{y}\dot{\sigma} + Cy\dot{\sigma} + C^2\dot{y}y \\
&= \sigma(-C\sigma - Dk,th(k_2y) - C \cdot c(t)) + C(\sigma + c(t)) \\
&+ Cy(-C\sigma - Dk,th(k_2y) - C \cdot c(t)) \\
&+ C^2y\sigma \cdot c(t) \\
&= -C\sigma^2 - Ck,th(k_2y) - C\sigma \cdot c(t) \\
&+ C^2\sigma \cdot c(t) - C^2y \cdot c(t) \\
&+ C^2y \cdot c(t) + C^2\sigma \cdot c(t) \\
&= -Dk,th(k_2y) - CDk,th(k_2y) \\
&\quad : C > 0, D > 0 \quad and \quad (y > 0 : k,th(k_2y) > 0 : -CDk,th(k_2y) < 0 \quad y < 0 : k,th(k_2y) < 0 : -CDk,th(k_2y) < 0) (7)
\end{align*}
\]

According to the order of magnitude:

\[
[Dk,th(k_2y), CDk,th(k_2y)] < [\sigma + Cy]
\]

\[
\begin{align*}
&\quad < [\sigma + Cy] \\
&\quad < [u \sin(\varphi)] < \frac{u \cos \varphi \cdot k_2}{1 + (k_0y)^2} \frac{u \cos \varphi \cdot k_3}{ch^2(\int (\dot{y} + k,th(k_2y))dy)}, y \\
&\quad < [\sin(\varphi)] < \frac{\cos \varphi \cdot k_2}{1 + (k_0y)^2} \frac{\cos \varphi \cdot k_3}{ch^2(\int (\dot{y} + k,th(k_2y))dy)}, y \\
&\quad < [\sin(\varphi)] < \frac{\cos \varphi \cdot k_2}{1 + (k_0y)^2} \frac{\cos \varphi \cdot k_3}{ch^2(\int (\dot{y} + k,th(k_2y))dy)}, y \\
&\quad < [\sin(\varphi)] < [\varphi + 1]
\end{align*}
\]

Formula (8) always holds.

Furthermore, if \( y = 0 \), then \( \dot{y} = \sigma + c(t) = 0 \) and \( \sigma = -c(t) \). We arrive at: \( u \sin \varphi = -c(t) \neq 0 \).

The above equation states that the ship has the heading angle \( \varphi \neq 0 \) in the presence of wind and current disturbances. Consequently, the ship needs to have a leeway angle while navigating at sea.

As such: \( V = -Dk,th(k_2y) - CDk,th(k_2y) < 0 \) always stands up.

In conclusion, when the global convergence of \( V \) approaches zero, \( V = \frac{1}{2}(\varphi + Cy)^2 > 0 \) and \( V < 0 \), and the system (5) is asymptotically stable. For formula (4), in the presence of wind and current disturbances, the cross-track error \( y \) is asymptotically stable and converging, and the track control is converted into the stabilization control of the intermediate variable \( Z \).

**DESIGN OF SHIP TRAJECTORY TRACKING CONTROLLER BASED ON SLIDING-MODE METHOD**

For formula (3), in order to reduce the calculative complexity level in the process of working out \( \delta \) as well as for the sake of simulation convenience, the control design is given in the following form:

\[
\begin{align*}
f_1(y) &= k_1,th(k_2y) + \dot{y} \\
f_1(f_1, \varphi) &= \varphi + k_1 \arctan(k_0y) + k_1 \int f_1(y)dy, \quad f_1(f_2, \varphi) = k_1,th(f_2) + r + k_1 \arctan'(k_0y) + k_1 \arctan(f_1(y)) \\
f_1(f_2) &= k_1,th(f_2) + f_2
\end{align*}
\]

Now, the control objective of the system is converted into the control of \( f_2 \), that is: when \( f_2 \) is zero, we get \( f_2 = 0 \), then \( \dot{y} = k_1,th(k_2y) + \dot{y} \). The control law in the presence of wind and current disturbances:

\[
\begin{align*}
\varphi^* &= k_2 \arctan(k_0y) - k_1,th \int_0^t \dot{y} + k_1,th(k_2y)dy, \quad f_2 \bigg|_{f_2 = 0}
\end{align*}
\]

In the case of \( f_2 = 0 \), we have proved that the system is stable.

After expanding \( f_2 \), we get:

\[
\begin{align*}
f_2 &= k_1,th(f_1) + f_2 \\
&= k_1,th(f_1) + k_1,th'(f_1) + r + k_2 \arctan'(k_0y) + k_1,th^2 \int f_1dy, \quad f_2 = k_1,th(f_1) + k_1,th(f_2) + f(u, v, r) + f(u, v, r, \delta) + d(t) \\
&+ k_1 \arctan'(k_0y) + k_1,th^2 \int f_1dy
\end{align*}
\]

Making \( f_2 = 0 \), we get the control law in the presence of wind and current disturbances:

\[
\delta = f^{-1}(u, v, r, \delta)(f(\varphi(t), y(t)) + f(u, v, r) + d(t))
\]

In this formula, \( f^{-1}(u, v, r) \) is the indeterminate and complicated smooth and continuous function; \( f(\varphi(t), y(t)) \) is the
nonlinear function with unknown structure which includes \( \varphi \) and \( y \); \( f(u,v,r) \) is the nonlinear function that includes \( r \); and \( d(t) \) is the unknown function.

Problems which appear in system (1) refer to parameter uncertainty as well as to wind and current disturbances. Moreover, the process of finding \( \delta \) in formula (11) is rather complicated, which may lead to “calculation inflation”. These disadvantages are not conducive to the realization of the simulation and actual application. To address the main goal of the path-following problem which is finding the nonlinear relationship between the rudder angle \( V \) and the track deviation \( y \), we can use the intermediate variable \( V \) to carry out the transition and ultimately determine this relationship. This paper adopts the following increment feedback control law [38]:

\[
\dot{\delta} = -p \cdot f_2 - \epsilon \text{sgn}(f_1) \quad p, \epsilon \in \mathbb{R}^+ \quad (13)
\]

In this formula, \( p \) is the proportional feedback coefficient with a constant value.

STABILITY ANALYSIS

Substituting \( \dot{y} = u \sin(\varphi) + c(t) \) into \( f_1(y) = k_1\vartheta(k_2y) + \dot{y} \), we get:

\[
f_1(y) = k_1\vartheta(k_2y) + u \sin \varphi + c(t) \quad (14)
\]

Let us assume that for all \( t \in [t_0, t_1] \), \( t_1 = t_0 + \Delta t \), if \( f_4(t_1) > 0 \), \( f_4(t) < 0 \), then, in line with function continuity, we can infer that there is a \( \xi \) which \( 0 \leq \xi \leq 1 \), when \( t = t_1 + \xi \Delta t \), \( f_4(t_1) = 0 \), and for all \( t \in [t_0, t_1] \), \( f_4(t) > 0 \). Then, when \( t = t_2 \), from formula (9) and the definition of derivative, we get:

\[
\Delta f_4 = k_1[\vartheta(f_1(t_1)) - \vartheta(f_2(t_1))] + k_2[\vartheta'(f_1(t_1)) - \vartheta'(f_2(t_1))] + f(u,v,r(t_1)) - f(u,v,r(t_2)) + f^{-1}(u,v,r,d(t_1) - u,v,r,d(t_2)) + d(t_1) - d(t_2) + k_1[\arctan(k_2y(t_1)) - \arctan(k_2y(t_2))]
\]

\[
+k_2[\vartheta'' f_1(t_1) d_1 - \vartheta'' f_1(t_1) d_1] \quad (15)
\]

From formula (13), we get:

\[
\Delta \delta = \dot{\delta}(t_1) \Delta t = -[p \cdot f_1(t_1) + \epsilon] \Delta t \quad 0 \leq \xi \leq 1 \quad 0 < f_1(t_1) \quad (16)
\]

Substituting formula (16) into formula (15), and considering formula (2), we get:

\[
\Delta f_4 \leq 2k_2 + k_4 + L \Delta T - f^{-1} (p \cdot f_1(t_1) + \epsilon) \Delta t + L_4 \Delta t
\]

\[
+2k_2k_1L_4 + (k_2k_3 + k_4u_{\text{max}} + k_4k_4 + k_4L_4) \Delta t
\]

\[
+2k_2k_2u_{\text{max}} + u_{\text{max}} + v_{\text{max}} + 2k_4(L_4 + k_4u_{\text{max}} + L_4) \quad (17)
\]

From the known conditions of Assumptions 1, 2, and 5, the leeway and drift angle does not exceed 90° during the voyage, as well as \( f_1(t) > 0 \). Consequently, when \( t \in [t_1, t_2] \), there is \( p \in \mathbb{R}^+ \), render \( \Delta f_4 < 0 \); and when \( t \in [t_0, t_1] \), there is \( f_1(t) < 0 \) and \( p \in \mathbb{R}^+ \), render \( \Delta f_4 > 0 \).

In conclusion, \( f_1 \) is successive in the closed interval \( [t_1, t_2] \), and \( f_1(t) \) and \( f_1(t) \) are of opposite sign, then in the open interval \( (t_1, t_2) \), there is at least one \( \xi \) which renders \( f_1(\xi) = 0 \) and \( f_1(t) \) are of opposite sign, then \( f_1(t) \) is stable and converges to zero. Q.E.D.

SHIP TRAJECTORY TRACKING CONTROL DESIGN SYSTEM

WAYPOINT LOS GUIDANCE

LOS is used to control the course between two waypoints, thus making the ship navigate towards the next waypoint along the set route. After setting the expected course, LOS defines it as a function of the cross-track error, then calculates the error and corrects it by steering. Finally, navigation is achieved by making the variable heading value convergent. See Figure 4 for detail.
It is used to judge whether the target point has been reached and the guidance for next waypoint can start. Therefore, the smaller the value $n$, the higher the accuracy of guidance, but the difficulty increases if $n$ is too small, so we set $n=1$ in this paper.

**SHIP’S ROUTE PLANNING**

A simple route analyzed in this paper is shown in Figure 5. A coordinate system has been established with the initial position of the ship as the origin $(0,0)$. The set waypoints shown in the figure have been selected assuming that there are two islands on both sides of the ship route into port, and a narrow channel between them. Starting from the initial position, the ship is expected to go through waypoints 1, 2, 3 and 4.

**SHIP TRAJECTORY TRACKING SIMULATION AND VALIDATION WITH AUTONOMOUS TRACKING SYSTEM**

To consider the accuracy and stability of the method, we used the fourth-order Runge-Kutta method to solve the ordinary differential equation system. The research object in this paper was the container ship A. The main parameters of hull, propeller and rudder for this ship are shown in Table 2.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Length between perpendiculars $L$/m</td>
<td>172</td>
</tr>
<tr>
<td>Beam $B$/m</td>
<td>32.2</td>
</tr>
<tr>
<td>Fore draft $d_f$/m</td>
<td>11.3</td>
</tr>
<tr>
<td>Aft draft $d_a$/m</td>
<td>11.3</td>
</tr>
<tr>
<td>Displacement $V$/m$^3$</td>
<td>50500</td>
</tr>
<tr>
<td>Wetted area $S$/m$^2$</td>
<td>8201</td>
</tr>
<tr>
<td>Block coefficient $C_b$</td>
<td>0.807</td>
</tr>
<tr>
<td>Propeller diameter $D$/m</td>
<td>5.9</td>
</tr>
<tr>
<td>Pitch diameter ratio $P/D$</td>
<td>0.8</td>
</tr>
<tr>
<td>Blade number $Z$</td>
<td>4</td>
</tr>
<tr>
<td>Area ratio $EAR$</td>
<td>0.7</td>
</tr>
<tr>
<td>Rudder aspect ratio $\lambda_\delta$</td>
<td>1.7872</td>
</tr>
<tr>
<td>Rudder height $H$/m</td>
<td>8.4</td>
</tr>
<tr>
<td>Rudder area $A_\delta$/m$^2$</td>
<td>39.48</td>
</tr>
</tbody>
</table>

The assumed initial parameters of ship motion were as follow: initial speed $u_0=7.2$m/s, course $\Psi_0=26.56$, and initial rotational speed of propeller $n_0=3$rps. The rotational speed of the propeller remained unchanged until the ship reached waypoint 4. The simulation route is shown in Figure 6.

From the simulation result we can see that the ship almost passes through all waypoints of the expected route. A relatively large offset is observed after passing waypoints 1 and 2 by the ship, because the heading changes noticeably in these areas. However, the offset is smaller than the ship length, and the simulation route corresponds to the expected route, so the route stability is maintained.

Time-dependent changes of longitudinal speed $u$, transverse speed $v$, rate of turning $\dot{\psi}$, heading $\Psi$, and rudder $\delta$ of the ship, which were recorded in the simulation are shown in Figure 7.

We can see in Figure 7 that all motion parameters change noticeably when the ship approaches each waypoint, especially waypoints 1, 2 and 3, because there is large course-alerting at
each point. The results of the simulation show that the cross-track error converges smoothly and without overshoot or buffeting, while the course and rudder angle oscillate around a fixed value. The control effect is satisfying and robust.

CONCLUSIONS

In the paper, mathematical modeling, numerical calculation method, programming language design, and graphic simulation method, are applied synthetically. The automatic navigation and trajectory tracking control system of 3-DOF in the plane for a ship is established. The following conclusions can be formulated based on the results of modeling and simulation.

With respect to ship nonlinear tracking control systems, this article has redesigned the cross-track error $y$ and the course deviation $\phi$ by taking advantage of strict bounds of the hyperbolic tangent function and the integral term to improve the effect of trajectory tracking. The nonlinear and robust control algorithm proposed in this work and the linear track controller derived using the sliding-mode and incremental feedback methods can effectively overcome the problem of unknown origin. Moreover, the tracking simulation performed for a container ship proves the feasibility of the sliding mode algorithm and the automatic navigation and trajectory tracking system. Good congruency is observed between the simulation route and the expected route. In a word, the simulation of ship trajectory tracking has been successfully conducted and the simulation system has been successfully established.

ACKNOWLEDGEMENT

This work was supported by “Natural Science Foundation of Liaoning Province, China” (Number:201602084) and “The Fundamental Research Funds for the Central Universities” (Number: 3132016002, 3132016006, 3132016008).

REFERENCES


CONTACT WITH THE AUTHORS

Yong Liu
e-mail: liuyong_dlmu@163.com
Dalian Maritime University
No.1, Linghai RD
116000 Dalian
China

Renxiang Bu
e-mail: liushuangou0@163.com
Dalian Maritime University
No.1, Linghai RD
116000 Dalian
China

Xiaori Gao
e-mail: 425339930@qq.com
Dalian Maritime University
No.1, Linghai RD
116000 Dalian
China
NUMERICAL SIMULATIONS OF SEA ICE CONDITIONS IN THE BALTIC SEA FOR 2010–2016 WINTERS USING THE 3D CEMBS MODEL

Maciej Janecki¹
Artur Nowicki¹
Alicja Kańska¹
Maria Golenko³
Lidia Dzierzbicka-Głowacka¹

¹ Institute of Oceanology, Polish Academy of Sciences, Poland
² Institute of Meteorology and Water Management National Research Institute, Maritime Branch, Hydrological Forecasting Office, Poland
³ P.P. Shirshov Institute of Oceanology, Atlantic Branch, Russia

ABSTRACT

Sea ice conditions in the Baltic Sea during six latest winters – 2010/2011 to 2015/2016 are analysed using coupled ice-ocean numerical model 3D CEMBS (3D Coupled Ecosystem Model of the Baltic Sea). Simulation results are compared with observations from monitoring stations, ice charts and satellite data. High correlation between model results and observations has been confirmed both in terms of spatial and temporal approach. The analysed period has a high interannual variability of ice extent, the number of ice days and ice thickness. Increasing number of relatively mild winters in the Northern Europe directly associated with climate change results in reduced ice concentration in the Baltic Sea. In this perspective, the implementation and development of the sea ice modelling approach (in addition to standard monitoring techniques) is critical to assess current state of the Baltic Sea environment and predict possible climate related changes in the ecosystem and their influence for human marine–related activities, such as fishery or transportation.

Keywords: Baltic Sea, sea ice, numerical modelling, remote sensing, climate change

INTRODUCTION

The Baltic Sea is a unique water area surrounded on all sides by land surface, connected with the ocean only by narrow straits. Including the Kattegat, Baltic Sea covers an area of approximately 420 000 km². As a result of abundant freshwater runoff from the surrounding land, Baltic Sea has very low salinity of 6 to 8 on the central basin's surface waters that drops below 3 in the northernmost regions. Apart from a few deeps, the sea is very shallow with an average depth of about 52 m. The Baltic Sea freezes at least partially in winter, which can directly affect navigability and maritime infrastructure. Usually, most of the regular marine transport routes are unaffected. However, a solid ice sheet forms during the most severe winters, making ports along the Polish and German coasts closed by ice for several days a year. Despite the negative effect of freezing for maritime traffic, the Baltic Sea ice is a very important factor regulating North European climate system [1, 2]. While open waters reflect only 3% to 10%, sea can reflect from 50% to even 70% of the incoming radiation [3]. Thus, ice cover affects not only optical, but also chemical and hydrographical properties of the ecosystem, as well as its biological part. Sea ice conditions in the Baltic Sea are dynamic with large interannual ice extent variability. Water freezing in the Baltic Sea usually starts at the end of October or at the beginning of November in the northern parts of the Bothnian Bay and in the inner regions of the Gulf of Finland. Thereafter, ice formation expands toward the central part of the Bothnian Bay and the coastal areas of the Bothnian Sea. During severe winters both the Bothnian Sea and the Gulf of Finland can
freeze completely up to the northern part of the Baltic Proper. In years of relatively mild winters a compact ice cover hardly exists on Bothnian Sea and occurs only locally in the Gulf of Finland. The melting season begins in April in the southern parts of the Baltic Sea progressing towards the north. In early May ice covers only the northern part of the Bothnian Bay and disappears completely in the first half of June at the latest. The annual maximum ice extent of the Baltic Sea (MIB) occurs roughly between January and March, usually in late February or early March. In average at this time, ice covers ~40% of the total area of the Baltic Sea, which is about 165 000 km² [4, 5]. Due to specific interannual variability of ice conditions in the Baltic there are periods of time when significant anomalies occur in this pattern. The lowest MIB of 49 000 km² was recorded in 2008, while the winter of 1986/1987 was the most severe in the history of observations with MIB of over 420 000 km². Taking into account that thermal memory of the Baltic Sea is only 2 to 3 months [6] and there is no correlation between consecutive ice seasons, a situation when a mild winter occurs directly after extremely severe one (or vice versa) is quite common.

Observations of ice covered area and ice volume are essential for understanding changes in sea ice mass balance, interactions between the ice, ocean and atmosphere. Reliable projections of sea ice response in a warming climate is important also due to the fact that ice extent and mean winter air temperature in Northern and Central Europe are highly correlated [2] which in terms of a mass–related severity index varies with the North Atlantic Oscillation (NAO) [1]. Multiple analysis of the long–term Baltic Sea ice observations report a significant decreasing trend of MIB for the past 100 years of ~2% per 10 years making mild ice seasons more common [4,7-8]. The MIB decrease appears to have accelerated since the 1980s but the large interannual variability prohibits a clear assessment as to whether this increase is statistically significant [4].

There is a great number of numerical models that have been used for the Baltic Sea over recent years to clarify different interactions between sea, sea ice and atmosphere [9-15]. Results from several models were also used to provide a regular sea–ice forecast to support the intense ship traffic on the Baltic Sea [14] and understand winter marine traffic and analyze winter ship navigation accidents [15]. Sea ice constitutes a natural barrier influencing the exchange of heat and nutrients as well as energy transfer between the ocean and the atmosphere. Therefore, numerical ocean–ice models have also been used to study and understand how changes in the climate system would impact the state of the Baltic Sea [16, 17], how ice cover dynamics affect biogeochemistry [18], or impact Baltic inhabitants such as ringed seals [19]. As the need to model sea–ice processes have been highlighted above, the aim of this article is to present and validate the coupled ocean–ice component of the 3D Coupled Ecosystem Model of The Baltic Sea (3D CEMBS) that has been implemented and developed in the Institute of Oceanology at the Polish Academy of Sciences. This is done by evaluating model results from the six winter seasons 2010/2011 to 2015/2016 compared with observations from monitoring stations, ice charts and satellite data.

MATERIALS AND METHODS

3D CEMBS MODEL

A coupled ice–ocean model has been used to calculate hydrodynamic conditions of the Baltic Sea area for the 2010–2016 period. This model origins from Community Climate System Model/Community Earth System Model (CCSM/CESM) coupled global climate model but has been downscaled and adapted for the Baltic Sea domain and further developed at the Institute of Oceanology, Polish Academy of Sciences and called the 3D Coupled Ecosystem Model of the Baltic Sea (3D CEMBS). It is a z–level coordinates, general circulation ocean model that solves the 3–dimensional primitive equations for stratified fluid using the hydrostatic and Boussinesq approximations. The model domain covers the whole area of the Baltic Sea including Kattegat and Skagerrak extended by a part of the North Sea in order to provide a buffer area for open boundary. However, for the purpose of this research, data outside of Kattegat border has been filtered out (Fig. 1a). There are four main source code blocks in 3D CEMBS (called modules or submodels):

- Ocean with Ecosystem – based on Parallel Ocean Program (POP) [20, 21];
- Sea Ice – based on Community Ice CodE (CICE) [22];
- Atmosphere – atmosphere forcing and deposition of nutrients (DATM);
- Land – freshwater inflow and nutrient loads from rivers and large coastal cities (DLND).

Those modules are responsible for the block–related processes simulation or data transfer with prior interpolation onto a domain grid. There is also an additional central block (coupler) responsible for synchronous exchange of information in the model and data stream control (Fig. 1b). 3D CEMBS is configured at approximately 2.3 km horizontal resolution (1/48°). The model grid has 21 irregular layers vertically. The thickness of the first four layers is 5 m and it grows with depth. Detailed model configuration (including
equations appendix) has been presented in separate papers [23, 24]. When water temperature within the grid cell drops below the freezing point (salinity dependency equation) POP calls for CICE, which is responsible for numerical calculations of processes, related with sea ice variables. CICE active configuration within the 3D CEMBS model can be found in [23]. Atmospheric forcing data used in this simulation run origins from Unified Weather Prediction Model (UM) run at the Interdisciplinary Centre for Mathematical and Computational Modelling, University of Warsaw, Poland (www.meteo.pl). 3D CEMBS simulation run has been performed for the period from 1 January 2010 to 31 May 2016 with 2 years of spin–up stage. The reason for such a short simulation is the lack of a reliable long-term in-situ dataset that could be used to make a comparison with 3D CEMBS. In addition, the IMGW database was available only as images. Therefore, it is difficult to compare it quantity-wise with model results. Since ice disappears completely on the Baltic Sea each summer, a longer spin-up period is not mandatory. The output files were recorded with daily frequency. The Cressman data assimilation scheme has been used within this configuration in order to improve overall model accuracy [25, 26]. Satellite–measured sea surface temperature (SST) values from the Moderate Resolution Imaging Spectroradiometer (MODIS, Aqua satellite) taken from SatBaltyk Database [27–29] were used for this process. To analyse the differences in ice formation on a spatial scale, model domain has been divided onto six regions shown on Fig. 1.

**FMI DATA**

Baltic Sea ice concentration database from Finnish Meteorological Institute (FMI) [30] has been used in this paper for comparison with 3D CEMBS model results. Files were downloaded from Copernicus Marine Environment Monitoring Service. Available FMI ice concentration data is based on ice charts produced on a daily basis during the Baltic Sea ice season and show the ice concentration in a 1 km grid. Sea ice service at FMI produces sea ice parameters based on a manual interpretation of satellite data and ground truth. The satellite data used are Synthetic Aperture Radar data from RADARSAT–2 and visual and infrared data from MODIS and National Oceanic and Atmospheric Administration (NOAA). Ground truth origin from Finnish and Swedish icebreakers, ice observation stations of the Baltic ice services and ports. The RADARSAT–2 data are in ScanSAR Wide mode dual polarization and each scene covers an area of about 500 km² and is resampled to a spatial resolution of 100 m. The scenes are mainly focused to the Baltic Sea, Kattegat and Skagerrak east of 9°E. The Envisat data were in Wide swath mode with a swath width of 450 km and were resampled to a spatial resolution of 150 m. The data covered the same area as RADARSAT–2 data. The MODIS and NOAA data covers the charting area several times each day and are resampled to 500 m².

**OSTIA DATA**

The Operational Sea Surface Temperature and Ice Analysis (OSTIA) system runs operationally at the UK Met Office since November 2006 [31]. Output is a daily global coverage combined SST and sea ice concentration product based on measurements from several satellite and in-situ SST data sets. OSTIA uses SST data in the common format developed by GHR SST and makes use of the uncertainty estimates and auxiliary fields as part of the quality control and analysis procedure. Satellite derived sea ice products from the EUMETSAT provide sea ice concentration data to the analysis system. After quality control of the SST observations, a bias correction is performed. To provide the final SST analysis, a multi-scale optimal interpolation is performed using the previous analysis with a slight relaxation to climatology as the basis for a first guess field. Global daily analyses of foundation SST together with sea ice concentration and analysis error estimates are produced on a 1/20° horizontal resolution grid.

**IMGW DATA**

The Office of Hydrological Forecast is the Maritime Branch of the Institute of Meteorology and Water Management (IMGW) in Gdynia, Poland. The ice charts of the Baltic Sea ice conditions are produced within the office in the forms of graphic illustrations of information provided in the Ice Bulletin. The chart is issued twice a week in regular or mild winters and on daily basis in severe winters. The type of ice, its distribution and concentration as well as the processes of ice decay are presented by the international codes and symbols, according to the terminology and ice symbol system of the WMO (hummocks, crevices, ice ridges and packs). The chart provides additional information on ice thickness, distribution and operating icebreakers. For the purposes of this paper, all of the released ice charts from the investigated years have been shared on the basis of a science agreement.

**RESULTS**

**ICE CONCENTRATION**

Ice concentration snapshots are compared in Figure 2 for all investigated winter seasons and available datasets. 3D CEMBS model results as well as FMI & OSTIA charts are shown as a percentage while IMGW Ice Charts are presented without making any changes, using World Meteorological Organisation Concentration Colour Code standards.

POLISH MARITIME RESEARCH, No 3/2018

37
A snapshot is taken on the day of the maximum ice extent (MIB). Detailed values of observed and simulated MIB's with the dates of occurrence as well as model errors are presented in Table 1. 3D CEMBS model results tend to have lower ice extent values than observations. The biggest difference can be seen during the most severe winter of 2010/2011 when MIB of 343,000 km² recorded by OSTIA was almost two times higher than the 3D CEMBS with 183,000 km². However, the day of MIB occurrence was the same for all datasets and occurred exactly on February 25th.

On a daily scale, extent values can have fairly large variations, both due to real changes in ice cover from growth, melt, or from motion of the ice edge, and due to ephemeral weather and surface effects. Figure 3 shows the annual cycle for 3D CEMBS and OSTIA results of ice extent. Both plots show that ice on the Baltic Sea starts to form usually somewhere at the end of December while melting completely in the middle of May. This is common for both severe and mild winters differing only in the size of the ice cover. MIB usually occur somewhere between the middle of January and March.

So far, focus was put mostly on the temporal scale of ice formation by showing either the conditions during the most extreme day in terms of extent per season (Fig. 2) or total ice extent area evolution without exploring its spatial distribution (Fig. 3). To get the full picture and distinguish potential high-risk zones for human activity (for example marine transportation or development) it is important to include the regional character of ice formation and see which areas are commonly ice-covered for the longest period, in what parts of the region as well as where and when the freezing/melting processes starts. To help achieve that goal, 3D CEMBS model simulated maps of ice days (Fig. 4) are presented. Each grid cell has a corresponding number of days when ice concentration was above 15%.

Tab. 1. Observed and simulated maximum ice extent (MIB) with the date of occurrence (Δ = model error)

<table>
<thead>
<tr>
<th>Winter season</th>
<th>Maximum ice extent (10⁹ m²)</th>
<th>Date</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>3D CEMBS</td>
<td>FMI</td>
</tr>
<tr>
<td>2011/2012</td>
<td>126</td>
<td>179</td>
</tr>
<tr>
<td>2012/2013</td>
<td>124</td>
<td>177</td>
</tr>
<tr>
<td>2013/2014</td>
<td>90</td>
<td>100</td>
</tr>
<tr>
<td>2014/2015</td>
<td>40</td>
<td>51</td>
</tr>
<tr>
<td>2015/2016</td>
<td>94</td>
<td>110</td>
</tr>
</tbody>
</table>
In order to connect both spatial and temporal aspects of ice evolution, the percentage of region covered by ice (Fig. 5) was calculated for six Baltic Sea regions according to the division in Fig. 1.

It is worth noticing, that even during mild winters northernmost areas of Bothnian Bay are covered by ice at least for a hundred days and ice covers eastern part of Gulf of Finland around the mouth of Neva River for at least a month or two. Also, Danish Straits & Kattegat as well as Baltic Proper are pretty much ice free most of the time with the exception of the coastal area.

**DETAILED ICE EVOLUTION**

**Winter 2010/2011**

Winter 2010/2011 was severe with the highest values for ice area covered since the extremely severe winter of 1986/1987. The end of November and December were exceptionally cold and the amount of sea ice began to increase. As mid–December approached, there was ice in all Finnish coastal areas. Frosty weather continued at the beginning of 2011 and the ice–covered area expanded to more than 100 000 km² (Fig. 3) at the beginning of January. At that time, ice covered the entire Bothnian Bay and the Quark as well as the Archipelago Sea. The latter half of February turned out to be exceptionally cold causing the amount of sea ice to increase rapidly, and the peak of the ice winter was reached on February 25th (Tab. 1).

**Winter 2011/2012**

In statistics based on MIB, the winter of 2011/2012 was average and shorter than usually, as it started exceptionally late, and the last pieces of ice disappeared earlier than average. Wintery weather had set in late January, and the cold conditions continued in February which caused the sea ice cover to expand, and the ice extent reached its peak somewhere between 5th and 12th February (depending on data source – see Table 1).

**Winter 2012/2013**

The Baltic Sea ice season 2012/2013 was average but the turning point of the winter was late. Ice formation in the innermost bays of the northern part of the Bothnian Bay began with the end of November. At the beginning of March cold arctic air started to flow to Scandinavia and the extent of ice began to grow reaching its maximum somewhere between 13th and 16th March (Tab. 1). From then on, the cold nights formed new ice but sunny days melted them and the extent of ice did not enlarge any more. The last ice melted from the southern Bothnian Bay and Gulf of Finland at the beginning of May. The Bothnian Bay was ice–free on the first days of June (Fig. 5 – Bothnian Bay section).

**Winter 2013/2014**

Season 2013/2014 was mild. The ice formation in the innermost bays of the northern part of the Bothnian Bay began with the end of November. Since January was exceptionally cold, the extent of ice reached its maximum at the beginning of February. After this, the weather became milder, and the southern winds pushed the ice fields together. The rest of February was unusually mild. In the beginning of March the ice extent was below 50 000 km². April was warmer than average and at the end of the month ice appeared only in the Bothnian Bay. Finally, The Baltic Sea became ice–free nearly two weeks earlier than usual having the last ice melted in the Bothnian Bay around May 15.

**Fig. 4. Number of ice days (ice conc. > 15%) simulated by 3D CEMBS model. From left: 2010/2011, 2011/2012, 2012/2013 (top) and 2013/2014, 2014/2015, 2015/2016 (bottom)**

**Fig. 5. Ice covered region percentage for all winters in Bothnian Bay (a), Bothnian Sea (b), Gulf of Finland (c), Gulf of Riga (d), Baltic Proper (e), Danish Straits and Kattegat (f). 5-day trailing mean has been used**
**Winter 2014/2015**

This ice season was exceptionally mild. The ice formation in this mild season began in the middle of November in the northern part of the Bothnian Bay. December and the beginning of January were milder than usual with only a short cold period till the last days of January resulting in the culmination ice extent around January 24th. March and April were warm which left the remaining ice formation only in the Bothnian Bay. The last ice melted in the first half of May. The Baltic Sea was ice-free two weeks earlier than usual.

**Winter 2015/2016**

The ice winter 2015/2016 was mild and considerably shorter than average. Around the turn of the year, the weather cooled considerably and ice started to form. However, January began with cold temperatures and cold weather continued for about three weeks. This resulted in the maximum ice extent forming at January 22nd (Tab. 1). Even though February and March were warmer than the long-term average, the extent of the ice cover started to decrease in the last week of March. The Gulf of Finland became free of ice in the middle of April and the warm weather at the beginning of May melted the remaining ice fairly quick making The Baltic Sea completely ice free around the middle of May.

**DISCUSSION**

Economic growth within the Baltic Sea region corresponds directly with a 34% marine traffic increase in the last 10 years (the number of icebreakers however has not increased). Therefore, the lack of high-resolution ice condition daily information and forecast can lead to serious economic losses [32, 33]. Using numerical modelling in addition to standard monitoring techniques (remote sensing and direct observations) brings a lot of benefits giving the ability to provide complex information of sea ice conditions with very high spatial and temporal resolution, which has become of vital importance in human marine activity.

Changes in the sea ice can have multiple effects in sea level rise, ecosystem, etc. Some species that use sea ice cover as a natural breeding area, for example seals (*Halichoerus grypus*) may have to migrate to the northern region of Bothnian Bay to survive [19].

While some northern populations are declining or migrating to new areas, there are evidences of multiple species that extend their breeding habitats becoming invasive in some areas [34].

Also, the natural hazards and risk for marine infrastructure can occur. At the same time, one could see rising opportunities for new navigation routes and exploration of areas that are currently restricted or limited to only a few months a year.

Accuracy of the satellite observation techniques and methods is heavily constrained by many limiting factors like (among others) weather, attached instruments or even algorithms used to process data. Numerical models on the other hand, (while not forgetting about their limitations) enable simultaneous projection of many sea ice cover parameters (even at low values) with a great temporal and spatial resolution including ice thickness.

Winter of 2010/2011 was the most extreme since 1986/1987 and it is not very likely that these events will happen often in the upcoming years especially in the climate change perspective. Therefore, good simulation of less severe winters proves that 3D CEMBS is a viable tool for ice conditions simulations and can be used for IPCCs scenario induced long-term runs. Since climate changes induce ocean’s temperature to rise, lots of the changes (including ice melting) happen in the water column and are invisible for instruments (e.g. satellites) that see and measure only the surface layer. Therefore, numerical simulations have an advantage in this field since they enable to predict the real volume changes and rate (speed) of investigated processes.

**CONCLUSIONS**

The analysed period 2010/2011 to 2016/2016 can be characterized as very dynamic in terms of interannual variability of ice parameters. The correspondence between 3D CEMBS model results and observations in terms of both temporal and spatial analysis is encouraging. The most severe winter season was 2010/2011 followed by mild winters 2013/2014–2015/2016. There are some differences between the model and FMI/OSTIA/IMGW data, especially during the severe winter 2010/2011. Even though the day of MIB in model (25th Feb 2011) was the same as the observation date, the simulated value of ice extent area was 39% lower in comparison with FMI (up to 45% lower for OSTIA). For other winters simulated values were usually only around 10–15% lower than observation.

The authors believe that the main reason for this underestimation is the 5 m surface layer thickness of the ocean model. To cool down such thick layer, so it’s temperature drops below the freezing point requires more time. In reality, fast ice starts to form in a thin surface layer. This should be solved by a model assimilation. However, there are periods of time, especially during winter, when satellite images are not available due to high cloudiness.

It is uncertain what might be the other potential reason for this underestimation, since there are many factors influencing it. Ice concentrations from satellite imaging are also sensitive to the algorithm used, and the resulting numbers for extent depend not only on algorithms but on other processing steps as well. Also, it is important to note that the extent values have uncertain significance when taken individually [34]. It is clearly visible that even though the model tends to slightly underestimate ice conditions it reacts very well to atmosphere forcing what can be confirmed by only minor differences in the dates for maximum extent at each winter season.
**ACKNOWLEDGEMENTS**

Partial support for this study was provided by the project "Integrated info-prediction Web Service WaterPUCK" – no. BIOSTRATEG3/343927/3/NCBR/2017 and “Knowledge transfer platform FindFISH” – no. RPPM.01.01.01-22-0025/16-00. The authors are grateful to the anonymous reviewers for valuable comments on earlier versions of the manuscript.

This study has been conducted using E.U. Copernicus Marine Service Information and Satellite Monitoring of the Baltic Sea Environment – SatBaltyk database.

Calculations were carried out at the Academic Computer Centre in Gdańsk.

**REFERENCES**


CONTACT WITH THE AUTHOR

Maciej Janecki
e-mail: mjanecki@iopan.gda.pl
Institute of Oceanology, Polish Academy of Sciences
Powstańców Warszawy 55, 81–712 Sopot
POLAND

Artur Nowicki
e-mail: anowicki@iopan.gda.pl
Institute of Oceanology, Polish Academy of Sciences
Powstańców Warszawy 55, 81–712 Sopot
POLAND

Alicja Kańska
e-mail: alicja.kanska@imgw.pl
Institute of Meteorology and Water Management
National Research Institute,
Maritime Branch, Hydrological Forecasting Office
Jerzego Waszyngtona 42, 81–342 Gdynia
POLAND

Maria Golenko
e-mail: m.golenko@yahoo.com
P.P. Shirshov Institute of Oceanology, Atlantic Branch
Prospect Mira 1, 236000 Kaliningrad
RUSSIA

Lidia Dzierzbicka-Głowacka
e-mail: dzierz@iopan.gda.pl
Institute of Oceanology, Polish Academy of Sciences
Powstańców Warszawy 55, 81–712 Sopot
POLAND
AN ANALYSIS OF BASIC PARAMETERS OF RO-PAX FERRIES IN THE BALTIC SEA AS GUIDELINES FOR ITS PRELIMINARY DESIGN

Lucjan Gucma
Julia Raczkowska
Maritime University of Szczecin, Poland

ABSTRACT

This paper presents results of an analysis of the operating and maneuvering parameters (length, width, draft, main propulsion power and maneuvering equipment, and windage area) of the Ro-Pax ferries over 150m in length. The analysis takes into accounts all ferries operating in the Baltic Sea in 2017. The work also includes route analysis of the collected Ro-Pax units. The influence of port infrastructure is also discussed. The mean optimal Baltic Sea Ro-Pax ferry and its parameters for short term and long term forecasts are presented. This paper may be used as a primary guideline for the determining the optimal Ro-Pax ferry dimensions size for the Baltic lines in terms of preliminary ship design.

Keywords: Ro-Pax ferry, ship design, ferry traffic, operating parameters, maneuvering parameters, analysis of ship dimensions

INTRODUCTION

Ship design is a complex process, consisting of several subsequent stages, compliant with Evans’s rule of design helix, and elaborated by Andrews in terms of helix loops. (Charchalis & Krefft, 2009) The essence of this design approach is illustrated in the Fig. 1. The expected solution is reached by verifying design decisions made at the preliminary design phase in the progressive course of design process. Traditionally, ship design may be considered in the four main phases process, namely:

1. Concept design – Feasibility study;
2. Preliminary design;
3. Contract design; and

Early stages of each modern ship design are concept design — feasibility study and preliminary design, also known as basic design (Papanikolaou, 2014). The base for ship design are very precisely defined in client’s requirements. They are prepared in form of technical and economical features of future ship. Next, client’s assumptions are checked for compliance with restrictions on the length, beam and draft in terms of the passage of canals and channels, ship’s approach to river estuaries, and calling at certain ports. The preliminary ship design encompasses the following more detailed objectives: Selection of main ship dimensions; Development of the ship’s hull form; Specification of main machinery and propulsion system type and size; Estimation of auxiliary machinery type and powering; Design of general arrangement of main and auxiliary spaces (cargo spaces, machinery, spaces and accommodation); Specification of cargo-handling equipment; Design of main structural elements for longitudinal and transverse strength; Control of floatability, stability, trim and freeboard, and Tonnage measurement (Papanikolaou, 2014).

The concept phase, which generally is an iteration process, plays a critical role in the final project results. This process requires the repetition of a single analysis procedure, or sometimes even all cycles, several times. The empirical equations form the base for analytical methods. They are usually determined using regressive analysis of collected data on similar, significant ships. The mathematical relationships between different dimensions, parameters and features
of specified type of vessel are shown in these equations. (Charchalis, 2013) The verification based on contemporary significant ship database is the key to achieve the reliable accuracy of analytical methods. Iteration methods require application of data coming from the significant ships database. The obtained relationships, which comply with the client’s ship assumptions and main dimensions coming from similar ship database, are used to determine general dimensions of the designed ship, and are the basis for further optimization studies. This stage gives fundamental data on ship design, like her main dimensions, hull form or powering, which have strong influence on further ship design and project cost, and consequently construction and exploitation costs. Any changes of ship data specified at this phase made at later stages can entrain great expenses for shipowner.

The contract design phase aims to prepare all necessary calculations and naval architectural drawings, as well as the drawing up of the technical specifications of the designed ship. The documents are the base for the formal shipbuilding contract between the ship owner and the shipyard.

The most significant benefit of well organized ship database infrastructure and proper software is that, the feasibility–concept design may be accomplished in 1 day or even less. (Papanikolaou, 2014)

This paper and the information held within it could be used as a method for preparing a primarily analysis for further studies aimed at designing new Ro–Pax vessels and predicting the trends of ship designing development.

### SELECTION OF FERRY’S MAIN DIMENSIONS

The design of ships dimensions is always a compromise between ship owner’s requirements, legal and technical restrictions and waterway and infrastructure limitations. When determining the main dimensions of ship, beam and draft in particular, topological limits of the route acts as a boundary line, often dictating the result value.

There are six dimensional relationships linking the four main ship dimensions of length, beam, depth and draft. To solve either the weight or volume equations three of these have to be used. The relationships are (Bertram, 2000):

1. \( B = f(L) \)
2. \( D = f(L) \)
3. \( T = f(B) \)

where:
- \( B \) – beam [m];
- \( T \) – draft [m];
- \( L \) - length [m]; and
- \( D \) – depth [m].

For the economy of construction, a ship should be perceived as a container. As the straight-side container which has the least surface area for a given volume is a cube, therefore, the vessel, according to this approach, should be as closely as possible a cube.
possible to this shape. This approach, according to (Watson, 1998), requires that draft (the smallest of the dimensions) should be the maximum permitted by length, breadth and depth; that depth (the next smallest dimension) should be the maximum permitted by length and breadth; that breadth should be the maximum permitted by length and finally that the block coefficient \( C_B \) should be as full as possible.

Designing a new vessel usually starts with determining the ship’s length. The length is usually specified based on similar ship data derived from a ship design database. In the case of deadweight carriers, next the displacement is estimated. After the preliminary estimation of the ship’s length as well as the block coefficient \( C_B \), the ship’s width in relation to the draft is determined. These dimensions are directly connected with each other through following formula:

\[
B \cdot T = \frac{V}{(L \cdot C_B)} \tag{4}
\]

where:

- \( B \) – beam [m];
- \( T \) – draft [m];
- \( L \) - length [m]; and
- \( C_B \) – block coefficient [-].

If neither restricted nor limited, the choice of proper width values usually results from the stability demands. Draft, which must correspond to the formula 4, is chosen in relation to the width such that the desired degree of stability results. For ships with restricted dimensions (particularly draught), the width required for stability is often exceeded. (Schneekluth & Bertram, 1998) Correlation between coefficients \( L/B \) and \( B/T \) are essential to define the impact of specified dimensions in ships. Coefficient \( B/T \) effects transverse stability of the ship and her hydrodynamic performance. Increasing of breadth has a positive influence on stability. However, it also increases resistance of a hull and requires higher propulsion power. \( L/B \) has impact on hull resistance and therefore its power. In terms of ship stability, the \( L/B \) ratio is less significant than the \( B/T \) ratio.

**BASIC DESIGN DATA OF RO-PAX FERRIES WITH CONSIDERATION OF PORT INFRASTRUCTURE**

The main design parameters of ferries are discussed below and their importance in terms of matching the ferry to the existing port infrastructure. The basic ferry operational parameters influencing predesign are (Gucma S., 2012):

1. total length \( (L_{oa}) \),
2. breadth of the ferry \( (B) \),
3. draft of the ferry \( (T) \),
4. transverse windage area \( (F_w) \),
5. power of the main propulsion \( (P) \),
6. power of side thrusters \( (P_s) \),
7. type of propulsion and steering devices.

The total length \( (L_{oa}) \) is the most important parameter of the Ro-Pax ferries, because it influences the loading capacity (length of the load line). From the navigational and operational point of view, the length of the ferry determines the necessary length of the mooring line of the quay. In most cases, the \( L_{oa} \) of the ferry should be shorter than the existing length of the quay, so that it is possible to provide proper angle of bow or stern lines. The length of the ferry also determines the minimum diameter of the turning place in the port. The minimum diameter of the turning place should be 1.5 \( L_{oa} \), however in special cases, the ferry is able to turn within a smaller area, i.e. up to 1.2 \( L_{oa} \), but this should be confirmed by individual tests usually carried out using simulation methods. Determining the minimum maneuvering area in this case involves the necessity to carry out real-time maneuvering simulation tests under varied conditions.

The ship’s breadth \( (B) \) is a parameter, which for Ro-Pax ferry is not very variable when changing the length of the ferry. It influences mainly on the speed characteristics of the ferry (appropriate slenderness) and, above all, on the number of loading lines per car deck. Typically, the width of a single load line lane for heavy goods vehicles is set at 3.1m – 3.6m. Narrowing the car lines increases the total length of the load line, but may cause difficulties in loading (increased loading time) and more difficult access in case of fire. In terms of the maneuvering area for the ferry, the width determines the minimum width of the approach waterways to the terminal. The minimum width of the waterway is defined as the multiplicity of the unit’s width and is taken as a value from 2 \( B \) (or sometimes even less) to even 4 \( B \) depending on the navigational conditions and the type of the water body. The average width of large Ro-Pax ferry is around \( B = 28 \)m. Such widths in the analyzed ports (Świnoujście, Ystad, Trelleborg) do not affect the restriction in this respect, therefore it can be considered as irrelevant in terms of navigation.

Draft of the ferry \( (T) \) is assumed that the safe depth of the basin should allow the ferry to enter with adequate reserve for water levels decrease. Significant squat of the ferry should be taken into account, because in difficult conditions, to keep the right momentum of the maneuver captain should maintain a significant speed (even up to 7 m) when entering the port. Ro-Pax ferries, however, due to the spatial cargo have relatively low draft compared to other types of ships of similar length. These values reach up to approximately \( T = 6.5 \)m and do not cause significant navigation restrictions at medium water levels in the majority of Baltic Sea ports.

The transverse windage area \( (F_w) \) is the result of providing adequate dimensions of the ferry and struggling to increase the space for the spacious ro-ro cargo and passengers, which highly increases the windage area. Properly selected parameters of the main and auxiliary propulsion enable, in turn, keeping the maneuverability of the ferry under given operating conditions.

The total power main propulsion \( (P) \) is selected to ensure the right speed characteristics. The power of propulsion devices also determines with the power of additional steering gears the possibility of independent maneuvering, including mainly overcoming by the ferry of aerodynamic drag, which is
related to the windage area of the ferry. Power of side thrusters ($P_s$). In most large Ro-Pax units, in order to generate forces and moments in the ship forebody (or balancing the moments in the stern section), a set of bow thrusters on (1 to 3) or stern thrusters (usually maximum 1 stern rudder) is used for maneuvers without using the tugs.

The following types of propulsion and steering devices are most frequently used on Baltic Ro-Pax ferries are:

1. classical: aft – two controllable pitch propellers (CPP), two unconventional rudders placed behind the propellers (usually active Becker style), optional stern thrusters; bow – bow thruster or thrusters (approximately on over 90% of ferries);
2. azimuth: aft – 2 azimuth propulsors (both gondola type and without gondola) placed on the extreme sides of ferry (less than 10% of ferries), bow – bow thrusters;
3. azimuth propulsors used on two-sided ferries (one or two on the bow and same on stern);
4. others, such as the VS drive, very rare on ferries.

To determine the operational parameters for ferries an individual method is used, which consists in analyzing a given ferry in relation to a given position and determining extreme conditions (wind force, current speed and direction etc.).

**MAIN ROUTES AND FERRY TRAFFIC IN THE BALTIC SEA**

As the question of the actual borders of the Baltic Sea Region remains open, for the purpose of this research, we include ten countries featuring their coastlines along both the Baltic Sea and Skagerrak Strait. These are (in alphabetical order): Denmark, Estonia, Finland, Germany, Latvia, Lithuania, Norway, Poland, Russia and Sweden. Based on given assumptions, the data were collected from two databases (http://matkustajalaivat.com, www.ferry-site.dk). The total of 62 ferries, operated by 16 ferry companies, were listed and processed statistically. The Fig. 2. illustrates the analysis of number of the operators’ fleet.

Baltic ferry operators currently serve 29 independent lines. The vast majority of the routes (97%) are international connections. The only domestic route is the connection between Swedish island Gotland and the port of Nynäshamn on the mainland. As a matter of fact, there are far more domestic lines connecting Swedish, Danish, Norwegian islands with their respective mainland’s. However, they are served by small local ferries operators with fleet usually under 150m in length. The Fig. 3. gives an overview of the Ro–Pax ferry connections in the Baltic Sea as of 2015.

According to collected data, Swedish ports handle nearly 71% of all Ro-Pax ferries’ Baltic voyages, what makes them the busiest in the BSR. This is not surprising, as Sweden has been the leader among the BSR countries in terms of the number of ferry connections with other countries of the Baltic Sea for many years. (Uriasz, 2010; Wiskulski, T. & Bar-Kołelis D., 2012) Sweden has developed ferry connections with nearly all Baltic countries, except for Norway and Denmark, thanks to a great number of ferry ports (11 ports listed). The most used Swedish routes lead to Germany, Finland and Poland (5, 4 and 4 routes respectively). Other rather traffic-dense BSR are Danish Straits and Kiel Canal. Denmark and Norway are connected by five different routes. However, these lines are often serviced by only one ferry.

The busiest international routes are Helsinki – Tallinn and Trelleborg – Rostock links. Both connections are serviced by 6 different ferries of 4 and 2 different operators respectively. Other very popular routes connect Stockholm with Finnish cities – Helsinki and Åbo. 4 different ferries of 2 different
ferry companies operate on these lines. The only connection with Russian port is St. Petersburg – Helsinki – Stockholm – Tallinn route served by a cruise Ro-Pax ferry of St. Peter Line Company. A little more than 40% of all BSR connections is serve by only one Ro-Pax ferry. This concerns mostly traffic between Sweden and Baltic States, and Denmark and Germany.

INFLUENCE OF PORT INFRASTRUCTURE ON PRELIMINARY FERRY PARAMETERS

Very specific in Ro-Pax ferry trade is strong influence of the existing and planned port infrastructure on design parameters. The major parameter here is length of the ferry quay \( L \) which will affect possibility of mooring of given ships. Another aspect is the turning possibility that is obviously the function also of ferry length (Gucma L., 2010). Table 2 presents of existing and planned parameters of three ports for operation of Ro-Pax ferry considered in this study.

<table>
<thead>
<tr>
<th>Port</th>
<th>Świnoujście</th>
<th>Ystad</th>
<th>Trelleborg</th>
</tr>
</thead>
<tbody>
<tr>
<td>Existing quays of ( 150m &lt; L &lt; 200m )</td>
<td>4</td>
<td>1</td>
<td>6</td>
</tr>
<tr>
<td>Existing quays of ( 200m &lt; L &lt; 230m )</td>
<td>1</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>Existing quays of ( L &gt; 230m )</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Planned quays of ( L &gt; 230m ) (within 3 years)</td>
<td>1</td>
<td>3</td>
<td>3</td>
</tr>
</tbody>
</table>

ANALYSIS OF SELECTED PARAMETERS OF RO-PAX FERRIES CURRENTLY OPERATING IN THE BALTIC SEA

The sample size covered by Baltic Sea operating ferry parameters. The data was collected in the beginning of 2017. Sample covers 62 ferries operated on the Baltic Sea at the beginning of 2017. Only ferries with a length of more than 150m were considered. It should be noted that ferry lines are often handled by sister units. Consequently, the sample size is reduced. The presented ferry parameters analysis allows to determine the average parameters of ferry currently operating, the ferry development trends in the Baltic Sea and indispensable power for equipment for controlling ship’s movements depending on the ferry’s dimensions. The following parameters of the vessels were analyzed:

- building year – \( b/y \),
- length overall – \( L_{oa} \) [m],
- bream – \( B \) [m],
- draft – \( T \) [m],
- main propulsion power - \( P \) [kW],
- thunders power \( P_{ss} \) [kW],
- transverse windage area \( F_{ny} \) [m²].

Table 2 presents selected quantitative statistical parameters of the examined data in the form of measurements of central values, variations, symmetry and concentration.

From the data in Table 2 we can see the mean values of each, examined vessel parameters. The data indicate that the average Baltic Ro-Pax ferry is 19 years old, 188m long, 28m wide and 6.4m deep. Her main propulsion power is 27,100kW, her stern and bow thunders power is 3,000 kW and her topside projected area is 4,600m².

The standard deviation values show relatively small statistical dispersion from mean value of all analyzed operating parameters. A similar situation occurs in the case of the kurtosis values, which suggest slight skewness of analyzed parameters’ values. Merely the kurtosis values for vessel main power indicate significant differences between the normal distribution and the kurtosis estimator, which results from a large number of vessels with an engine of the power ranging from 20,000kW to 30,000 kW.

VESSEL’S YEAR OF BUILD

It should be noted that the mean age of Ro-Pax ferries currently operating in the Baltic Sea is around 18 years. The detailed frequency distribution of ferries in the form of a histogram is shown in Fig. 4. An analysis of data shows that a significant number of ferries was built between 1985 and 1995 (22 vessels). These are vessels that need replacing with newer ones. It can be seen that in 2005-2010 a considerable number of new ferries (16 vessels) were placed into service.

The changing dynamics of length overall in relation to the building year of ferries in the Baltic Sea (Fig. 5) shows no significant changes over time. The average ferry length grow is only 0.86m per year. This is caused by the market demand and the slow expansion of port infrastructure, which has an
average lifespan in excess of 25 years. At that time of lifespan, the average length of the ferries in the Baltic Sea has increased by only 21m. In the Fig.5, one can also notice a group of several new ferries, built between 2005 and 2010, whose average length is approximately 200m if not slightly less.

Fig. 5. Changing dynamics of length overall (Loa) of Baltic ferries in relation to their building year (as of 2017)

VESSELS' LENGTH

The length overall (Loa) of the ferries is shown as the frequency distribution in Fig. 6. It can be seen that a significant part of the units are from 170m to 200m (39%). Ferries of 200m to 220m make up only 16%. Ferries over 230m are a rarity. This is mainly due to the fact, that port infrastructure is not properly prepared to receive such ships.

Fig. 6. Analysis of length overall (Loa) of Baltic Ferries (as of 2017)

VESSELS' BREADTH

A histogram of the vessel’s breadth is presented in Fig. 7. The average breadth of ferries with length overall over 150m is $B = 28$m and the ferries with width ranging from 27.5m to 29.0m make up the majority (42%). Noteworthy, ferries rarely exceed 30m (11%).

Fig. 7. Analysis of width (B) of Baltic Ferries (as of 2017)

Though the analysis of the ferries' breadth (Fig. 7) indicates a small effect of the length of ferries on their breadth, all of the ships are within the coefficient $L/B$ range for fast and passenger vessels, which value ranges between 6.5 and 8.5.

Fig. 8. Analysis of width (B) as a function of length overall (Loa) of Baltic ferry (as of 2017)

VESSELS' DRAFT

Analysis, in a form of histogram, of draft ($T$) of the ferries is shown in the Fig. 9. It indicates that the draft of significant part of the vessels ranges from 6.2m to 6.9m (66%). The average draft of Baltic Ro-Pax is about 6.4m and rarely exceeds 6.9m (11%).

Fig. 9. Analysis of draft (T) of Baltic Ferries (as of 2017)
The $L/T$ ratio is essentially a secondary relationship resulting from formula in 4. Analyzing ferry’s draft one comes to similar conclusions as in the case of the width (Figure 8). Vessel’s draft does not depend significantly on its length overall.

VESSEL’S PARAMETERS INFLUENCING ITS MANEUVERING POSSIBILITY

The transverse windage area of the ferry ($F_{yn}$) is a basic parameter that determines the vessel’s maneuverability and power requirements for self-maneuvering. There is a strong correlation between a topside projected area and length overall of the ferries. This phenomenon was proven in the Fig. 11.

Figure 12 presents the dependence of the total power of the main propulsion ($P$) on the transverse topside projected area ($F_{yn}$) of the Baltic ferries. It shows a strong correlation between these parameters, not including fast ferries, which were designed for speeds higher than average ($V_{n} = 22\text{kn}$).

Transverse windage area of the ferry ($F_{yn}$) of the ferries is shown as the frequency distribution in the Fig. 13. It can be seen that a significant part of the units range from 4,000$m^2$ to 5,000$m^2$. The histogram slightly corresponds with an analysis of ferries’ length overall ($L_{oa}$), as there is a correlation between these two parameters.

An analysis main propulsion power ($P$) of Baltic Ferries is illustrated in the Fig. 14, including a bar chart. An analysis of data in figure above shows that the most frequently present in the BSR vessels were those propelled by an engine with power ranging from 20,000$kW$ to 30,000 $kW$, as they account for nearly 55% of all vessels.
The dependence of the total power of bow and stern thunders \((P_s)\) from length overall \((L_{oe})\) is shown in the Fig. 15. It indicates the correlation between a thunder power and a ferry’s length. This is understandable, as they are designed to balance wind torque, which increases with a vessel’s length.

By assuming that the dynamics of infrastructure development will not change significantly, it will be able to be operated for about 25 years. Such a ferry will be able to use the existing infrastructure in the Baltic Sea during that time and can be easily moved over Baltic Sea lines.

As a maximum ship – the future one, which will be able to be served only in some ports and will be slowly coming into operation over next 10 years is Ro-Pax with the parameters:

**Name:** Ro-Pax300m  
\(L_{oe} = 230m\)  
\(B = 30m\)  
\(T = 6.7m\)  
\(F_w = 5,800\ m^2\)  
\(P = 33,000\ kW\)  
\(P_s = 6,000\ kW\)

**CRITICAL WIND SPEED FOR FERRY OPERATION**

Critical wind speed \((v_k)\) – the maximum transverse wind speed for safety operation) is a crucial parameter for Ro-Pax ferries determining the no-delay operation. Expert investigations have been made to determine parameters of critical wind speed by ferry captains. The values are subjective but its estimation is averaged by ferry crew in long time operation experience. The Tab. 7 shows these critical wind speeds \((v_k)\) for safe operations of various ferries exploited in the South-West Baltic area. The linear model is created to establish dependence of such an indicator on the critical speed of the ferry (Fig.15) in form:

\[
v_k = f \left( \frac{c_p}{F_{yn}} \right) = f \left( \frac{P + kP_s}{F_{yn}} \right)
\]

where:

- \(v_k\) – critical speed \([m/s]\)
- \(P\) – main propulsion power \([kW]\);
- \(P_s\) – total power of thrusters \([kW]\);
- \(c_p\) – coefficient of power \([-]\);
- \(k\) – empirical coefficient assumed as 12.

The coefficient \(k\) is optimized to achieve highest determination coefficient \((R)\) in linear model presented in the Fig. 17. Similar methodology was used in (Gucma S., 2012).

**DETERMINATION OF FUTURE MEAN FERRY PARAMETERS**

Statistical analysis of selected parameters of ferries currently operating on the Baltic Sea leads to the conclusion that the medium, future-oriented ferry on the Baltic Sea is a ship with the following parameters:

**Name:** Ro-Pax200m  
\(L_{oe} = 200m\)  
\(B = 28m\)  
\(T = 6.5m\)  
\(F_w = 5.000\ m^2\)  
\(P = 27,000\ kW\)  
\(P_s = 4,000\ kW\)
For ferries identified in this study, the critical wind values will therefore be respectively:
1. Ro-Pax200m: 
   \[ c_p = \frac{(P + kP_{ss})}{F_{yn}} = 15.0 \text{ kW/m}^2, \text{ thus } \nu_k = 0.59c_p + 10.2 = 19.0 \text{ m/s}; \]
2. Ro-Pax230m: 
   \[ c_p = \frac{(P + kP_{ss})}{F_{yn}} = 18.1 \text{ kW/m}^2, \text{ thus } \nu_k = 18.1 \text{ m/s}. \]

The adopted parameters of the ferries show their high maneuverability and the possibility of operation all year round in Świnoujście (less than 1 day with wind higher than 20m/s) and almost all year round in Ystad and Trelleborg (abt. 4 days with wind higher than 20m/s).

It should be remembered that there are two types of maneuvers performed by ferries: static and dynamic. The first of these relies on the force and static interaction of propulsions for external extortion (mainly wind). The dynamic maneuver is to keep the ferry moving and to generate fast-changing moments of forces in order to counteract external influences.

**CONCLUSIONS**

The study analyzed the parameters of Ro-Pax ferries currently operated in the Baltic Sea of length overall over 150m. The performed analyses are vital in reference to plans of designing new vessels and modernizing port infrastructure. The trend of ferry development was determined and on its basis the parameters of two optimal ferries were determined for next 25 and 35 years of operation.

The study shows that Ro-Pax ferries on the Baltic Sea are replacing by new units relatively slowly and lot of them are still more than 20 years old.

The analysis showed that major parameter of Ro-Pax ferries is length overall. In detailed analysis the simplified models were presented which connect ferry length with other important parameters like: main engine power, thruster’s power and windage area.

The development trend of the ferries was determined on this basis and the parameters of the future 25 years and 35 years were defined as:
1. The medium-maximum unit (Ro_Pax 200m) is a Ro-Pax ferry with the length \( L_{oa} = 200m \) (other parameters are in the study) that can be maneuvered to the wind with a force of about 19m/s. Such ferry, assuming that the dynamics of infrastructure development will not change significantly, could be operated on the Baltic Sea for about 25 years. Such a unit will be able to use the existing infrastructure in the Baltic Sea during this time and exchange outdated ferries.
2. The maximum-future unit (Ro_Pax 230m) with the total length \( L_{oa} = 230m \) and other parameters specified in the study, which will be able to be operated only in some Baltic Sea ports and will enter slowly into operation in the perspective of several or more years.

The method of critical wind speed determination based on expert (captain’s) opinion was presented. The method was applied to future ferries to find critical wind, assumed for given solution and later on economical effectiveness of ferries (yearly delay time in given route).

The study analyzed also the existing and planned port infrastructure in Świnoujście, Ystad and Trelleborg in terms of current and future capability of operating large ferries (Loa> 200m) stating the following:
- In the current state of infrastructure, it is possible to operate determined Ro-Pax200m ferry in all three analyzed ports, regardless of conditions, but only one quay is available in Świnoujście;
- On the Świnoujście – Trelleborg line it will be possible to operate the Ro_Pax 230m unit in 2020;
- On the Świnoujście – Ystad line it will be possible to operate a Ro_Pax 230m unit after 2020 or later, depending on the pace of modernization works at the Ystad Port which are about to start in 2018.

![Fig. 17. Dependence of critical wind (\( \nu_k \)) of selected ferries from the power coefficient (\( c_p \))]
ACKNOWLEDGMENTS

This research outcome has been achieved under the grant No 1/S/CIRM/16 financed from a subsidy of the Ministry of Science and Higher Education for statutory activities.

REFERENCES


15. www.ferry-site.dk (access January - February 2017)


CONTACT WITH THE AUTHORS

Lucjan Gucma

e-mail: l.gucma@am.szczecin.pl

Maritime University of Szczecin
Waly Chrobrego 1/2
70-500 Szczecin

POLAND

Julia Raczkowska

e-mail: j.raczkowska@am.szczecin.pl

Maritime University of Szczecin
Waly Chrobrego 1/2
70-500 Szczecin

POLAND
BUSINESS PROCESS RE-ENGINEERING OF A MARITIME CAR TERMINAL: CHANGING FROM INBOUND TO OUTBOUND INTERMODAL NODE IN FINISHED VEHICLE LOGISTICS (FVL)

Bojan Beškovnik
Marina Zanne
Faculty of Maritime Studies and Transportation, Pot pomorščakov, Slovenia

ABSTRACT

This article presents a broad perspective on a maritime car terminal's role in finished vehicle logistics with an aim to elaborate the process when a terminal changes its role from an inbound intermodal node to an outbound intermodal platform. Through the analysis of the maritime car terminal in Koper that occupies the 9th position among top European ports in handling light vehicles, a comprehensive view of process re-engineering is provided. The research analyses specific inbound and outbound processes and real-time decisions that influence the terminal's productivity and satisfies stakeholders' business expectations. The proposed business process re-engineering (BPR) model for an outbound car terminal highlights the directions of future management decisions to position the terminal as an export oriented intermodal node. The article offers a theoretical and practical contribution to efficient FVL management.

Keywords: finished vehicle logistics; automotive industry; maritime car terminal; intermodal node; business process re-engineering

INTRODUCTION

The automotive industry is very often exposed to cyclical and turbulent fluctuations. A decrease in GDP development or a global financial crisis has a direct impact on the automotive industry. The final product (car) is a consumable of long-term usage. As a consequence, demand which has a strong influence on production and global supply chains, is lower. Of course, the opposite (positive) direction of demand has global impacts on different supply chains as well, because the time gap between market demand and adjusted production applies pressure on supply elements, production and intermodal terminals.

Intermodal (car) terminals are exposed to both situations, when they have to secure increasing buffers for already produced vehicles, that are hardly sold on the market, and when they have to perform agile and fast services for already paid vehicles.

Car terminals are directly dependent on finished vehicle logistics (FVL) and its strategy. Every car production plant uses personalized logistics chains. Based on a plant’s position, the chains are established and later constantly optimised. FVL chains should be short, lean, and agile, and provide all the necessary services expected by both the automotive industry and the end consumer. Mendonça and Dias [11] state that within FVL a car terminal must secure different operational services. Besides loading, discharging and storage services the terminal is very often requested to provide vehicle inspection in different manipulation stages, pre-delivery inspections (PDI), maintenance, car customization, damage repair and in some countries also dissembling and assembling operations.

All requested services in FVL are organised in a certain logistics node that can be closer to the production point or the end consumer. Consequently, car terminal for inbound or outbound flow of new cars may be quite distinct from one another. Very often, more services related to car logistics and customisation are provided in inbound flow, where the node (terminal or dealer’s depot) is closer to the end consumers. Such a position within the FVL usually requires different buffer strategies that are tailored to groups of customers. On this basis,
every car terminal develops its market strategy, strategy of building an adequate infrastructure and internal organisation.

Although there are few literature and scientific articles about car terminals and FVL chains, this specific field of intermodal transport is becoming interesting area of research. The position and role of a car terminal within a global FVL has been analysed by Chandra et al., Dias et al., Fischer and Gehring, Mattfeld and Kopfer, Mendonça and Dias, and Torbianelli [4, 6, 7, 10, 11, 14]. They all see FVL as a very complex chain and express the need of proper process management by the car terminal management to follow FVL expectations. The berth subsystem at a maritime car terminal is managed with higher priority, compared to the storage subsystem or inland delivery zone subsystem. Carriers and their vessels are served with priority status. Moon and Woo [12] claim that from the point of view of total costs, they represent the most important part and their shorter stay in the port significantly reduces terminal and port total costs. Van der Horst and de Langen [15] note that hinterland transport also influences port performance, but operational time-windows are predominantly focused on berthing activities.

Moreover, the inbound and outbound flows usually require different processes and different exploitation of infrastructure. Consequently, the management must monitor and adapt internal processes to the changing market and stakeholders’ expectations. Paik and Bagci [13] see the re-engineering process as an important method in securing better customer services and port’s performance improvements.

The study of the Koper port car terminal highlights strategic and operational changes that the port adopted in changing its role in FVL through the northern Adriatic transport route. The terminal’s initial role as an inbound maritime car terminal for the European market changed into a predominantly outbound platform, serving as an export gateway for European car production. On this basis, the management and operational approach for terminal process re-engineering is highlighted and elaborated in a proposed business process re-engineering (BPR) model for an outgoing oriented car terminal.

**THEORETICAL ACKNOWLEDGMENTS AND RESEARCH OBJECTS**

**THEORETICAL BASIS**

The FVL chain can be organised as a simple service in moving cars or a higher value chain where additional services are performed on the moving object – the car. Lambert et al. [9] expose that a complex logistics process requires a lot of additional knowledge and skills, in order to secure safety and security levels for cargo and logistics providers. Torbianelli [14] highlights the point that FVL can be observed as an aggregate of very complex services, especially when RO-RO and deep-sea transport has to be arranged by operators and different stakeholders. Chandra et al. [4] also point out the complexity and role of processes needed to accommodate the vessel, where these operations directly influence the system’s hierarchy decisions.

Mattfeld and Kopfer [10] analyse the transhipment operations and also see the berth subsystem as the most important one, where operational processes and the operational timetable of the entire terminal are connected with berth operational time windows. Undoubtedly, the other subsystems for storage and inland delivery zone are important subsystems with different operational and strategic problems. By defining process hierarchy, Böse and Windt [2] see potentials for autonomously executed processes in the two subsystems. From the flow-managing point of view, the inter-relations between berth-storage subsystems and vice-versa are the most important strategic and operational decisions. Namely, the inbound and outbound flows from land transport that are managed by the inland delivery zone are less intensive and the time pressure is lower. According to Dias et al. [6], terminals and ports must be flexible in offering other services such as postponement services on cars. Carmakers pose special expectations for additional services to be provided by workshops as a part of the unique intermodal node.

From the infrastructural point of view, a maritime RO-RO or car terminal is not a complex system [8]. Even though it uses a large area for storage services, the infrastructural elements are not management’s main development priorities. The terminal owners or the management are rather focused on manpower and how to fulfil RO-RO and deep-sea carrier’s expectations of shorter waiting times for free berths, shorter stays in the system and higher productivity, with fewer damages to discharged or loaded cars. This is acknowledged by Van der Horst and Van der Lugh [16], who analyse management’s focus in managing maritime terminals. According to their research, management and owners are primarily focused on manpower utilisation. They are not oriented towards better process coordination between the subsystems and horizontal communication.

Carbone and De Martino [3] claim that higher productivity can be secured only by understanding and converging the interests of all stakeholders. Certainly, RO-RO operators are an important group of stakeholders, but land transport hauliers (rail operators or trucking companies) have significant role in the system’s performance and utilisation. Vilkelis and Jakovlev [17] confirm with their study that they very often suffer from empty runs and low utilisation as a consequence of imbalanced import and export of cars. Their failure in managing inbound or outbound flows directly influences the carrier’s operational satisfaction on the berth side, especially when cars must be delivered to the port for just in time loading on the vessel. Consequently, RO-RO carriers, carmakers and inland transport operators dedicate much more attention to the outbound terminal’s lean and agile processes.

**OBJECTIVE TARGET**

The main orientation of this study is to provide an overview of strategic and tactical decisions which car terminal management takes when cargo flows significantly change direction, quantity, value and scope of additional services.
The research is not focused on multi-agent simulations, discrete-events simulations or simulation algorithms [1, 7], rather on management's decisions to provide BPR, in order to achieve operational optimisation and at the same time planned financial results.

The research provides an overview of processes re-engineering at a maritime car terminal that is crucial to following a carmaker's expectations and new trends in FVL. Trends in FVL show a constant increase in buying new cars, rising operational costs and changing environmental regulations [18]. Moreover, EU market trends indicate production increases in EU countries, tailor made production based on customer's demand, postponement of services in car modifications, shifting transport of vehicles from RO-RO transport to containers and an increase in exporting luxury cars from the EU market. All these trends directly and indirectly influence a maritime car terminal's market position and performance.

The car terminal in the port of Koper is used as an observed intermodal node. This maritime car terminal is ranked among the top 10 maritime car terminals in Europe [5]. Moreover, the terminal achieved an average yearly throughput increase of 16.3% over the last 6 years. The cargo flow drastically changed its direction during this period and on this basis the terminal transformed its logistics role from a primarily inbound car port to an outbound one. Today almost 70% of the terminal's throughput is connected to the export of European car production. In addition, the increase brought higher amounts of luxury cars that require special handling and additional services prior to loading for the new Asian markets.

The research into the last ten years development period and transformation processes yields answers to the following research hypothesis that help build an adequate BPR model for an outbound oriented maritime car terminal:

- Inbound and outbound flow of cars require different processes and real-time decisions that influence terminal productivity,
- Inbound flow uses a maritime car terminal as a mid- or long-term buffer and technology of delivering cars by road transport means are preferred,
- Outbound flows are more complex and difficult to manage, thus the terminal management shall use the BPR model when changing from an inbound to an outbound intermodal node in FVL.

THE STUDY OF KOPER PORT CAR TERMINAL

KEY TECHNICAL DATA OF THE PORT AND TERMINAL

The port of Koper is a relatively young port, as it celebrated 60 years of existence in 2017. Presently the port handles over 22 million tons of different cargo that is managed by twelve different and specialised terminals. The port covers 274 ha of land. Out of this 111 ha are used as open storage areas and 49.2 ha for covered warehouses. The total berth length surpasses 3,280 meters and the port offers 28 berths to accommodate different types of vessels. All six profit centres use the same internal road and rail infrastructure. There is no physical delimitation of each terminal area as a unique system; consequently, terminals use storage areas at different locations in the port and combine berth use according to free berths closer to a storage area.

The car terminal uses 7 berths that are situated in all 3 port basins. Pure car carriers (PCC) or RO-RO ships operators are served at the general cargo berth, the container terminal berth or in the basin 3 – used exclusively for RO-RO or deep-sea car vessels. The total berth length is around 800 m and it is equipped with four RO-RO ramps. The berth subsystem is therefore very complex, which must very often combine the use of a berth with other terminals. Moreover, internal transport processes between the berth and storage subsystem and vice-versa cross other internal flows of different cargo, such as containers, general cargo etc.

The storage subsystem is also very complex in that it uses approximately 750,000 m² of open space located in various spaces of the port grounds. The biggest storage areas are in the northern and eastern parts of the port. The terminal also handles a five-floor car garage, with a total capacity of 125,000 m². The last floor is directly connected with RO-RO berths in the first basin by a dedicated bridge. All together, the terminal can accommodate approximately 50,000 vehicles at once. Out of this, 6,000 cars can be stored in the garage.

The car terminal has been operating for 20 years. The first dedicated terminal area was built in 1996. At the beginning, only import cars were handled. Later, the first outbound cars were accommodated by the port. Last year the port of Koper was ranked in the 9th position among top European ports in handling light vehicles [5]. The terminal is ranked immediately after the biggest maritime car terminals in the Mediterranean Sea: Barcelona and Valencia. With a yearly throughput of 749,006 cars handled in 2016, the port of Koper exceeds that of the port of Bristol which handled 719,000 cars.

TERMINAL THROUGHPUT ANALYSES

During the last 10 years the car terminal handled over 5.6 million cars. In 2016, the terminal achieved a new milestone by surpassing 700,000 cars handled. The total throughput was 749,006 cars. As shown in Fig. 1 the terminal experienced rapid growth in total throughput in the last 6 years. From 2010, when yearly throughput was at the same level as in 2006, the average yearly increase up to 2016 was 16.3%. The forecasted throughput for 2017 is a further increase of 2% over the 2016 results.

An in-depth analysis shows that the high growth is primarily driven by outbound flows of new cars produced in central European markets. The outbound flow or the European export of cars through the Koper port represents 67.4% of the terminal's throughput. The terminal handled 504,228 cars in the outbound direction (see Fig. 2).
The main overseas markets are Far East markets and Turkey. The positive trend in outbound flows should continue in the coming years, because the port agreed with the European carmakers to accommodate additional volumes to be shipped to the Far East, where the new market in Japan will be served with a direct deep-sea line.

The inbound flow represented 32.6% of the terminal’s yearly throughput in 2016, when the terminal handled 244,778 cars for destination markets in central Europe. As shown in Fig. 2 the situation up to 2008 was just the opposite. The terminal was predominantly in function as an inbound intermodal node. The share of discharged cars from the vessels was 65.2%, compared to just 34.8% loaded onto vessels. The car terminal in Koper was still an important intermodal node of inbound cars with a final destination to Russia. In 2011, the inbound and outbound flows of cars were equal. Later, beginning from 2012, the terminal changed its position to an export intermodal platform for light vehicles, largely.

Further analysis of the terminal’s intermodal role shows that inbound cars for hinterland markets in Austria, Hungary, Slovakia, the Czech Republic, Poland and Slovenia are predominantly loaded on trucks. The share of loaded cars on trucks represent an enormous 95% of total inbound volume. Rail transport is used for just 5% of inbound volume. Considering that in 2009 the share of cars delivered to the hinterland markets was 7.2% and in 2013 was just 2.6%, the terminal remains in the same position, where the rail transport is not the first transport choice for car importers in Europe.

According to the analysed data the situation in outbound is just the opposite. Outbound cars arrive by rail to the terminal at the rate of 68% of total outbound volume. Regular block trains to the Koper port are used from production plants in Germany, Hungary, Slovakia and Czech Republic. Arrivals by truck thus represent 32%. Considering the inbound and outbound flows in numbers, 232,500 cars are picked up by trucking companies; meanwhile just 161,350 cars are delivered to Koper by trucks. On the other hand, 342,870 cars arrive to Koper by rail and just 12,280 cars leave the terminal on the wagons. There is a huge imbalance in full and empty transport means in arrival or departure from Koper. To some extent, this penalizes intermodal operators, because one-way transport greatly reduces income. As shown in Fig. 3, the inland delivery subsystem handles 393,900 cars that arrive or depart by trucks and 355,106 cars that arrive or leave the port by rail.

From the terminal’s perspective, this imbalance in arrival and departure by rail or truck has several negative impacts on terminal’s performance, such as:

- Wagons spend more time waiting for potential outbound volume, because rail operators would like to avoid empty return transport,
- Very often empty trucks must be positioned at Koper, causing higher congestion at the truck terminal, as the arrivals are not always just-in-time to pick up ordered cars,
- A high share of road transport causes a higher degree of random arrivals of a single truck into the system,
- A high volume of cars departing by trucks causes the need for larger loading spaces to be used for preparing cars for each truck (preparing loading lines consisting on average of 6 to 9 cars for each truck).

TERMINAL’S ADAPTATION IN CHANGING ITS ROLE AS INTERMODAL NODE

The terminal is under strong pressure from the European automotive industry’s strategy to increase the European export of new light vehicles to the Asian market. Due to congestion in northern European ports, the port of Koper must find a way to accommodate the increasing number of outbound cars. Consequently, the maritime car terminal is
constantly adapting its business and operational model from the mostly inbound oriented intermodal node in the past to an outbound one.

This study shows that several changes at different organisational levels must be made to follow market expectations. Firstly, the terminal must handle an increased number of cars that is primarily connected to static capacity of accepting a certain number of cars in the storage subsystem. Secondly, the throughput increase creates pressure on both of the other subsystems because the inland delivery subsystem must accept a higher volume of cars and the berth subsystem must provide enough free berths to accommodate RO-RO and deep-sea vessels.

As an inbound intermodal node, the maritime car terminal at Koper is often used as a storage buffer. Vehicles stay in the system longer (between 20 to 40 days) to be sold while waiting and finally delivered to the end customer. On this basis, the preferred delivery transport mode is road transport. Such a logistic process enables more flexibility and reduces additional costs within FVL, as cars are delivered directly to the dealers or end customers. Of course, this model puts more pressure on the terminal to secure enough storage positions and at the same time to offer longer free storage time.

In the past the terminal was very often requested to provide additional services on cars, such as washing, pre-delivery inspection, car customisation, repairing minor damage caused by transport or by hail; but in recent years, these services have shifted to the dealers’ offices. Moreover, in the inbound direction the terminal just unloads the cars from the vessel and positions them at the storage subsystem. This is a simple process, as cars are sorted and shifted later, when they must be prepared for the picking-up by truck or rail or delivered to the workshop (Autoservice Ltd.) within the port area, for additional services on the cars.

On the contrary, as the outbound intermodal node the terminal must prepare a higher volume of cars to be loaded onto certain vessels in advance. This is a more complex process, because the higher volume of cars to be prepared takes a longer time and larger vessels’ pre-loading storage area (last point of rest – LPR) must be secured and finally used. The last cars to arrive that are delivered in a last-minute timeframe must be taken into consideration as well. This puts pressure on the inland delivery zone and especially on truck gates. On the other hand, cars spend a shorter time in the system and this reduces the need of car positions in the storage subsystem. Consequently, the terminal faces lower pressure for static storage demand. The use of storage areas for outbound cars is in correlation with the frequency of RO-RO and deep sea lines calling the port. A higher frequency of calls reduces the need for storage space, but as presently Koper has just a biweekly deep-sea line to the Far East, the pressure on the storage area remains at a high level.

Another important element that shapes the FVL chain is connected to the special requirements of certain groups of cars. Luxury cars require a higher level of services and very often a covered storage area. The port of Koper is experiencing an increase in luxury cars movements, therefore the pressure to accommodate cars in the garage and to perform personalised services has increased. The processes involved in parking the cars in a five-floor garage are more complex, require additional transportation time and more work force. In addition, the pre-loading process of preparing the cars for loading requires additional services such as washing, providing fastening eyes, mounting the eye on both front and rear bumper and later also fastening the car on the vessel just by using the mounted eyes. To follow the export trend of luxury cars, the port of Koper strategically decided to build an additional five-floor garage and to enlarge workshop infrastructure with additional washing lines by 2018 or 2019. With this, the port will secure additional 5,000 positions for cars.

This study also highlights adaptations in rail infrastructure. Namely, the changing of the port’s role from inbound intermodal node to outbound one increased the use of rail transportation. Currently, over 350,000 cars are conveyed by rail wagons; while in 2010 less than one third of this volume was handled. Therefore, the terminal and the port made the important strategic decision to invest in new rail tracks and new rail ramps. The operational adaptation was also needed as the terminal is responsible for unloading cars from the wagon, while truckers unload cars from the trucks by themselves. The terminal also took the operational decision to prolong the working time in the inland delivery zone subsystem and extend the working week to the weekend as well.

DEVELOPING MODEL TO IMPROVE SUBSYSTEMS’ ADAPTATION AND EFFICIENCY

Based on a study of the port of Koper car terminal’s development over the last 10 years, a model of business process reengineering (BPR) to secure a reasonable adaptation of processes and infrastructure utilisation, has been worked out. The model highlights certain critical nodes in the terminal’s process workflow that have an important influence on lean processes. The outbound and inbound workflows are different from point of view of various aspects. The outbound workflow consists of the following important processes that have different impacts on management’s strategic and operational decisions:

- Very often cars arrive into the system by truck and rail until the last day, to be loaded onto the eventually berthed vessel, requiring time-sensitive services,
- Delivery of cars to the workshop can be carried out in different time-steps (directly from the truck gates or rail tracks or later from the storage subsystem or even from the LPR area for last minute services – see Fig. 4),
- Internal shunting of serviced cars from the workshop can be performed in different time periods (to storage yard area, to LPR positions or even directly to berth or vessel),
- The pre-loading process includes a higher number of cars, requires more infrastructure (large storage area for LPR closer to the preselected berth) and labour force,
• Loading on vessel requires more time and more labour compared to the opposite process of discharging the vessel, causing a longer stay in the system and higher berth occupancy (time-sensitive process).

Other elements, such as the commercial ones, have strong impacts on securing efficient outbound operations that are not present in inbound processes. Storage of cars and pre-loading processes are more complex when different groups of cars are handled (luxury cars, light goods vehicles, large goods vehicles, etc.) and when different services are necessary.

Moreover, before being loaded onto the vessel, cars must be grouped per each port of discharge (POD), in order to load them adequately on the vessel. Being the last outbound terminal on a certain RO-RO line, it undergoes commercial pressure to load the maximum possible number of cars, sometimes even more than planned. Thus, the terminal is very often called to accommodate and deliver cars with emergency status (status assigned due to available free space on the vessel or due to the agreed business).

The proposed BPR model for a maritime car terminal, when it is changing from an inbound intermodal node to outbound one, consists of five elements that are managed on different levels (see Fig. 5). Compared to the inbound processes, the described elements pose different focus on their processes due to prioritisation of vessel’s accommodation and loading plan. Fig. 5 shows main processes that should be considered in managing outgoing flows.

The inland delivery zone, consisting of truck gates and rail tracks, must secure the following processes:
- Planning working timetables to service cars destined for each vessel,
- Monitoring arrivals of cars by truck or rail with an emergency status,
- Coordinating activities with workshops, storage subsystems, LPR areas and berth subsystems in real-time,
- Prioritising acceptance and control services for cars with an emergency status.

The second element that requires the re-engineering of processes is the workshop. The workshop must secure additional storage space for emergency cars, follow the timetable of berths for servicing cars in concordance with the temporal factors of vessels, prioritise services for cars with emergency status, coordinate activities for lean internal shifting of cars to or from LPR areas, and follow customers’ demands for additional services for outgoing cars.

The storage subsystem with an LPR area forms the third and the fourth element. The storage subsystem must accommodate all outgoing cars according to the RO-RO lines’ frequency. Higher frequency decreases the need for static storage space, but places greater pressure on LPR areas, the inland delivery zone processes and the workshops. In such cases, more cars with emergency status are anticipated. The processes in an LPR area must be performed more frequently, requiring more workforce, but at the same time using less space. In connection to the different groups of cars to be handled through the terminal, the storage subsystem must provide storage in a covered area (garage or hail protected area) and a dedicated storage area as close as possible to the planned berth.

The fifth element is the crucial element in the process of a maritime car terminal. The berth, with loading processes, determines the system’s efficiency. The longer the vessel is in the system the more the pressure is put on other elements, as it results in higher berth occupancy and consequently higher congestion of the entire system. The processes at the berth, when an outgoing vessel has to be served, must provide clear communication with the carrier about free space and loading plans, loading services to secure the utilisation of the maximum vessel capacity and efficiency of loading.

The BPR model elucidates the main bottlenecks a maritime car terminal meets when it must set-up lean outbound processes that follow the commercial interests of stakeholders in FVL. The processes are different from inbound ones; thus, the management must establish a creative environment to introduce BPR as fast as possible. Partial implementation of the processes causes inefficiency and stakeholder dissatisfaction.

CONCLUSIONS

The study provides a concise view of process managing at a maritime car terminal. Through the research into the first hypothesis that inbound and outbound flow of cars require different processes and real-time decisions, the key elements in process management of inbound and outbound flows are
highlighted. Based on the study of the port of Koper, the car terminal organisation of the outbound flow of light vehicles is a more complex process, with higher pressure from the perspective of time.

The second research hypothesis that inbound flow of cars uses a maritime car terminal as a mid- or long-term buffer and that the technology of delivering cars by road transport means are preferred, is confirmed. The inbound flow of cars in arrival by vessel for hinterland EU markets is primarily served by road transport. Rail transport represents just 5% of total throughput. On the other hand, rail transport is used for 68% of incoming cars for overseas markets. Such imbalance puts operational and commercial pressure on all stakeholders. The transport companies very often use the port’s infrastructure as a time buffer in order to optimize inland transport routes.

The research also confirms that outbound flows of cars are more complex and difficult to manage, thus the terminal's management is forced to apply a BPR model when changing from an inbound to an outbound intermodal node. On this basis, a BPR model for primarily outgoing car terminal is designed. The BPR model can be used in all major car terminals, regardless of infrastructure and terminal layout in use. The presented model can be used in a process set-up basis in ports that are establishing car terminals in south Eastern Europe, like the port of Bar, which might become a regular outbound platform for car production in Serbia.

REFERENCES


CONTACT WITH THE AUTHOR

Bojan Beškovnik
e-mail: bojan.beskovnik@fpp.uni-lj.si

Faculty of Maritime Studies and Transportation
Pot pomorščakov 4
6320 Portorož
SLOVENIA

Marina Zanne
e-mail: marina.zanne@fpp.uni-lj.si

Faculty of Maritime Studies and Transportation
Pot pomorščakov 4
6320 Portorož
SLOVENIA
ECONOMICAL AND SAFE METHOD OF GRANULAR MATERIAL STORAGE IN SILOS IN OFFSHORE PORT TERMINALS

Mateusz Sondej¹
Chandana Ratnayake²
Michał Wójcik³
¹ Gdańsk University of Technology, Faculty of Civil and Environmental Engineering, Poland
² Telemark Technological Research and Development Centre (Tel-Tek), Department of Powder Science and Technology (POSTEC), Norway

ABSTRACT

The article discusses issues related with storage of granular materials in silos made of corrugated sheets and reinforced with vertical ribs. Advantages and disadvantages of these structures are named, and typical technological solutions used by largest silo producers are presented. Moreover, basic assumptions of Eurocode 3 are discussed in the context of determining the buckling load capacity of a ribbed jacket. Alternative methods are indicated to determine the silo stability using analytical and FEM based methods. General conclusions are formulated with respect to designing of silos made of corrugated sheets.

Keywords: silo, corrugated sheet, granular material, port terminal

INTRODUCTION

Bulk terminals in harbours are an important element of agricultural industry, involved in operations with various types of grains (wheat, barley, rye, rapeseed, etc.) and feed materials (soya meal, rape meal, beetroot pulp). The infrastructure of these terminals comprises silos, halls, conveyors, dryers, cleaners, mixers, laboratories, and ventilation and control systems. The basic structures used for cargo storage are large-capacity silos. The total capacity of typical terminals ranges from several tens to several hundreds of thousands of tonnes. For instance, the cubic capacity of flat-bottom silos at the Bytom Quay in the Port of Gdansk amounts to 55 thousand tonnes. The silos can be constructed of different materials, such as: steel, aluminium, reinforced concrete, brick, wood, or – rarely used – flexible materials [1]. The cross-section shape of a silo is usually circular, due to favourable stress distributions (mainly of membrane type – without bending). However, in the cases of area limitation, it may turn out economically justified to use rectangular cross-sections. The silos can have a hopper for gravitational and direct cargo discharge to road or rail transport means. However, at larger diameters (d > 8 m) they are less effective due to large hopper loads. Large-scale structures usually rest directly on a foundation plate (flat-bottom silos) and are emptied gravitationally or mechanically. In bulk terminals, the silos are most often grouped into batteries.

Reasonable choice of an optimal silo structure should take into consideration not only advantages and disadvantages of the structure alone, but also logistics, infrastructure, company's strategy, specific nature of stored materials, and climate (strong winds, earthquakes, thermal loads). Selecting an appropriate building material involves the analysis of numerous factors, such as: prices of materials, costs of building and operation, reliability and safety of the object, its durability, time of building, warranty, overall dimensions, and/or past experience of the potential owner.

At present, silos made of corrugated sheets (with horizontal folding) and reinforced with thin-walled columns, uniformly distributed around the silo perimeter and screwed to the sheets, are frequently used structural solutions of terminals used for granular material storage (Fig. 1). Their advantages include: economical steel consumption (low weight of the structure), prefabrication,
easy and very fast assembly (from several to slightly more than ten days), aesthetics, higher buckling load capacity and lower sensitivity to initial geometric imperfections, compared to silos made of flat sheets, elasticity of the structure which ensures high resistance to earthquakes, possibility for modification of the existing object, or replacement of damaged elements, or total disassembly. On the other hand, basic disadvantages of these silo structures include susceptibility to corrosion, low resistance to negative pressure and explosions, and low load resistance at non-uniform material outflow. The diameters of silos made of corrugated sheets which are offered by various producers range from about 3.5 m to 42 m. However, the majority of producers limit the overall dimensions of silos to 32 m, mainly due to roof structure limitations. The slenderness of the produced silos, defined as the height-to-diameter ratio of the cylindrical part [2], is within \( \lambda = 0.4-3.5 \). The slenderness ratios of large-diameter silos usually do not exceed \( \lambda = 1.0 \). The thickness of corrugated sheets is usually from 0.7 mm to 3 mm (thicker sheets are obtained by uniting several sheets together). The height of the sheet fold is within 10<\(d<20 \) mm, while its length is within 68<\(d<120 \) mm (Fig. 2). It is difficult to decide unambiguously which type of folding is better. The higher the sheet fold profile, the higher the bulk load capacity of the ribbed jacket. On the other hand, a high profile is characterized by higher wall roughness, which in turn increases the vertical load of the columns. The silo columns have usually the form of open thin-walled profiles made of higher-strength steel. Various cross-section shapes of folds are used to increase the resistance to local buckling, (Fig. 3). The shape of the column profile is usually selected such as to obtain high bending capacity, at simultaneous small sheet thickness (usually between 1.5 mm and 10 mm) and taking into account assembly processes (for instance, possibility of economical storage and easy fixing to corrugated sheets). In axial compression conditions, the load capacity of the column cross-section varies from 50 kN (Fig. 3j) to 2700 kN (Fig. 3i). The jackets of small-size silos can maintain sufficient load capacity without vertical reinforcement [3, 4]. For these solutions, it is more profitable to use corrugated sheets with relatively small folding.

The task of the silo structure is to carry loads, most of all coming from the stored granular material. The horizontal pressure generates circumferential tensile forces in the jacket, while the friction on the wall is the source of vertical compressive forces. The latter forces are carried by vertical columns, due to insufficient load capacity of corrugated sheets alone in this direction. The jacket ribbed in the above way is an orthotropic shell with very favourable strength parameters. The corrugated sheet serves a dual purpose: it carries the horizontal pressure of granular material and acts as elastic support for columns, protecting them against buckling. Unlike traditional silos made of flat sheets, here the silo wall only carries the tensile membrane forces, thanks to which its thickness and, consequently, weight are much smaller.

The stability loss of the structure is one of basic causes of failures in metal silos. It is noteworthy that the probability of failure in a silo is much greater than in other construction works. The silo failure can be not only a source of serious financial problems, but it also creates a potential threat to human life in direct vicinity of the silo. Despite numerous real experiments performed in silos [9] as well as numerical simulations [5–8], the main cause of silo failures is still insufficient knowledge on the complex behaviour of granular materials. The silos made of corrugated sheets are not completely free from failures [10], but they occur much more rarely, compared to failures in structures made of flat sheets. For structures whose failure would bring severe consequences, it can be reasonable to perform continuous measurements of forces in columns or jacket stresses, or direct pressure measurements. For instance, this type of monitoring is widely used in bridge structures [11] or public.
utility buildings [12, 13]. The silo geometry assessment can be made using digital photogrammetry or laser scanning [14, 15]—including scanning from harbour canals and directly from the sea in the case of silos situated in harbours [16].

LOAD CAPACITY OF SILOS MADE OF CORRUGATED SHEETS AND REINFORCED WITH COLUMNS

Determining the buckling load capacity of a silo made of corrugated sheets and reinforced with columns around the perimeter is more difficult than for the silo made of flat sheets, as it requires the stability analysis of a ribbed shell, with functional load and stepwise thickness change of jacket and column walls. At present, designing of steel silos is based on the Eurocode 3 [2]. It is noteworthy that the past Polish norm PN-B-03202:1996 concerning silos did not provide any recommendations on how to check the buckling load capacity of a silo made of corrugated sheets. Part of silos which still work safely and without failure do not meet the requirements of current standards in force concerning the stability of silo structures [2], which may suggest that these requirements are too restrictive.

The standard Eurocode 3 [2] provides two concepts of determining the buckling load capacity, depending on the distance $d_s$ between vertical stiffeners. When this distance is smaller than the limit $d_{s,max}$, the silo jacket is regarded as a shell with orthotropic properties, for which the critical buckling stress is calculated based on the Donnell-Mushtari-Vlasov theory (Fig. 4). The limiting distance between the columns is defined as [2]:

$$d_{s,max} = k_{dx} \left( \frac{r^2 D_y}{C_y} \right)^{0.25},$$

where $C_y$ is the membrane stiffness in the direction of folding, $D_y$ is the bending stiffness in the direction of folding, $k_{dx}$ is the coefficient (recommended value $k_{dx}$=7.4), and $r$ is the silo radius.

The assumptions adopted in the so-called of equivalent orthotropic shell method are as follows: uniform and linear compressive load is applied to both cylinder edges (on the central surface); the stiffness of the jacket made of corrugated sheets and reinforced with ribs is averaged (smoothed) and constant along the entire height; the central surface of the equivalent shell coincides with that of the corrugated sheet; the eccentricity of ribs and stiffening rings is measured from the central surface; the edges of the cylindrical shell are not allowed to move in radial and circumferential directions; the forms of buckling are described by trigonometric functions [2].

When the distance between the columns is larger, i.e. when $d_s > d_{s,max}$, the buckling load capacity is checked for a single vertical rib, based on the analogy to a compressed bar resting on an elastic foundation (Fig. 5). If more detailed calculations are not required, the critical force can be calculated from formula (2), which corresponds to the scheme shown in Fig. 5c:

$$N_{b,Rd} = 2\sqrt{EI_y K} / \gamma_{M1},$$

where $EI_y$ is the bending stiffness of the vertical rib, $K$ is the stiffness of the elastic foundation, and $\gamma_{M1}$ is the safety factor.

Formula (2) was obtained for the following assumptions, illustrated in Fig. 5c: buckling of the column takes place only in the plane perpendicular to the silo wall (i.e. the corrugated sheet stiffens sufficiently the column in the tangent plane to the wall); the column is only loaded with the vertical force applied to its end; the column rests on elastic foundation along its length; the column is pin-supported at its ends; the solution does not depend on column height. The stepwise change of bending stiffness of ribs and jacket can be taken into account in a more detailed analysis, making use of the Finite Element Method (FEM) [17, 18, 19], for instance. Due to relatively conservative values of buckling load capacities obtained from the model of bar resting on elastic foundation, the reducing effect of imperfections is omitted.

Simultaneous existence of these two completely different methods to calculate the buckling load capacity (depending on the distance between columns) leads to the discontinuity between them. The difference in buckling load capacity values
at the boundary of applicability of these two methods can amount to as much as 700% [18], which testifies to the imperfection of the procedure given in the standard [2]. The results of calculations made using the FEM method [10, 17–19] confirm that the approach described in the standard is too conservative, especially for large distances between columns, when the applicable column stability analysis is based on the model of bar resting on elastic foundation. For some silo geometries characterised by small distances between columns, the standard-based load capacity of a silo without imperfections is nearly twice as low as that calculated using the FEM method. For large distances, this difference is even much more dramatic and can be as much as 10 times lower than the result obtained from FEM calculations. That is why the standard-based approach should not be applied to the latter case, due to the use of an oversimplified model leading to unrealistic results. The silos with buckling load capacity calculated based on the standards are unnecessarily oversized, by even as much as 2–3 times. The buckling load capacity of a silo made of corrugated sheets and reinforced with columns can be determined correctly using the approach described in [18], which proposes some modifications of the procedure calculating the buckling load capacity for silos with a relatively large distance between columns. The use of this method eliminates the load capacity jump at the applicability boundary. It also provides results close to those obtained using the FEM method. Nevertheless, a more detailed analysis which would take into account such important phenomena as local stability loss of thin-walled elements of the structure, or column bending at the support should be based on the FEM method which provides opportunities for broad static and dynamic, linear or nonlinear, analysis of buckling load capacity. The linear analysis is simple and can be used to assess critical loads, while the nonlinear analysis, based on the equilibrium path between load and strain, makes it possible to determine the critical load taking into account the effect of material and geometric nonlinearity [20].

EFFECT OF GRANULAR MATERIAL ON BUCKLING LOAD CAPACITY OF A SILO

Another disadvantage of the standard-based approach to determining critical forces in steel silos is neglecting the stabilised effect of stiffness of the stored granular material and the normal load on the load capacity of a cylindrical silo [21]. The results of experimental examination of silos have revealed that the buckling load capacity of a silo filled with granular material is much greater than that of an empty silo (in some cases this difference can even exceed 100%). The numerical calculations [3, 4, 21] made using an advanced model describing the behaviour of granular materials (non-local hypoplastic model [22–24]) have confirmed the positive effect of the presence of material on the load capacity of the full silo, compared to the empty one. The pressure of granular material reduces the amplitude of initial geometric imperfections. Moreover, the granular material act as lateral support for cylindrical silo walls, which move inward during the buckling process. The favourable effect of reinforcement in the cylindrical silo depends most of all on: stiffness of the granular material, flexibility of the wall, type and scale of initial geometric imperfections, yield point of the steel, and type of flow in the silo. As a rule, the buckling load capacity is higher for thinner walls, larger amplitudes of imperfections, and stiffer grains of the granular material [4]. At present, this effect is not taken into account in the standards.

EFFECT OF GEOMETRIC IMPERFECTIONS

Compared to silos made of flat sheets, the advantage of silo structures made of corrugated sheets and reinforced with columns is their relatively small susceptibility to initial geometric imperfections. Only the presence of larger geometric imperfections of silo walls, of several centimetres in size, (for instance 50 mm, which is usually 30-100-fold thickness of the corrugated sheet composing the wall) leads to noticeable decrease of buckling load capacity of the structure. Nevertheless, taking into account initial geometric imperfections in steel silo designing is of high importance [25, 26]. Based on geodetically measured real imperfections [19], being the result of specific method of erection of a silo made of corrugated sheets and reinforced with columns (the silo erection begins with upper elements, which are successively elevated), we can observe that they differ considerably from the first form of buckling of the structure. Nevertheless, the effect of both types of imperfections on the buckling load capacity is similar, with differences not exceeding 20% [19]. It is therefore advisable in FEM calculations to assume the imperfection having the shape of the first form of
buckling (simpler application, compared to really measured imperfections) and the amplitude approximately equal to 5 cm, to preserve an appropriate safety margin.

SILOS MADE OF CORRUGATED SHEETS WITHOUT COLUMNS

Although they are even more economical due to lower steel consumption, the silos made only of corrugated sheets are rather rarely used, since the absence of columns in combination with small vertical stiffness of corrugated sheets result in small load capacity of the silo with respect to vertical loads. At present, there are no realistic dimensioning methods for silos made solely of corrugated sheets. The standard Eurocode 3 contains a simple procedure for determining the load capacity of a corrugated sheet:

\[ n_{\text{v,Rx}} = \max \left\{ \frac{R_y f_y}{2d}, \frac{t}{R_p f_p} \right\} \tag{3} \]

where:
- \( d \) is the height of fold, \( R_y \) is the local curvature radius of the corrugated sheet, \( r \) is the silo radius, \( t \) is the sheet thickness, and \( f_y \) is the yield point of the steel. However, this procedure is very conservative, as it does not take into account the positive effect of granular material on silo's load capacity. According to this procedure, the load capacity of the silo made of corrugated sheets and loaded with granular material is sufficient for designing a small silo, of several meters in height and dimension. However, there are much larger silos made of corrugated sheets without columns which exist and function without problems, but for which the load capacity condition calculated from the standard is several times too low.

The numerical results [3] have revealed that empty silos made of corrugated sheets (without columns) are sensitive to wind and seismic loads, which generate additional vertical forces at the support and large displacements of the upper edge of the silo. However, it was observed that the presence of granular material has a huge positive effect on the load capacity of the silo during its filling. This effect increases with the decrease of wall thickness. The buckling load capacity also increases with the increase of grain stiffness and initial material densification. In contrast to the standard-based calculations performed to determine the load capacity of a sample silo made of corrugated sheets, of 2.67 m in diameter and sheet thickness of 1 mm, which assessed this capacity as equal to 15% of sand load, the numerical calculations taking into account material stiffness have revealed that this silo can carry full sand load with large margin, of 230% [4]. For wheat, this margin was equal to 50%. While for the silo made of flat sheets, the normal and tangential pressure distributions calculated using the FEM method [4] for the time of silo filling were similar to those based on the standard Eurocode 1 [27], for the corrugated sheets they differed significantly. The tangential pressure was much smaller, which resulted in much smaller reaction of the support, compared to that obtained using the standard-based approach. On the other hand, the normal pressure distribution during filling had the shape similar to hydrostatic, which gave higher pressure than that calculated based on the standard in the lower part of the silo [4]. Due to high elasticity of silo walls, the majority of vertical load is passed directly to the silo bottom, which creates a serious threat to silos with hoppers, the presence of which is not taken into account in standard-based dimensioning procedures.

CONCLUSIONS

Based on the performed theoretical and experimental examination, the following conclusions can be made:

- silos made of corrugated sheets and reinforced with vertical thin-walled columns around the perimeter are a safe and economical method of storage of various granular materials. When properly designed, these silos have twice as high buckling load capacity and twice as small weight as silos made of flat sheets,
- analytical dimensioning of silos made of corrugated sheets which is based on present silo standards is too simplified and uneconomical, especially for silos with relatively large distances between columns (it neglects real, three-dimensional operation of the entire structure), and should not be applied,
- silos made of corrugated sheets should be designed using the Finite Element Method (FEM), which provides opportunities for more realistic assessment of silo's buckling load capacity, compared to the standard-based calculations. This approach leads to massive material savings, at the same time making it possible to analyse in detail such important phenomena as, for instance, bending of silo columns at the foundation,
- the buckling load capacity can be approximately assessed using the modified procedure described in [18],
- the granular material in the silo significantly increases its load capacity. The scale of this increase depends on: elasticity of the structure, stiffness of the stored material, and type of flow, among other factors. The standard-based procedures neglect this fact,
- the FEM analysis of a silo structure should take into account initial geometric imperfections. For simplification purposes and to be on the safe side, it is advisable to assume the imperfection having the first global form of silo buckling and the amplitude equal to 5 cm.

REFERENCES


CONTACT WITH THE AUTHOR

Mateusz Sondej
Gdańsk University of Technology
Faculty of Civil and Environmental Engineering
G. Narutowicza 11/12
80–233 Gdańsk
POLAND

Chandana Ratnayake
Telemark Technological Research and Development Centre (Tel-Tek), Department of Powder Science and Technology (POSTEC), Porsgrunn
Norway

Michał Wójcik
Gdańsk University of Technology
Faculty of Civil and Environmental Engineering
G. Narutowicza 11/12
80–233 Gdańsk
POLAND
DETERMINATION OF SHEAR MODULUS OF SOIL IN THE RC / TS APPARATUS FOR DESIGNING OFFSHORE WIND POWER PLANT FOUNDATIONS

Piotr Srokosz
Ireneusz Dyka
Marcin Bujko
The University of Warmia and Mazury in Olsztyn, Poland

ABSTRACT

The paper presents a selected aspect of the determining the initial soil shear modulus value on the research example in resonant column – torsional shear apparatus (RC / TS). There are presented the significance of the initial value of shear modulus in design of offshore wind power plant foundations and the importance of its variability in the function of cyclical shear strains of soil related to the impact of sea and atmosphere on the designed structures. Based on the conducted analyses, a new methodology for interpreting the TS test results of soil has been proposed. It allows estimating the values of the shear modulus in the full range of shear strains occurring in issues closely related to the design and construction work of offshore wind power plant foundations.

Keywords: offshore wind power, shear modulus, back analysis, numerical modelling

INTRODUCTION

Problems associated with the depletion of raw materials used for energy production motivate to search for new natural sources and technologies. The main field of these activities is intensive work related to the development of technologies for energy production using its renewable sources. These sources include wind, solar, rainfall, tides, sea waves and geothermal energy. In 2016, renewable energy sources met approximately 10% of human energy demand. In Poland, it is planned to increase production of energy from renewable sources up to 15% of energy in the national energy balance by 2020. Over the last two decades, the use of wind and solar power has been the most intensive.

Wind turbines are built both on land and at sea. They can be constructed individually or in groups called wind farms. As the placement of farms on land near areas inhabited by people raises a lot of controversy, offshore wind farms are becoming more and more popular. This is a great alternative to conventional energy. The main advantages of offshore farms compared to land based are:

- greater wind strength and stability, which allows to increase the efficiency of the wind farm and minimize the negative impact on the national electricity network,
- no constraints on the size of the structure and easier transport of large elements, making it possible to obtain more power of the farm,
- minimized impact on people and landscape, so there is no problem of social opposition,
- favored the development of many species of fish and marine mammals because they are areas with limited navigation,
- much larger space where it is possible to install devices,
- creation of much more jobs at the construction stage than in the case of land based farms.

The most important disadvantages include:

- many times higher costs of construction, connection to the power grid, maintenance and operation of the power plant,
- highly corrosive and erosive environment,
- relatively difficult, limited access during renovation or maintenance,
- difficulties in navigation.
Offshore wind energy is the future of the European renewable energy sector. The world’s first offshore wind farm was built in 1991. The efficiency of offshore wind technologies is growing every year. Over 91% of investments in offshore wind energy are located in Europe. According to the latest report [28], unquestionable leaders in Europe in terms of installed capacity are Great Britain: 6,835 MW (1,753 wind turbines) and Germany, with installed capacity of 5,355 MW (1,169 wind turbines). The following places are occupied by Denmark: 1,266 MW (506 wind turbines), the Netherlands (1,180 MW, 365 wind turbines), Belgium (877 MW, 232 wind turbines). There is great potential in Poland for the construction of offshore wind farms. According to information provided by the Ministry of Maritime Economy and Inland Navigation, as of 04.05.2016, 13 permits for the construction of offshore wind farms were in force, of which 9 were paid. Significant interest of investors, including the largest international and domestic energy concerns, in the Polish market of the IMF, provides the basis for asserting that offshore wind energy may in the 2025-2050 perspective constitute an extremely important element of the national power system. The very favorable wind conditions (Fig. 1) and the shallow waters of the Baltic Sea significantly reduce the investment costs of offshore wind farms. The construction of offshore wind farms in Poland may contribute to the creation of approx. 31.5 thousand new jobs in the perspective of 2030 [8].

Development of offshore wind farms requires appropriate calculation tools as well as normative and legal studies regulating the principles of designing these structures. This demand also includes data acquisition, design methods and analysis of wind farm construction foundations. Until 2010, the development of offshore wind farm projects focused on small depths (up to 20 m) and close to land areas (up to 20 km). It is expected that by 2030, the location will be the standard on the seas with a depth of 60 m and 60 km from the mainland.

Offshore wind power foundations are mainly constructed as: monopile, gravity based foundation or gravity based structure (GBF or GBS), tripod, jacket or tripile (Fig. 2). Currently, most solutions for wind farm foundations are based on monopile technology. However, the more wind farms will move deeper into the sea to greater depths, the foundations will require major improvements, which also include floating platform technologies [27].

During the whole period of use of the wind farm structure, different phases of the subsoil’s work should be taken into account: earthworks, reinforcing or modifying the ground, foundation construction, assembly and disassembly of structures, launching, energy production, extreme wind, failure and others. In the case of offshore wind farms, impacts are associated with the nature of the marine environment. In this case, the main impacts are dynamic and in the design process it is necessary to perform numerical simulations taking into account possible impacts. For this purpose, advanced computational methods are used, striving for the most accurate assessment of the construction cooperation with the soil in a wide range of stress state occurring at various stages of construction and operation of the structure and in the limit state.

The main aim of this work is to present a new method of numerical interpretation of torsional shear test results in the aspect of acquiring material parameters for designing offshore wind farm foundation. The article is divided into the following sections: introduction, importance of the subject in terms of current achievements in the methods of estimating the initial value of the G modulus and its significance in the issues of founding offshore objects, description of basic parameters characterizing soil stiffness, description of the resonant column – torsional shear apparatus (RC/TS) with its advantages and disadvantages, presentation of results of exemplary analyses conducted in the RC/TS apparatus, proposal of an alternative methodology for interpretation of TS test results, presentation of several sample results of analyzes carried out with the proposed back analysis methodology, summary of the results obtained and a summary of conclusions.

**IMPORTANCE OF THE SUBJECT**

The wind power plant foundation is constantly subjected to dynamic loads, which results from the specificity of turbine operation and wind gusts. Turbine producers pay attention to the specific conditions for maximum settlement and tilting of the structure, the crossing of which may lead to high tipping moments. They also mention the important role of economic aspects [11].

At the same time, according to [11], there is a lack of generally available and specific guidelines for design in Polish
standards. Author highlights that there were no instructions for conducting soil tests for those specific constructions. Therefore it was necessary to use the guidelines of the producer of a particular turbine or results obtained from experience in similar conditions.

According to the Polish Regulation of the Minister of Transport, Construction and Maritime Economy of April 27th 2012 regarding the determination of geotechnical conditions for the foundation of construction objects [25], objects of the third geotechnical category are distinguished, including atypical building structures, regardless of the complexity of soil conditions, the performance or use of which may pose a serious danger to users. That includes wind farms. At the same time, the regulation says that for construction objects of the third geotechnical category the scope of research should depend on the predicted degree of complexity of soil conditions as well as the specificity and character of the construction object or the type of planned geotechnical works and should specify mechanical soil parameters such as: internal friction angle, cohesion, undrained shear strength, and constrained or shear modulus, obtained in laboratory or field tests.

On the other hand, national publication highlights that it was common for the designer to have only one or two boreholes for each wind turbine and to determine the soil parameters using the B method according to PN-81 / B-03020. Not only that the amount of soil characteristics obtained in this way was insufficient to reliably design the foundation of the turbine, then the parameter values may deviate from the real state. Author warn that the geotechnical designer is forced to use further empirical formulas to determine the required parameters. The sizes obtained in this way may be far from the truth. The consequence of these calculations is the strengthening of the soil or piling in places where the structure can be directly placed. Therefore, attention should be paid to the fact that savings at the stage of geological or geological engineering documentation may contribute to the increase of the foundation costs in the final effect. At the end author mentions that designing the foundation of wind turbines, additional geotechnical problems should be considered in the form of soil behavior analysis in response to the dynamic / cyclical impact of the structure. Soil in the range of small deformations is involved in small change in soil stiffness results in only 2% change in natural frequency (eg a 45% change in soil stiffness results in only 2% change in natural frequency). In the article [1], the authors point out that the prediction of cyclic monopile rotation is a leading aspect in the design of a wind turbine. For this reason, it is recommended to conduct soil tests in the resonance column for full, reliable forecasting of the long-term behavior of the monopile using the concept of threshold deformation (see [13]). An example of the practical use of resonance column test results for calculating wind power plant foundations is the work of Yu, Wang and Guo [20], which investigates the long-term dynamic behavior of a offshore monopile wind power plant structure settled in the sand and the effect of long-term cyclic and dynamic load on the subsoil. The research involved the use of a scaled model of a wind turbine planted on a monopole and subjected to various types of cyclic and dynamic loads using the mass suspended at the top of the model (Fig. 4). The analyzes used model soil research results based on RC tests carried out on sands by Drnevich, Hall and Richard [7] (Fig. 5). A relationship between the natural frequency of the structure and the change in soil stiffness was observed. The results of dynamic construction research have shown that the natural frequency of wind turbine construction increases with the number of cycles, but with a reduced rate of increase in the level of soil deformation. It was also found that this change is dependent on the level of shear strain of the subsoil. It has been noted that the results of the RC test can be used to explain the behavior of the structure model during tests and well reflect the soil reaction to the dynamic load on the structure.

The article by Arany and others [1] proposes a simplified procedure for the design of wind turbine foundations, where it is recommended to use, among others, RC tests. The article points out that the dynamic stability of the structure can be endanger by changing the natural frequency of the structure over the entire lifetime of the turbine. Under the influence of environmental loads, a resonance phenomenon may occur causing a decrease in fatigue durability, worsening of SLS conditions and even a catastrophic loss of structural stability. Therefore, an important aspect is the observation of the impact of changes in soil stiffness on the natural frequency of the structure.

Figure 3 shows that large change in soil stiffness results in small change in natural frequency (eg a 45% change in soil stiffness results in only 2% change in natural frequency). In the article [1], the authors point out that the prediction of cyclic monopile rotation is a leading aspect in the design of a wind turbine. For this reason, it is recommended to conduct soil tests in the resonance column for full, reliable forecasting of the long-term behavior of the monopile using the concept of threshold deformation (see [13]). An example of the practical use of resonance column test results for calculating wind power plant foundations is the work of Yu, Wang and Guo [20], which investigates the long-term dynamic behavior of a offshore monopile wind power plant structure settled in the sand and the effect of long-term cyclic and dynamic load on the subsoil. The research involved the use of a scaled model of a wind turbine planted on a monopole and subjected to various types of cyclic and dynamic loads using the mass suspended at the top of the model (Fig. 4). The analyzes used model soil research results based on RC tests carried out on sands by Drnevich, Hall and Richard [7] (Fig. 5). A relationship between the natural frequency of the structure and the change in soil stiffness was observed. The results of dynamic construction research have shown that the natural frequency of wind turbine construction increases with the number of cycles, but with a reduced rate of increase in the level of soil deformation. It was also found that this change is dependent on the level of shear strain of the subsoil. It has been noted that the results of the RC test can be used to explain the behavior of the structure model during tests and well reflect the soil reaction to the dynamic load on the structure.
Despite the available model results and advantages of RC research, the most common practice is to use empirical correlation formulas, neglecting direct laboratory tests, which may not lead to optimal design effects.

**SOIL STIFFNESS PARAMETERS**

Modeling the behavior of structures in individual stages of loading requires knowledge of the mechanical parameters of the soil material appropriate to the calculation methods used, in particular those that describe its stiffness/deformability.

For wind power plants it shall be ensured that the properties of the soil correspond to the assumptions in the static and dynamic calculation. For the dynamic analysis, the distance of the natural frequency for the overall structure from the excitation frequencies is decisive in avoiding resonance. In the assessment of the expected natural frequencies, a parameter study is needed for the dynamic soil parameters; this shall be defined so that a range of possible soil types and soil properties is covered.

The stiffness/shear modulus is the basic mechanical parameter of the ground that allows the adoption of computational techniques used for materials exhibiting elastic properties. This parameter, however, is the instantaneous value in given conditions defined by the level of strain and stress state, for a specific load history, its speed and character (static, dynamic, cyclic), dependent on the drainage conditions, and the type of structure which transfers the load. In addition, different types of stiffness modules are used, which correspond to the definitions determined by the theoretical basis of the used analytical method. The shear modulus \( G \) (Fig. 6) is now considered to be the basic parameter characterizing the stiffness of soils.

As the level of stress and strain increases, the stiffness modulus degrades. With more complex load modes like cyclic, dynamic, monotonically variable, the value of the stiffness modulus changes with the next load step (Fig. 7). The occurring deformations \( \gamma \) can be divided into several characteristic ranges (Fig. 8). In the range of small and very small strains, the sample...
is characterized by the fully reversible behavior. The value of threshold strain (cyclic / volumetric threshold shear strain − γtv) is determined depending on the type of soil, (from 0.01% for coarse soils to 0.06% for fine-grained soils).

The value of the G modulus corresponding to this threshold deformation falls to the value 0.85-0.60 of the initial value of the module Gmax. Above this threshold value, the shape of the shear modulus degradation curve G depends on the type of load, it may be static (monotonic) or dynamic (cyclic). The range of strains above the threshold deformation γtv is called the range of plastic strain. The degradation of the shear modulus as the strain increases is higher for cyclic and dynamic loads and increases with the number of N cycles. The range of these deformations according to some authors is called medium and large strains (Fig. 8), and above 1-3% – the residual (very large) deformations. The value of the shear modulus G drops below 10% of the initial value.

The initial Gmax shear modulus is widely recognized as a parameter of the state of a given soil [14]. The shear modulus for very small deformations in dynamic problems is the same as the module Gmax for static problems both “drained” and “undrained” conditions. In dynamic tests based on the measurement of propagation velocity of a shear mechanical wave in a solid medium, the initial shear modulus is determined from a simple dependence:

$$G_0 = G_{\text{max}} = G_{\text{dyn}} = \rho \frac{V_s^2}{2}$$

where:
- ρ – material density,
- Vs – shear wave propagation speed.

This dependence is used in many research methods whose main purpose is to measure the time of the waveform of the excited elastic wave on a separate section of the subsoil or in the soil sample. The module determined in this way corresponds to very small strains. Increasingly, the use of soil test using standard penetrometers (Cone Penetration Test CPT, Dilatometer Marchetti Test DMT, Pressuremeter Menard Test PMT) equipped with geophones and devices for generating a mechanical wave in the soil is more and more popular, thanks to which it is possible to measure in situ the velocity at which a seismic wave propagates in a separate layer of soil.

More precise measurements of the propagation speed of the shear wave in the ground are obtained by using indirect measurement methods. This method is a laboratory test in a resonant column, which uses the relationship between the frequency of natural vibrations and the speed of propagation of the elastic shear wave. The RC resonant column test is based on the relationship between the shear modulus G and the resonant frequency of the tested soil sample. The functionality of the RC apparatus is extended for testing in various modes of loads and measurements, eg monotonic shear mode (TS) – torsional shearing.

According to Project Requirements [24] minimum values for the soil shear modulus G may be referred to DIN 4178: 2005-04 [30]. Designing the foundation of wind turbines towers, various limit states and computational situations should be considered. Standard limit states are not a big problem. The demanding problem is the analysis of subsoil reactions in response to dynamic and cyclic interactions from the structure as well as the analysis of mutual interaction of the substrate and structure, taking into account specific types of loads occurring in the marine environment. For land construction of wind farm towers, it is allowed to check the minimum dynamic rotational stiffness condition of the circular subsoil of the shallow foundation:

$$k_{\phi,\text{dyn}} > k_{\phi,\text{lim, dyn}}$$

where:
- $k_{\phi,\text{dyn}}$ – dynamic rotational stiffness of the soil, determined from the formula:

$$k_{\phi,\text{dyn}} = \frac{8G_0r^2}{3(1-\nu)}$$

$v$ – average Poisson’s ratio of soil,
$r$ – foundation radius,
$G_0$ – dynamic shear modulus for the ground,
$k_{\phi,\text{lim, dyn}}$ – limit dynamic rotational stiffness of the subsoil defined by the turbine producer.

The fulfillment of the condition of dynamic stiffness of the soil for the foundation of offshore wind farm structures is insufficient. A comprehensive analysis of the behavior of these structures requires the performance of analytical work based on the results of extensive and advanced geotechnical laboratory tests, in situ as well as model tests (eg using geotechnical centrifuges [2]).
RC / TS APPARATUS

Modern laboratory methods make it possible to research the soil in terms of small strains and allow better insight into the complex nature of the soil. One of the devices that enables multi-aspect analysis of mechanical properties of soil in the field of small deformations is resonant column (RC) combined with torsional shear apparatus (TS).

The RC / TS is precise measuring device that has been used in geotechnical engineering all over the world since the 1960s. The construction of the device is constantly modernized in order to broaden the spectrum of applications. WF8500 model from the British company Wykeham Farrance is a device belonging to the Institute of Building Engineering UWM in Olsztyn. The WF8500 is a RC apparatus – a resonant column – with the possibility of working in TS shear mode (Fig. 9). The apparatus is used to determine the mechanical features of the soil related to its stiffness, including the values of the strain modulus $G$ and the damping coefficient $D$. A detailed description of the of conducting the research and interpretation of results was included, among others in [18]. In the applied method, there is used propagation of elastic waves in the soil of the tested sample caused by its cyclic rotating. The device works in the frequency range of 10–300 Hz, which allows generating a shear wave corresponding to the range of small and very small shear strains. The apparatus allows testing of a full cylindrical soil sample with a diameter of 50 mm or 70 mm. The most important advantages of the device are:

- the ability to conduct RC, TS and FD (free decay) tests on the same soil sample,
- fully automated processing of results and an immediate result ($G$, $D$),
- smooth adjustment of the level of deformation and the value of the confining pressure at which the parameters $G$ and $D$ are determined (this allows to determine the relationship $G(\gamma)$ and $D(\gamma)$ for different values of confined pressure).

The device is not without flaws – the most important from the point of view of scientific research is the completely closed device control software that prevents setting individual load / strain paths (eg it is not possible to extend the time of registration of geometric parameters of the sample after TS testing to observe plastic-durable part of the shear strains of the tested soil).

Figures 10-12 show screen shots of exemplary RC, TS and FD test results, on which graphs of measured physical quantities were graphically depicted and the results of automatically interpreted mechanical features of the materials examined were tabulated.

Despite the very complex and difficult to implement the soil sample installation process in the apparatus, there should be emphasized that the test in the resonance column is considered to be highly reliable, practical and relatively convenient in terms of interpretation of measurements (see [14]).

NEW METHODOLOGY FOR DETERMINING THE SHEAR MODULUS

Obtaining a full description of the behavior of the tested material under a cyclic load requires reconstruction of the full load path taking into account the incremental changes in stiffness (in feedback) of the path function (see [9]). In numerical calculations, the simulation of the hysteresis loop phenomenon (the Masing rule) is implemented by implementing in computer applications special cases of constitutive laws of non-linear elasticity or elasticity-plasticity (eg [3] and [5]). In the state-of-the-art, you can find many interesting mathematical / numerical solutions that allow you to simulate even quite complex (non-elliptical) shapes of these loops (eg [16] and [4]) as well as their experimental modeling (eg [23]). To carry out numerical calculations related to the back analysis of TS test results, a modeling concept for the Masing rule was developed, which was programmed in c++ in the form of the computer application TS.exe. Fig. 13 shows the internal structure of the TS.exe program – calculations are carried out in the form of sequentially executed procedures and functions (names are marked in italics).

Forcing the shearing of the numerical representation of soil sample is conducting by applying increments of the angle of rotation of its upper surface ($\Theta$) calculated in each, the contractual time step ($t$) and, according to the formulas:

$$ t = \frac{2\pi n}{N} $$

(3)

$$ \Theta^{i+1} = \Theta_{max} \sin(t) $$

(4)

$$ \Delta \Theta^{i+1} = \Theta^{i+1} - \Theta^{i} $$

(5)

where:

- $t$ – contractual time calculated from the value of the step $i$,
- $n$ – number of steps set in the input data,
- $\Theta^{i+1}$ – rotation angle, respectively: previous ($i$) and next ($i+1$),
- $\Theta_{max}$ – the maximum torsion angle given in the input data.

The phenomenon of stiffness degradation is obtained by reducing the value of the Kirchhoff modulus $G(\gamma)$, and as
Fig. 10. View of the screen with RC test results

Fig. 11. Screen view with TS test results

Fig. 12. View of the screen with the results of the FD test
Fig. 13. The internal structure of the TS.exe application
a consequence the Young modulus \( E(G(\gamma), \nu) \) keeping the Poisson’s ratio \( \nu \) constant.

\[
E(\gamma) = 2G(\gamma)(1+\nu)
\]

The anisotropic changes in the mechanical properties of the sample material induced by shearing are simulated by the resultant effect of differing stiffnesses at individual integration points in the volume of a given finite element. The procedure for the modified method of tangent stiffness is shown in Fig. 14.

The adopted simplification regarding the modeling of non-uniformly degraded stiffness in the volume of material is completely acceptable, taking into account the aspects of the target use of the obtained calculation results to solve practical problems occurring in marine engineering. Additionally allows to consider the local variation of stiffness changes within each element independently and to observe the effects of their homogeneity within the entire system.

The proposed concept of determining the non-linear relation \( G(\gamma) \) consists in iteratively matching the relation \( \tau(\gamma) \), having the form of a hysteresis loop obtained from FEM simulation calculations to results obtained from TS soil research. The most commonly used technique is to perform numerical simulation of the experiment and to estimate the quality of the solution based on the adopted objective function, eg:

\[
L = \sum w (\tau^e - \tau^c)^2
\]

in which: \( \tau^e \) - measured (experimental) values \( \tau \) for given values \( \gamma \), \( \tau^c \) - determined (calculated) from the theoretical / numerical model of values \( \tau \) for given values \( \gamma \), \( w \) - weights assigned to analyzed points (computing nodes), \( n \) - number of points (nodes) with compared values \( \tau(\gamma) \).

The task constructed in this way is a complex problem and it may turn out that it is wrongly posed in the sense of Hadamard, which may result directly in difficulties in obtaining the convergence of the iterative process. However, adjusting the results of calculations to the results of experimental research can be carried out in stages, limiting the optimization process to finding coordinates of a selected subset of points associated with previously corresponding points in a subset of experimental data. Analysis of subsequent points and optimization of the local relation \( G(\gamma) \) allows to simplify the entire process and significantly improve its convergence. The proposed algorithm is shown in Fig. 15.

![Fig. 14. The algorithm of the modified tangent stiffness method](image_url)
The first stage of the algorithm implementation is to discretize the results of the experimental study. It should strictly correspond to the calculation points realized by the numerical application of the TS test. Considering the fact that the experimental data already have the form of a set of points (the sampling rate of the RC / TS apparatus is 100 Hz), the preparation of the set of computational nodes consists in the selection of measuring points with the possible application of linear interpolation in the case when the node's elevation is between measurement points.

Due to the assumption that $G$ values depend on only one argument – shear strain, a set of selected points was divided into four subsets representing four stages of sample shearing (Fig. 16):

- primary torque load: $y_{AB} \in [y_a = 0, y_b = +y_{\max}]$;
- primary unloading and secondary torque load with the opposite return until reaching $y_c = 0$: $y_{BC} \in [y_b = +y_{\max}, y_c = 0]$;
- continuation of the secondary load: $y_{CD} \in [y_c = 0, y_d = -y_{\max}]$;
- secondary unloading and subsequent loading with a torque with a positive sign to achieve $y_e = 0$: $y_{DE} \in [y_d = -y_{\max}, y_e = 0]$,

where the relations are determined separately $G(y)$. These relationships can be obtained in the form of explicit relationship after assuming the general form of the function $G(y)$ – arbitrarily assumed the function in [16] on the form:

$$G(y) = G_0 \Gamma (y, \alpha, y_{ref})$$  

$$\Gamma (y, \alpha, y_{ref}) = \frac{1 - \alpha}{1 + \frac{\gamma}{y_{ref}}} - \alpha$$  

in which:

$G_0$ - maximum value of the modulus
$G$, $y_{ref}$ – reference value of shear strain,
$\alpha$ – a constant that can be interpreted in general as [16]:

$$\alpha = \frac{G_{\min}}{G_0}$$  

Fig. 15. The algorithm of searching for the local dependence of $G(\gamma)$ by back analysis

Fig. 16. Split of the shear cycle of a material sample into four phases
where: $G_{\text{min}}$ means the minimum value of the $G$ modulus corresponding to $G(\gamma = +\infty)$.

Assuming the form of variability of shear modulus values according to formulas (8) and (9), three basic parameters were defined: $G_0$, $\alpha$, $\gamma_{\text{ref}}$, which determine the path and values of $G(\gamma)$.

In the process of minimizing the purpose function, the method of non-gradient optimization, belonging to the group of direct search methods, the Nelder-Mead algorithm was used [15]. The algorithm consists in the sequential generation of symplexes defined $n+1$ vertices in the case of optimization of the $n$-argument objective function.

**EXAMPLES OF APPLICATION OF PROPOSED METHODOLOGY**

In order to demonstrate the effectiveness of the proposed methodology for determining the value of the shear modulus, several TS tests were carried out on selected cohesive and non-cohesive soil samples. Selected samples are characterized in Table 1. The research team in Institute of Civil Engineering in Olsztyn carried out experiments in a wide range of confining pressure in the TS apparatus – from 0 kPa to 500 kPa. Some of them were selected for the numerical back analysis. All fully cylindrical samples were not saturated, silt samples had natural structure and water content, sand samples were reconstituted. Table 2 presents the parameters of numerical models of the analyzed soil samples.

Numerical models of samples discretised with tetrahedral elements with ten nodes and fifteen Gaussian points (TH10G15). The adopted number of nodes and elements in numerical analyzes are not critical values – the numeric application used in back analyzes is not sensitive to the density of discretization. The adopted number of load steps is the minimum value guaranteeing a smooth path of stiffness variability in the obtained numerical results. The other parameters result from the features measured during the TS tests.

Figures 17-22 show experimental and computational relations $\tau(\gamma)$ with the resultant dependencies $G(\gamma)$ for individual load / unload steps for one full Masinga cycle. As can be seen, reconstructing the results of an experimental study by calculations requires the use of a function $G(\gamma)$ the function describing the degradation of the stiffness of the material progressing along with the increase in the value of the shear strain. It should be emphasized that, by interpreting the obtained results, the relation $G(\gamma)$ referring to $\frac{1}{2}$–$\frac{3}{4}$ of hysteresis loop as supplementation of stiffness changes initiated by the primary material unloading corresponding to the $\frac{1}{4}$–$\frac{1}{2}$ fragment of this loop.

However, it should be remembered that forms of relation $\tau(\gamma)$ on both sides of the point $\gamma = 0$, connecting the second and third stage of twisting the sample (point C in Figure 16), can significantly differ and then separate relations $G(\gamma)$ can represent well the variable mechanical properties of the material being tested.

The obtained relations can be used for designing the offshore wind farm foundations taking into considerations shear modulus values for $\gamma = 10^{-6}$ independently in the case of positive

<table>
<thead>
<tr>
<th>Sample symbol</th>
<th>Soil</th>
<th>Diameter [cm]</th>
<th>Height [cm]</th>
<th>Weigh [g]</th>
<th>$D_{50,*}$ [mm]</th>
<th>$e^*$ [-]</th>
<th>Water content [%]</th>
<th>Volumetric density [g/cm$^3$]</th>
<th>$I_c^<em>/I_s^</em>$</th>
<th>Cell pressure [kPa]</th>
</tr>
</thead>
<tbody>
<tr>
<td>TS3057 Silt</td>
<td>7.02</td>
<td>14.14</td>
<td>1016.8</td>
<td>0.047</td>
<td>0.43</td>
<td>26.8</td>
<td>1.858</td>
<td>0.4</td>
<td>100.0</td>
<td></td>
</tr>
<tr>
<td>TS3249 Silt</td>
<td>6.96</td>
<td>14.16</td>
<td>1007.9</td>
<td>0.047</td>
<td>0.42</td>
<td>26.8</td>
<td>1.871</td>
<td>0.4</td>
<td>120.6</td>
<td></td>
</tr>
<tr>
<td>TS3935 Sand</td>
<td>7.28</td>
<td>14.31</td>
<td>1093.0</td>
<td>0.33</td>
<td>0.41</td>
<td>dry</td>
<td>1.835</td>
<td>0.94</td>
<td>48.3</td>
<td></td>
</tr>
<tr>
<td>TS4913 Sand</td>
<td>7.28</td>
<td>14.30</td>
<td>1093.1</td>
<td>0.33</td>
<td>0.41</td>
<td>dry</td>
<td>1.836</td>
<td>0.94</td>
<td>48.6</td>
<td></td>
</tr>
<tr>
<td>TS5558 Sand</td>
<td>7.00</td>
<td>14.30</td>
<td>1109.7</td>
<td>0.33</td>
<td>0.38</td>
<td>7.8</td>
<td>2.016</td>
<td>0.99</td>
<td>28.6</td>
<td></td>
</tr>
<tr>
<td>TS5919 Sand</td>
<td>7.00</td>
<td>14.31</td>
<td>1029.0</td>
<td>0.33</td>
<td>0.38</td>
<td>dry</td>
<td>1.868</td>
<td>0.97</td>
<td>28.0</td>
<td></td>
</tr>
</tbody>
</table>

* $D_{50,*}$ = nominal particle size, $e$ = void ratio, $I_c^*$ = consistency index, $I_s^*$ = compaction index

<table>
<thead>
<tr>
<th>Sample symbol</th>
<th>Number of nods</th>
<th>Number of elements</th>
<th>Number of load increments in the cycle</th>
<th>The minimum angle of twist [mRad]</th>
<th>The maximum angle of twist [mRad]</th>
</tr>
</thead>
<tbody>
<tr>
<td>TS3057</td>
<td>305</td>
<td>142</td>
<td>80</td>
<td>-16.6</td>
<td>18.7</td>
</tr>
<tr>
<td>TS3249</td>
<td>284</td>
<td>125</td>
<td>80</td>
<td>-7.9</td>
<td>9.7</td>
</tr>
<tr>
<td>TS3935</td>
<td>295</td>
<td>134</td>
<td>80</td>
<td>-0.8</td>
<td>1.3</td>
</tr>
<tr>
<td>TS4913</td>
<td>295</td>
<td>134</td>
<td>80</td>
<td>-1.3</td>
<td>1.9</td>
</tr>
<tr>
<td>TS5558</td>
<td>303</td>
<td>140</td>
<td>80</td>
<td>-2.7</td>
<td>2.8</td>
</tr>
<tr>
<td>TS5919</td>
<td>303</td>
<td>140</td>
<td>80</td>
<td>-1.7</td>
<td>2.3</td>
</tr>
</tbody>
</table>
loading, unloading, negative re-loading and unloading of the soil (in the case of the sample TS5558: 27 MPa, 30 MPa, 10 MPa and 20 MPa, respectively). However, the best approach is to use full $G(\gamma)$ analytical relationships in foundation design, using numerical applications based on FEM.

Although the obtained relations are in the form of mathematical formulas allowing their use in any range of arguments, it is necessary to avoid extrapolating Kirchhoff’s module values from obtained $G(\gamma)$ relations beyond the range of maximum values of shear strains generated during the test. The only possibility of extending the scope of interpretation is to increase the value of the angle of shearing the material sample in the TS test (the device allows the measurement of the angle of torsion up to ± 30 mRad, but due to the control of the electric voltage generating the torque – the achievement of certain deformations is conditioned by the momentary stiffness of the sample).

Obviously, it is necessary to emphasize the extremely important fact that each extrapolation of results carries the risk
of strong disturbances and distortions in the area that goes beyond the permissible range of arguments for the designated functions and therefore, it’s not advisable. The results in the range of large shear strains can be successfully supplemented with traditional laboratory tests (e.g., in a triaxial compression apparatus with on-specimen sensors).

In the presented examples of automated interpretation of laboratory test results, selected possibilities of conducting numerical analyses with the use of MES have been demonstrated, which in their application are limited only by the imagination of experimenters (see numerical analyses based on the inclusion of soil microstructure, e.g., [22]).

The back analysis of TS test results for subsequent cycles proceeds in the same way as for the first cycle, but such analyses have not been carried out and will be the subject of further research.
SUMMARY AND CONCLUSIONS

Presented in this paper subject estimating the value of soil strain parameters is certainly an important part of the process of implementing the concept of foundation engineering of offshore objects. The proposed method of interpretation of the results of the torsional shear test can be successfully applied in practical issues, giving information not only about the initial value of the shear modulus, but also about the nature of its degradation. Performing back analyses based on FEM simulations can be used without any limitations, especially in the case of numerical tool calibration, which will be carried out with calculations that dimension the foundations of the designed offshore building. Unfortunately, the method is not without flaws. Its basic drawback is the lack of proof for the universality of the designated relations $G(y)$ – they are characteristic of the specific implementation of constitutive law used in the numerical tool, which has been calibrated using a back-analysis technique.

In the near term, the authors plan to modify their RC/TS apparatus to allow long-term observation of plastic deformations of the soil (see [12]) during the TS test after the impact of the torque moment has ceased. This will allow a quantitative assessment of the share of permanent deformation in the total deformation of the material under investigation, which will allow to clarify the mathematical description of phenomena occurring during torsional shearing of soil materials.

BIBLIOGRAPHY


82 POLISH MARITIME RESEARCH, No 3/2018


25. Rozporządzenie Ministra transportu, budownictwa i gospodarki morskiej z dnia 25 kwietnia 2012 r. w sprawie ustalania geotechnicznych warunków posadowienia obiektów budowlanych, Dziennik ustaw Rzeczypospolitej Polskiej.


30. DIN 4178:2005-04, Glockentürme (Bell Towers), DIN - Deutsches Institut für Normung e.V.
FEM CALCULATIONS IN ANALYSIS OF STEEL SUBSEA WATER INJECTION FLOWLINES DESIGNING PROCESS

Michał Wodtke1
Artur Olszewski1
Artur Wójcikowski1
1 Gdansk University of Technology, Poland
2 LOTOS Petrobaltic S.A., Poland

ABSTRACT

The article describes the result of theoretical research aimed at assessing the loads and operating conditions of a Coiled Tubing pipeline injecting water, suspended to the mining platform of Lotos Petrobaltic. For this purpose, appropriate calculation models have been developed using the Finite Element Method (FEM), taking into account the nature of the analyzed object and its loads. The analyzes were carried out for two pipes (previously operated and newly proposed) differing in geometrical and strength parameters. The research was carried out for selected directions of load on the pipeline (originating from sea waves) and various variants of attaching the suspended pipeline to the mining platform.

Keywords: offshore engineering, water injection flowlines analysis, FEM simulations

ORIGIN AND PURPOSE OF WORK

Due to the planned increase in the injection pressure during the extraction of B8 deposit resources on the Baltic Sea operated by Lotos Petrobaltic, there was a need to select new pipes for water injection meeting the requirements of the implemented project. The injected water pipes are used to connect high pressure heads placed on the seabed (located in the place of the drilled wells) with the injection system located on the mining platform (Fig. 1). The solutions made in CT roll-up pipelines technology, the so-called “coil tubing” are used for this type of elements. The pipe used so far with a diameter of 4.5” (bed B3, Baltic Beta platform) was characterized by a maximum working pressure of 25 MPa, while the expected operating pressure of water after modernization is to be about 30 MPa. For this reason, after analyzing the available solutions, a new pipe with a diameter of 4” (with a maximum working pressure of 33 MPa) was initially selected.

The aim of the work was to perform a comparative analysis of CT pipes (currently used 4.5” and new 4”) during their operation under extreme conditions of stress related to sea waves. The obtained results will allow to assess the safety of the newly proposed 4” pipe and using it under conditions of increased injection water pressure. In addition, as a result of the conducted analyzes, the values of loads transmitted by the injection pipelines to the platform structure as a result of the marine environment and operational factors will be determined. Knowledge of these stresses allowed for a further process of designing pipe holders as part of the modernization works of the Petrobaltic mining platform.
ASSUMPTIONS FOR ANALYSIS

The exact determination of the actual working time (in years) for an object immersed in water (CT pipe) subjected to continuous variable loads caused by sea waves and variable internal pressure related to the water injection process is practically impossible. Of course, there are various methods available in the literature for determining fatigue strength (e.g. described in [1], [2], [3]), but they are characterized by a significant spread of results and a certain probability. An additional, separate problem is the determination of the actual values of variable loads due to sea waves, which are irregular in time.

Due to the above and due to the fact that Lotos Petrobaltic has significant operational experience related to the use of CT pipes of identical construction and similar diameter (4.5") under similar conditions, it was decided that the safety assessment of the proposed new 4" pipe will take place on the basis of comparison of working conditions and stresses in the current 4.5" and new 4" pipes. In the case of analysis of objects subjected to continuous cyclic stresses of variable amplitude, this method is definitely more reliable in relation to the determination of absolute stress values and estimation of fatigue life in years or months.

The following dimensions of the analyzed section of the injection pipe were assumed for analysis: \(A = 16\, \text{m}, B = 82\, \text{m}, R = 20\, \text{m}\) and \(E = 70\, \text{m}\) (Fig. 1). On the basis of a separate study [4], the parameters of the wave loading the CT pipe suspended on the platform were set. Table 1 presents the values adopted for the needs of analyzes (the meaning of symbols in Fig. 1). These parameters correspond to the significant wave height at the return period of 100 years.

<table>
<thead>
<tr>
<th>Wave height</th>
<th>H</th>
<th>[m]</th>
<th>9.9</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wavelength</td>
<td>L</td>
<td>[m]</td>
<td>152.6</td>
</tr>
<tr>
<td>Wave period</td>
<td>T</td>
<td>[s]</td>
<td>9.7</td>
</tr>
</tbody>
</table>

To describe the wave motion using data for the wave according to Table 1 (due to the values of the \(L/B\) and \(H/B\) parameters) the theory of Stockes wave (approximating the fifth order) [5], [6] was used. It was used to determine the instantaneous values of velocity and acceleration of horizontal water particles (in the vertical coordinate \(z\) function), on which the load of system components depends according to the dependence (1) [7]:

\[
q = C_a \rho_w \frac{\pi d^2}{4} (-a_s) + C_M \rho_w \frac{\pi d^2}{4} a_w + \\
C_d \rho_w \frac{d}{2} |u_w - u_s|(u_w - u_s)
\]  

where:
- \(q\) – unit load [N/m],
- \(\rho_w\) – density of seawater [kg/m\(^3\)],
- \(d\) – outer diameter of the pipe [m],
- \(u_w\) – horizontal velocity of water particles (dependent on \(z\)) [m/s],
- \(u_s\) – horizontal pipe speed (depending on) [m/s],
- \(a_w\) – horizontal acceleration of water particles (dependent on) [m/s\(^2\)],
- \(a_s\) – horizontal tube acceleration (dependent on) [m/s],
- \(C_a\) – weight factor added [-],
- \(C_M\) –coefficient of inertia [-],
- \(C_d\) –coefficient of normal drag [-].

The values of \(C_a\) and \(C_d\) parameters depend on the local value of the Reynolds number in accordance with the guidelines contained in [8]. For analyzes it was assumed that the seawater density is \(\rho_w = 1025\, \text{kg/m}^3\) and its dynamic viscosity \(\mu = 1.6 \times 10^{-3}\) [Belt] (to determine the Reynolds number). The calculations also take into account the additional inertia of the system (the so-called added mass) with the \(C_s = 1\) ratio [9].

ANALYSIS OBJECT - CT PIPES

During operation, injection water pipes are exposed to various loads. The most important of them are: bending moments and cutting forces (caused by sea waves); forces stretching the pipe wall (axial caused by weight and circumferential from the internal pressure of the injection water) as well as torsional moments caused by the action of offset force relative to the point of engagement of the pipe (depending on the direction of the wave inflow). The most common causes of pipeline failures are [10]: mechanical damage (due to impacts or external factors), internal or external corrosion, structural and material errors, and natural operational factors. CT pipes do not usually require joining, they are delivered in sections of appropriate length and wound, for transport purposes, on a roll of large diameter [11]. The technology of rolled pipelines has found wide application in the mining industry, for this reason they are also the subject of detailed research works, e.g. [12], [13]. Table 2 presents the most important geometrical and strength parameters of
CT pipes analyzed in this work. The pipe used so far with a diameter of 4.5” was made of X65C material. The material of the new 4” pipe (X70C) was characterized by increased durability resulting from the need to operate at an increased pressure level of water. As can be seen from the comparison of parameters, in addition to the increased strength of the material, the new pipe was also characterized by a greater wall thickness (at a smaller diameter).

Tab. 2. Comparison of selected geometric and material parameters of the analyzed CT pipes with sizes 4.5” and 4”

<table>
<thead>
<tr>
<th>parameter</th>
<th>unit</th>
<th>4.5” pipe</th>
<th>4” pipe</th>
<th>difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>material</td>
<td></td>
<td>X65C</td>
<td>X70C</td>
<td></td>
</tr>
<tr>
<td>yield point $R_y$</td>
<td>[MPa]</td>
<td>448</td>
<td>483</td>
<td>$\Delta 8%$</td>
</tr>
<tr>
<td>strength limit $R_m$</td>
<td>[MPa]</td>
<td>530</td>
<td>552</td>
<td>$\Delta 4%$</td>
</tr>
<tr>
<td>outer diameter of the pipe</td>
<td>[mm]</td>
<td>114.3</td>
<td>101.6</td>
<td>$\Delta 12%$</td>
</tr>
<tr>
<td>wall thickness of the pipe</td>
<td>[mm]</td>
<td>6.35</td>
<td>7.62</td>
<td>$\Delta 20%$</td>
</tr>
<tr>
<td>section strength index for bending</td>
<td>[mm$^2$]</td>
<td>55051</td>
<td>49191</td>
<td>$\Delta 12%$</td>
</tr>
<tr>
<td>cross-section of the pipe</td>
<td>[mm$^3$]</td>
<td>2152</td>
<td>2249</td>
<td>$\Delta 4%$</td>
</tr>
<tr>
<td>pipe material density (steel)</td>
<td>[kg/m$^3$]</td>
<td>7810</td>
<td></td>
<td></td>
</tr>
<tr>
<td>stiffness module of the pipe</td>
<td>[Pa]</td>
<td>2.1 x 1011</td>
<td></td>
<td></td>
</tr>
<tr>
<td>insulation thickness</td>
<td>[mm]</td>
<td>1.89</td>
<td></td>
<td></td>
</tr>
<tr>
<td>density of insulation material</td>
<td>[kg/m$^3$]</td>
<td>1200</td>
<td></td>
<td></td>
</tr>
<tr>
<td>working water pressure</td>
<td>[MPa]</td>
<td>20</td>
<td>30</td>
<td>$\Delta 50%$</td>
</tr>
<tr>
<td>maximum pressure of the water being pumped</td>
<td>[MPa]</td>
<td>25</td>
<td>33</td>
<td>$\Delta 32%$</td>
</tr>
</tbody>
</table>

Preliminary results of the analysis show that the largest share in the value of stresses reduced in the pipe wall is bending and internal pressure of water. The value of circumferential stresses in the pipe wall (so-called hoop stress) induced by the action of internal pressure can be estimated using a simple relationship for thin-walled pressure vessels [14]. These stresses are respectively about 212 MPa for a 4.5” pipe at a pressure of 25 MPa and about 204 MPa for a 4” pipe at a pressure of 33 MPa.

**CALCULATION MODEL**

CT pipe calculations were made using the Finite Element Method (FEM) [15]. FEM is widely used in the analysis of oil and gas extraction pipelines, e.g. to determine the impact of the seabed geometry on its loads [16] or the analysis of the pipeline installation process [17]. In this work, the analysis of the system on which the load variables associated with sea waves are working, the dynamic analysis of FEM (transient analysis) with activated large displacement effect has been used. The developed model together with the boundary conditions is shown in Fig. 2.

The model was composed of a CT tube, which was mapped using two-dimensional beam finite elements and a non-deformable seabed. It was assumed that the modeled pipe consists of a horizontal section (lying on the bottom of the sea), archways from a horizontal to vertical section and a vertical section protruding above the sea level (see Fig. 1). A regular division into finite elements with an average finite element length of approx. 0.2 m was applied. Geometric parameters of the pipe cross-sections and material data were adopted in accordance with the data contained in Table 2.

To take into account the impact of the analyzed pipeline with the seabed, contact elements were used. Contact pairs are defined between the horizontal and curved pipeline section and the surface located at the sea bottom ($z = -82$ m). The function of the contact elements was to mediate the transfer of loads between the pipe and the bottom as well as to limit the possibility of the analyzed elements moving below the seabed.

---

![Fig. 2. Boundary conditions of CT tube models used in analyzes; a) without additional restraint, b) with additional sliding restraint ($z = -10$ m); c) with additional fixed fixation ($z = -10$ m)](image-url)
level of the bottom as a result of the load. It was assumed that the coefficient of friction for contact elements is 0.2.

The model was confirmed by taking all degrees of freedom (translational and rotational) to the nodes located at the ends of the pipeline fragment included in the calculations. This boundary condition is a simplification of the reality of the real object operation, because, for example, during extreme sea waves the tube holder moves with the platform to which it is attached. Simplification, however, goes towards increasing the loads carried by the analyzed object (it increases the security of analyzes). The load of the model was (depending on z coordinate) hydrodynamic unit pressure of water calculated in accordance with relationship (1). The calculations were made for two wave directions (see Fig. 2a): direction 0 (consistent with the direction of the X axis) and direction 90 - perpendicular to the previous one (consistent with the direction of the Y axis). In the next stage of work, for the variant of loading the pipe with the action wave from the 0 direction, the influence of additional restraint of the 4” pipe (located on the coordinate z = -10 [m]) in two configurations was also analyzed: sliding restraint (Fig. 2 b) and fixed (Fig. 2 c).

The solution of the task with periodically acting variable force in time, which is the load on the wave, requires defining not only the boundary conditions, but also the initial conditions. The initial state of movement of the CT pipe due to the waving affects the values of the relative velocities and accelerations of water loading the pipe (the hydrodynamic load depends on the instantaneous water velocities and the pipe). Knowledge of the exact initial conditions would allow limiting the required analysis time to a single period of action of the T wave. In the lack of knowledge about the initial state of the system (as in the case analyzed) it was assumed that it was at rest and the total analysis time was extended to 3T. During the calculations from 0 to 2T the model of the system was put into motion and was intended only to obtain the initial conditions. In this way, repeatability of structure movements (as well as hydrodynamic loads) was obtained despite the lack of knowledge about initial conditions. All results presented in the study were obtained from calculations for the time from 2T to 3T. The time step of the analysis was set to 0.004 s. The analysis results, due to the size limitations of the database, were recorded every 0.02 s, which for the time T = 9.7 s gives 485 sets of simulation results.

SELECTED ANALYSIS RESULTS

Fig. 3 shows the calculated time variation of the level of the wave profile with the parameters according to Table 1. For the presentation of the results 4 characteristic time moments were selected (points in Fig. 3): t₁ = 0.1 s (sea level movement ⬆️), t₂ = 2.5 s (wave ridge), t₃ = 4.76 s (sea level movement ⬇️) and t₄ = 7.2 s (wave valley) for which selected analysis results were presented.

Fig. 4 compares the exemplary results of FEM calculation of the determined hydrodynamic unit load distributions of 4.5” and 4” pipes. The obtained load course was characterized by irregularity, and was different from the expected typical “smooth” distribution [2]. This can be explained by the appearance in the local model of flexible pipeline vibrations (of small amplitude) that were revealed in the results of FEM analysis. These vibrations change the continuous velocity distribution and relative acceleration of the fluid elements that wash the pipe, on which the hydrodynamic load components (related to flow resistance and inertia) depend.

In the model used, the vibrations are triggered by cyclic collisions of fragments of the horizontal part of the pipeline with a model bottom. The collisions, due to the use of contact elements, are elastic in nature. This simplification is far from reality in relation to the seabed, which exhibits plasticity (sandy-clay soil [16]), and in the place where the pipe meets the seabed, as a result of their interaction, bottom depression may occur [18]. In addition, the model does not take into account the pre-stress of the system caused by stresses in the pipe walls after it has been laid and bent at the bottom. In fact, as the post-assembly observations indicate, the pipeline adopts due to its stiffness quite a gentle arch at the transition from the vertical to the horizontal section (chain line). As a result, the phenomenon of vibration in the real object observed in the analysis results probably does not occur at all or at least has less intensity.

a)
DIRECTION 0

This point compares the results of the FEM analysis obtained for both sizes of the analyzed pipes and the direction of the inflow of the wave (direction consistent with the pipeline line on the bottom, see Fig. 2 a).

Fig. 5 compares the results of displacements in the x direction of two selected pipeline points with coordinate z = 0 m (calm level) and z = -41 m (half of the sea depth) as a function of analysis time. It can be noticed that the obtained displacement results as a function of analysis time for both pipes are almost the same. Only a slightly larger displacement, maximum by approx. 0.1 m (for point z = -41 m) was calculated for the 4.5” pipe.

The obtained stress waveforms are not continuous and they are characterized by a step change in values in successive time moments. This is due to the aggregation of the effects of global bending of the pipe due to waving and local bending caused by vibrations of the pipeline. Both effects have different frequency. The change of stresses caused by local vibrations (generated by the elastic collision of the model bottom with the pipe) is probably only a feature of the model and this phenomenon does not occur in the real object. For this reason, to establish the maximum value of bending stresses, resulting only from waving, an exponential trend line for the stresses result determined from the time range near the expected maximum value (polynomial of the second degree, blue line in Fig. 6) was used. A similar methodology for determining bending stresses in the pipeline was used for the other analyzed cases. The maximum stresses calculated in this way occurred for time t ≈ 3.3 s and it was equal to approx. 426 MPa for 4” pipe and approx. 435 MPa for 4.5” pipe. The values of tensile stresses were comparable in both cases (on average about 11 MPa for 4” and about 12.6 MPa for 4.5”).
This part of the work compares the results of FEM analysis obtained for both sizes of analyzed pipes and the direction of 90° wave inflow (direction perpendicular to the pipeline line on the bottom, see Fig. 2 a).

The obtained result of displacements in the y direction from the calculations of both pipes for two selected time moments (t1 and t3) was compared in Fig. 7. The obtained deflections of both pipes have the same shape and similar values.

In Fig. 8a results of displacements in the y direction of two selected pipeline points with z = 0 m (calmness level) and z = -41 m (half of the sea depth) as a function of analysis time were compared. The obtained results of calculations indicate that the displacements y of the selected points of both pipes are almost identical. In Fig. 8b the movement path of the pipeline point is compared with the coordinate z = -41 m in the x-y plane due to waving. Only slight differences in the results of the calculated displacement values can be observed (especially in the x direction).

Fig. 9 shows a comparison of bending and tensile stress results in a 4" (a) and 4.5" pipe (b) at the anchoring point at the platform (z = 16 m). The maximum bending stress values calculated using the trend line were observed for time t = 3.2 s. A slightly smaller value of extreme bending stresses (about 327 MPa) was calculated for a 4" pipe, for a 4.5" pipe it was approx. 348 MPa. Similarly to the direction of load action, tensile stress values were significantly lower than bending stresses and comparable for both pipes (on average about 11 MPa for 4" and approx. 12.6 MPa for 4.5").
IMPACT OF ADDITIONAL PIPE FIXATION (DIRECTION 0)

One of the ways to minimize the impact of bending the pipe caused by sea waves may be the use of an underwater holder that secures the pipe to the stationary leg of the platform. For this reason, it was computationally verified how the load and deformation of the pipe will change as a result of the additional immobilization at the coordinate z = -10 m. The analysis was performed for the 4” pipe and load action from the 0 direction and two types of additional pipeline restraint: sliding (Fig. 2 b) and fixed (Fig. 2 c).

The comparison of the results of the pipeline displacement calculations in the direction of the x axis for the three restraint configurations and the two selected time moments (t1 and t3) is shown in Fig. 10. It can be noticed that the additional restraint reduces the calculated value of the system displacements due to the waving effect. This is particularly evident in the area of <-10 m, in which the largest displacement of the pipeline is observed without additional restraint. In this fragment, the pipeline with additional restraint achieves maximum displacements in the x-direction from approx. -1.8 m to 0.7 m, while without additional restraint it is maximum approx. ± 2.5 m.

In Fig. 11 the results of displacements in the x direction of two selected pipeline points were compared: z = 0 m (calm level, Fig. A) and z = -41 m (half of the sea depth, Fig. B) as a function of analysis time. The largest displacements were characterized by the result obtained for the pipe without
additional restraints, the smallest system with additional fixed restraint. The determined displacements of the point $z = 0$ m for models with additional restraint were of different character, while the result obtained for the pipeline coordinate with $z = -41$ m was very similar for both mentioned cases.

Fig. 12 shows a comparison of the variation in time of bending and tensile stress for models with additional restraint at the point $z = 16$ m (fixing the tube at the platform) and in Fig. 13 for the point $z = -10$ m (additional restraint).

![Fig. 11. Comparison of FEM displacement results for a 4" pipe in the x direction (load direction 0) for different restraint options: a) $z$ point = 0m; b) $z$ point = -41m](image1)

![Fig. 12. The results of FEM calculation of bending and tensile stresses [MPa] in fixing the pipe to the platform with $z = 16$ m; a) with sliding restraint; b) with fixed fixation](image2)

The calculated maximum value of bending stresses in fixing the pipe to the platform ($z = 16$ m) equals approx. 262 MPa (for $t = 2.6$ s) for the case with additional sliding restraint and approx. 103 MPa (for $t = 2.5$ s) for the case with an additional fixed restraint.

The obtained course of variation of bending stresses for a model with additional fixed restraint is continuous, because the full bonding of the model does not allow propagation of local vibrations above the level $z = -10$ m. Another character has the course of tensile stresses that reach a comparable level to bending stresses (max. about 78 MPa). The peak of the tensile stress values for the moment of maximum bending stress is clearly visible (for $t = 2.5$ s).

The time course of values of bending and tensile stress in the pipe at the level of additional restraint is shown in Fig. 13. The maximum stress value for the case of the floating restraint is about 260 MPa (for $t = 2.8$ s) and for fixed restraint about 161 MPa (for $t = 2.5$ s).
DISCUSSION OF RESULTS

In Table 3, the calculated values of loads and stresses in pipe restraints for all analyzed cases are summarized.

Tab. 3. Synthetic summary of the analysis results

<table>
<thead>
<tr>
<th></th>
<th>direction 0</th>
<th>direction 90</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>4.5&quot;</td>
<td>4&quot;</td>
</tr>
<tr>
<td>max σg (z=16 m) [MPa]</td>
<td>23950</td>
<td>20950</td>
</tr>
<tr>
<td>max σg (z=10 m) [MPa]</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>max Mz (z=16 m) [Nm]</td>
<td>176900</td>
<td>176900</td>
</tr>
<tr>
<td>max Mz (z=10 m) [Nm]</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>max σr (z=16 m) [MPa]</td>
<td>12.6</td>
<td>11.2</td>
</tr>
<tr>
<td>max σr (z=10 m) [MPa]</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>max Fz (z=16 m) [N]</td>
<td>22700</td>
<td>25100</td>
</tr>
<tr>
<td>max Fz (z=10 m) [N]</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>σflow (dia. Dmax) [MPa]</td>
<td>212</td>
<td>204</td>
</tr>
</tbody>
</table>

The symbols in the table mean: σg - bending stress, Mz - bending moment, σr - axial (tensile) stress, Fz - axial force and σobw - internal pressure stresses. Analyzing the results for a pipe without additional restraint, it can be concluded that stresses originating from bending and internal pressure are the dominant stresses. The level of bending stress is very similar in both pipes (4.5" and 4"), although small differences (in the order of 9 - 21 MPa) are noticeable in favor of the 4" pipe. On the basis of the obtained results, it can also be stated that:

- the most unfavorable direction of the wave is the inflow of the wave along the x-axis, i.e. along the fragment of the pipe laid on the bottom,
- calculations showed that the best solution limiting the bending stress caused by sea waves would be an additional fixing of the pipe in the holder mounted under the surface of the water, this solution effectively reduces the acceptable displacement of the pipe, thereby reducing bending stress,
- in the case of using an underwater holder for the sliding restraint, the horizontal reaction component is about 10 kN, while for the holder with non-displacement restraint the horizontal component is about 5 kN, the vertical component about 176 kN and the moment of force about 7900 Nm,
- unfortunately, the use of additional fixation of the CT pipe is difficult to implement, because the mounting of the holder and the connection of the holder and pipeline would have to be made below the sea level in immersion by a team of divers. In addition, a separate issue to consider in this variant would be the durability of the underwater grip,
- no large displacement of the pipeline is currently observed. It seems that the decisive factor is its relatively small diameter and the recorded small wave pressure.

CONCLUSION

The calculations indicate that the working conditions chosen by Lotos Petrobaltic and 4" working on the B8 bed made of CT X70C material will be similar to the conditions in which the previously used CT 4.5" pipes work on the Baltic Beta platform (B3 bed). On this basis, it can be concluded that the durability of the new 4" pipes will be similar to the durability of the existing CT 4.5" pipes. In addition, it can be stated that:

- both currently used 4.5" pipes as well as new 4" pipes intended for installation on the modernized Petrobaltic platform work in difficult conditions and are heavily strenuous in extreme sea conditions. However, as the experience of exploitation indicates, the main problem of using CT pipes is their internal corrosion, which reduces the wall thickness and can lead to the breaking of the pipeline,
- all parameters of the new CT pipe to be installed on the modernized Petrobaltic platform were correctly selected: the material was changed to have a larger R_n and Rm, the wall thickness was increased and the diameter was reduced. As a result, it was possible to increase the injection pressure to about 30 MPa, maintaining the minimum permissible flow section. Of course, in order to increase reliability, it would be advisable to use a 4" pipe with even thicker wall or greater strength. However, this could give rise to additional technical difficulties related to the applied pipeline laying technology (the impact of increased stiffness of the pipe on its rewinding and laying process).
The obtained results of analyzes were used in further design work related to the installation of new CT 4” pipes as part of the water injection system. Fig. 14 shows the CT tubes installed on the LOTOS Petrobaltic platform, which were the subject of the analyzes described.

Fig. 14. CT 4” pipes operating in the water injection system on the LOTOS Petrobaltic platform

ACKNOWLEDGMENT

The authors would like to thank LOTOS Petrobaltic S.A. and A&O EXPERT for cooperation and consent to publish the results of the work.

LITERATURE


CONTACT WITH THE AUTHOR

Michał Wodtke
e-mail: mwodtke@pg.edu.pl

Gdańsk University of Technology
11/12 Narutowicza St.
80 - 233 Gdańsk

POLAND
THERMO-ECONOMIC ANALYSIS AND ENVIRONMENTAL ASPECTS
OF ABSORPTION REFRIGERATION UNIT OPERATION ONBOARD
MARINE VEHICLES: RO- PAX VESSEL CASE STUDY

N R Ammar1
I S Sediek2
1 Alexandria University, Faculty of Engineering, Egypt
2 College of Maritime Transport & Technology Arab Academy for Science, Technology & Maritime Transport,
Alexandria, Egypt

ABSTRACT

Marine diesel engines lose a huge amount of fuel heat content in the form of exhaust gas and jacket cooling water, especially onboard high-powered marine vehicles such as Ro-Pax ships. In this paper, the possibility of using the waste heat of marine diesel engines as a source of heat for air conditioning absorption system is investigated. The thermodynamic analysis, in addition to the environmental and economic analysis of the air condition absorption cycle operated with two heat sources using lithium bromide as absorbent, are performed using the Engineering Equation Solver (EES) software. The last 10 years have seen a steady growth in the passenger ferry and Ro-Pax market, with particularly strong growth in passenger numbers. As a case study, a Ro-Pax vessel operating in the Red Sea area is considered, regarding the profitability of using air conditioning absorption system. The results show specific economic benefits of the jacket cooling water operated absorption refrigeration unit (ARU) over the exhaust gas operated unit, with annual costs of capital money recovery of 51,870 $/year and 54,836 $/year, respectively. Environmentally, applying an ARU machine during cruising will reduce fuel consumption by 104 ton/year. This, in turn, will result in reducing NOx, SOx, and CO2 emissions with cost-effectiveness of 7.73 $/kg, 20.39 $/kg, and 0.13 $/kg, respectively.

Keywords: ship emissions, IMO, lithium bromide-water ARU, thermodynamic analysis, economic and environmental analysis

 NOMENCLATURE ABBREVIATIONS

\[ h \] enthalpy, kJ/kg
\[ m_r \] mass flow rate of refrigerant, kg/s
\[ m_{r1} \] mass flow rate of steam out of the evaporator, kg/s
\[ m_{r2} \] mass flow rate of liquid carryover from the evaporator, kg/s
\[ m_{s} \] mass flow rate of strong solution, kg/s
\[ m_{w} \] mass flow rate of weak solution, kg/s
\[ v \] specific volume, m³/kg

ARU Absorption Refrigeration Unit
CO\(_2\) Carbon Dioxide
HC Hydrocarbon Emissions
IMO International Maritime Organization
NO\(_x\) Nitrogen Oxide Emissions
PM Particulate Matter
SO\(_x\) Sulfur Oxide Emissions
LiBr Lithium Bromide
INTRODUCTION

For the year 2012, total shipping emissions were approximately 938 million tonnes CO$_2$ and 961 million tonnes CO$_2$e for GHGs combining CO$_2$, CH$_4$ and N$_2$O. International shipping emissions for 2012 are estimated to be 796 million tonnes CO$_2$ and 816 million tonnes CO$_2$e for GHGs combining CO$_2$, CH$_4$ and N$_2$O. International shipping accounts for approximately 2.2% and 2.1% of global CO$_2$ and GHG emissions on a CO$_2$ equivalent (CO$_2$e) basis, respectively [1]. This study estimates multi-year (2007–2012) average annual totals of 20.9 million and 11.3 million tonnes for NO$_x$ (as NO$_2$) and SO$_x$ (as SO$_2$) from all shipping, respectively [2, 3]. Moreover, recent studies of ship emissions state that the shipping-related particulate matter (PM) emissions are responsible for approximately 60,000 of cardiopulmonary and lung cancer deaths annually, with most of those deaths occurring along the coasts [4].

In addition to the abovementioned emissions, the refrigerants used onboard vessels for air conditioning and cargo cooling purposes are another source of ship emission. These refrigerants are either ozone-depleting substances, such as chlorofluorocarbons (CFCs), or their replacements, such as hydro fluorocarbons (HFCs) 1,1,1,2-tetrafluoroethane (R134a), and the mixture of pentafluoroethane, trifluoroethane and tetrafluoroethane (R404a). All these refrigerants have significant global warming potential [5]. Based on most recent statistics, the average annual loss of refrigerants from the global fleet makes the air conditioning equipment responsible for about 69.8% of the total loss of refrigerants, and 30.2% of this loss is related with the cooling equipment. Fig. 1 shows the contribution of each ship type, based on the loss of refrigerants due to air conditioning. It can be noticed that the largest amounts of loss of refrigerants are from general cargo and cruise ships [1]. Economically, there is a notified increment of fuel oil prices which presents a challenge for the marine industry.

REGULATIONS CONTROLLING SHIP EMISSIONS

Due to the continuous increase of emissions from ships, IMO issued a set of regulations regarding this concern in the form of the International Convention for the Prevention of Pollution from Ships (MARPOL), Annex VI. Regulation 14 limits the emissions of SO$_x$ and particulate matter (PM), while Regulation 13 and the energy efficiency design indexes limit the NO$_x$ and CO$_2$ emissions, respectively [6-8]. A solution which leads to the reduction of ship emissions bases on absorption refrigeration units (ARUs), which utilize engine heat losses [9, 10]. The heat balance of the marine diesel engine, being the most frequently used prime mover on ships [11], shows that approximately more than 50% of the fuel input is lost in heat losses, including the exhaust gas (25.5%), fresh water cooling (5.2%), charge air cooling (2.9%), and radiation (0.6%) [9, 10, 12]. Three of these heat losses can be utilized, at different temperature levels, as a heat source for ARU: exhaust gas (300-600°C), charge air (200°C), and jacket cooling water (80-100°C) [12]. Exact amounts of these heat losses will depend mainly on the cycle of operation and engine speed.

Vapor compression refrigeration systems are in common use onboard ships. These systems are powered with electric energy [13]. Although they have the advantages of high Coefficient of Performance (COP) and low purchase price, their use will be phased out with time due to their contribution to the greenhouse effect and ozone layer depletion [14]. Regulation 12 of the MARPOL convention, Annex VI, issued by IMO, states that: new installations containing ozone-depleting substances are prohibited on all ships from January 2020 [15]. On the other hand, absorption cooling is an environmentally friendly cooling method. It uses a number of refrigerant–absorbent pairs, the most common of which are water-lithium bromide and ammonia–water. They offer good thermodynamic performance [10] and can be used in single-stage and double-stage coolers, depending on the input heat source. Single-effect types are suitable for waste heat applications. The ammonia absorption cooling system can provide cooling down to -60°C [16]. On the other hand, water-lithium bromide absorption coolers can be used mainly in air conditioning to cool down to temperatures above zero Celsius [12].

This paper presents the thermodynamic analysis of a single-effect water-lithium bromide absorption refrigeration machine, starting with the cycle description and validation. The analysis includes investigating the sensitivity of cycle performance indicators. The paper also aims to study both the economic and environmental effects of using absorption refrigeration for reducing both the running costs of air-conditioning during the ship cruise, and the harmful emissions. A high-speed passenger ship is used as a case study.
RO-PAX SHIP CASE STUDY

The Ro-Pax AlKahera is one of passenger ships operating in the Red sea area. The ferry sails between the port of Duba in the Kingdom of Saudi Arabia and the port of Safaga in Egypt. The main technical data of the ferry are given in Table (1) [17]. The increased number of these ships sailing on the Red Sea have improved the maritime transport in this area by increasing the number of voyages per year, which, however, also increased the amounts of gaseous emissions [13, 18]. These emissions have harmful impact on the Red Sea environment.

<table>
<thead>
<tr>
<th>Ship Name</th>
<th>AlKahera</th>
</tr>
</thead>
<tbody>
<tr>
<td>Type</td>
<td>Ro-Pax</td>
</tr>
<tr>
<td>IMO number</td>
<td>9441776</td>
</tr>
<tr>
<td>Year of build</td>
<td>2008</td>
</tr>
<tr>
<td>Flag</td>
<td>Egypt</td>
</tr>
<tr>
<td>Passengers/crew</td>
<td>1200/18</td>
</tr>
<tr>
<td>Main Engine</td>
<td>4 x MTU 20V 8000 M71R</td>
</tr>
<tr>
<td>Maximum Continuous Rating</td>
<td>4 x 7,200 kW @ 1,150 rpm ± 1.5%</td>
</tr>
<tr>
<td>Service Speed</td>
<td>34 knots</td>
</tr>
<tr>
<td>Fuel Consumption at 90% MCR</td>
<td>5,731 m³/h</td>
</tr>
<tr>
<td>Generating Sets</td>
<td>4x MAN D2866 LXE LSAM 46.2 VL12 , 228kW @ 1500 RPM</td>
</tr>
<tr>
<td>Specific fuel consumption of generators</td>
<td>260 g/kWh</td>
</tr>
<tr>
<td>Air condition cooling capacity</td>
<td>200 kW (4 units)</td>
</tr>
<tr>
<td>Number of trips per year</td>
<td>250</td>
</tr>
</tbody>
</table>

The heat flow analysis for one of the main engines indicates that there are two high waste energy streams that have potential to be recovered using ARU. These two streams are the heat losses in the exhaust gas and in the jacket cooling water. With the exhaust mass flow rate of 8.7 kg/s and the temperature of 460°C, the exhaust gas represents 25.5% of the total input power of 16,093 kW. On the other hand, the amount of heat transferred to the jacket cooling water is 1,091 kW at the mass flow rate of 26.11 kg/s, the output temperature of 85°C, and the temperature difference of 10°C [19]. Based on the technical data of the case study, the Carrier absorption unit (16LJ11) with nominal cooling capacity of 264 kW [20] can be used onboard to cover the required refrigeration load of 250 kW. The required mass flow rates of hot water flow and cooling water flow for the selected model are 10.4 kg/s and 17 kg/s, respectively, while the inlet and outlet temperatures of these flows range within 95°C – 75°C and 29.4°C – 38.4°C, respectively. The mass flow rate of cooled water is 11.4 kg/s, with inlet and outlet temperatures of 12.2°C and 6.7°C, respectively. Both, the exhaust gas source and the jacket cooling waste heat source can provide the system with the required hot water mass flow rate of 10.4 kg/s, and the temperature difference between 95°C and 75°C. The absorption machine is assumed to be used only during the cruise period of 8 hrs.

THERMODYNAMIC ANALYSIS OF THE ABSORPTION AIR CONDITIONING CYCLE

The most popular refrigeration and air conditioning systems are based on the vapour absorption systems. These single-effect absorption refrigeration cycle consists of eight components: generator, condenser, evaporator, absorber, pump, two throttle valves, and heat exchanger (HX), as shown in Fig. 2 [21]. The working fluid is a mixture of water and lithium-bromide. The generator provides the cycle with the heat (Q_g) to evaporate the water from the water-LiBr solution to high-pressure steam. The produced steam flows to the condenser, in which the thermal energy (Q_c) is rejected to the cooling medium.

A throttle valve reduces the steam pressure and allows it to return to the vapor phase. The evaporator represents the cooling capacity (Q_e) of the absorption machine. The exit low- pressure steam of the evaporator is absorbed into strong lithium-bromide solution coming from the generator by rejecting the heat energy (Q_a) to the cooling medium. The output temperatures of the absorber and the condenser may have the same values, depending on the design of their cooling cycles [22-25]. Before the solution returns to the generator for a new cycle, its pressure at absorber exit is raised using a pump. The solutions leaving the generator and the absorber are referred to as strong and weak solutions, respectively, with reference to the percentage of lithium-bromide [26]. The solution heat exchanger preheats the weak solution using part of the heat energy of the boiling strong solution leaving the generator, which improves the cycle efficiency.

![Fig. 2. Schematic diagram of the water-lithium bromide absorption cycle driven by main engine waste heat](image)

With the intention of conduct the performance evaluation of the water-lithium bromide absorption cooler, some assumptions and initial values are to be considered. The required initial values are: the evaporator capacity (Q_e), output temperatures of the generator, condenser, evaporator,
and absorber, and the effectiveness of the solution heat exchanger. The basic assumptions for the thermodynamic model are: the cycle is steady-state, the refrigerants at evaporator and condenser outlets are assumed saturated vapors and saturated liquids, respectively, the flow restrictors are adiabatic, the heat loss to the surroundings is negligible, and there are no pressure losses in pipes and heat exchangers.

The mass fractions of strong and weak solutions, measured in kg LiBr/kg, are $X_{ss}$ and $X_{ws}$, respectively. These fractions can be expressed in terms of temperatures (°C) of the generator ($T_g$), condenser ($T_c$), absorber ($T_a$), and evaporator ($T_e$), using Eqs. (1) and (2) [9]. The indexes presented in Fig. 2 are used to nominate the state points in the cycle.

$$X_{ss} = \frac{49.04 + 1.125T_g - T_c}{134.65 + 0.47T_g} \quad (1)$$

$$X_{ws} = \frac{49.04 + 1.125T_a - T_e}{134.65 + 0.47T_a} \quad (2)$$

Based on the above assumptions, the mass and energy balance equations can be applied to the absorption cycle shown in Fig. 2, taking into account the operating conditions at each point. The liquid carryover from evaporator is assumed to be 2.5% [27]. Applying the energy balance equations to the absorber, the condenser, and the evaporator, gives the following expressions:

$$\dot{Q}_a = \dot{m}_1h_{10} + \dot{m}_2h_{11} + \dot{m}_{ss}h_6 - \dot{m}_{ws}h_1 \quad (3)$$

$$\dot{Q}_c = \dot{m}_f(h_7 - h_6) \quad (4)$$

$$\dot{Q}_e = \dot{m}_1h_{10} + \dot{m}_2h_{11} - \dot{m}_f h_9 \quad (5)$$

The ratio of the solution mass flow rate leaving the absorber to the vapor mass flow rate entering the condenser is called the cycle circulation ratio ($F$).

$$F = \frac{\dot{m}_{ws}}{\dot{m}_f} \quad (6)$$

The temperatures of the solution heat exchanger exit points can be determined from the energy balance (7) and the heat exchanger effectiveness ($\varepsilon_{HE}$) equation (8).

$$\dot{m}_{ws}h_2 + \dot{m}_{ss}h_4 = \dot{m}_{ws}h_3 + \dot{m}_{ss}h_5 \quad (7)$$

$$\varepsilon_{HE} = \frac{T_3 - T_2}{T_4 - T_2} \quad (8)$$

The coefficient of performance (COP) measures the refrigeration cycle performance. It can be calculated using Eq. (9).

$$COP = \frac{\dot{Q}_g}{\dot{Q}_g + \dot{W}_p} \quad (9)$$

where $\dot{Q}_g$ is the heat input to the generator. It can be determined from the generator energy balance, Eq. (10).

$$\dot{Q}_g = \dot{m}_{ss}h_4 + \dot{m}_f h_7 - \dot{m}_{ws}h_3 \quad (10)$$

$\dot{W}_p$ is the power required to pump the solution from the exit absorber low pressure ($P_L$) side to the generator high pressure ($P_H$) side.

$$\dot{W}_p = \dot{m}_{ws}v_1(P_H - P_L) \quad (11)$$

First law of thermodynamics has been used to analyze and optimize inclusively the performance of absorption heat transformer operating with LiBr/H2O as the working pair. The mass and energy equations for the water-lithium bromide absorption cycle are analyzed using the Engineering Equation Solver (EES) software under steady-state operation. The solver has the built-in functions of thermodynamic properties of steam and water-lithium bromide mixtures. It performs multiple iterations using the mass and energy balance equations to determine the values of enthalpy (h) and mass flow rate (m) at each point of the absorption cycle based on the input data.

The performance of steam and water-lithium bromide mixtures for the current EES model were verified using the cases provided in [28, 29]. In addition, the current water-lithium bromide absorption cycle, modelled using the EES program, was validated on theoretical and experimental data provided by [27]. The main input parameters were: cooling capacity equal to 11 kW, temperatures of the generator, absorber, condenser, and evaporator, equal to 90°C, 34.9°C, 30°C, and 6.7°C, respectively, strong and weak mass fractions amounting to 60% and 55.5% respectively, and the heat exchanger effectiveness equal to 60%. Table (2) compares the results of the current study with those published in the literature. The maximum percentage difference between these results was 1.13% for the cycle COP.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Published study [27]</th>
<th>Present study</th>
<th>Percentage difference (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Condenser (kW)</td>
<td>10.78</td>
<td>10.90</td>
<td>+ 1.11</td>
</tr>
<tr>
<td>Generator (kW)</td>
<td>14.2</td>
<td>14.35</td>
<td>+ 1.05</td>
</tr>
<tr>
<td>Absorber (kW)</td>
<td>13.42</td>
<td>13.44</td>
<td>+ 0.15</td>
</tr>
<tr>
<td>Pump (kW)</td>
<td>0.29</td>
<td>0.289</td>
<td>- 0.34</td>
</tr>
<tr>
<td>COP</td>
<td>0.704</td>
<td>0.696</td>
<td>- 1.13</td>
</tr>
</tbody>
</table>
ENVIRONMENTAL AND ECONOMIC ANALYSIS OF THE ABSORPTION AIR CONDITIONING CYCLE

The reduction in the quantity of annual emission during cruise after applying ARU (ER_{ARU}) can be calculated using Eq. (12)

\[ ER_{ARU} = P_E \cdot T_s \cdot F_e \]  

(12)

where \( P_E \) is the saved electric power in kW at cruise, \( T_s \) is the number of sailing hours per year, and \( F_e \) is the emission factor for the engine in (g/kWh).

ARU can be considered as one of the emission reduction measures. The annual emission cost-effectiveness after applying ARU onboard the ship (ACE_{ARU}) can be calculated with respect to the reduction of each type of pollutant emissions using Eq. (13) [6, 30].

\[ ACE_{ARU} = \frac{C_{ti}}{ER_{ARU}} \]  

(13)

where, \( C_{ti} \) is the total annual cost of ARU machine which includes capital, operating, and maintenance costs in ($/year).

The total annual installation cost (\( C_{ins} \)) of the ARU system depends on the annual initial money recovery (AMR), the installation cost (\( C_{ins} \)), the annual maintenance and operating cost (\( C_{m&o} \)), and the heat exchanger cost (\( C_{HE} \)) in the case of the exhaust gas operated ARU. \( C_{ti} \) can be calculated using Eq. (14)

\[ C_{ti} = AMR + C_{ins} + C_{m&o} + C_{HE} \]  

(14)

The annual money recovery (AMR) when applying ARU depends on the unit cost (UC), the expected ship age after applying the absorption system (\( n \)), and the interest rate (\( i \)) [31]. AMR can be calculated using Eq. (15).

\[ AMR = UC \times \frac{i(1 + i)^n}{(1 + i)^n - 1} \]  

(15)

Fuel saving due to the use of ARU instead of CRU during sailing (\( m_{fs} \)) depends on the electric power saved during the trip (\( P_E \)) in (kW), the specific fuel consumption of the generator (\( b_{eg} \)) in (g/kWh), and the number of sailing hours per year (\( T_s \)). Thus, \( m_{fs} \) can be calculated using Eq. (16).

\[ m_{fs} = b_{eg} \cdot P_E \cdot T_s \]  

(16)

The fuel saving cost (\( C_{fs} \)) will depend on three main factors, including: the amount of saved fuel, fuel prices (\( C_f \) in $/ton), and the yearly fuel price change (PI). This can be expressed as the following equation:

\[ C_{fs} = m_{fs} \cdot C_f \cdot [1 \pm PI]^n \]  

(17)

RESULTS AND DISCUSSION

The results compare the performance of two ARU waste heat recovery systems. The first system uses the main engine exhaust gas as the heat source for the generator, while the second uses the jacket cooling water. In the study, the operating temperatures of the condenser and the absorber were assumed within the range from 20 to 40°C, based on the used cooling water conditions [9, 32]. The assumed hot water temperatures were within the range of 95°C to 86°C, with the average generator output temperature of 90.5°C, for the exhaust gas operated generator, and within the range of 85°C to 75°C, with the average generator output temperature of 80°C, for the jacket cooling water operated generator. This was based on the main engine heat balance data and the selected carrier model for ARU unit. The evaporator temperature for this unit is 6.7°C [20]. The generator temperature for a single-stage water-lithium bromide ARU unit was assumed within the range from 75°C to 120°C [21, 33, 34]. The cooling capacity of the ARU unit was assumed equal to 250 kW based on the ALKahera high-speed passenger ship technical data shown in Table (1).

The results include the effect of condenser and evaporator temperatures on the performance of the absorption refrigeration cycle. In addition, the environmental and economic results of applying the above two waste heat recovery operated ARU machines onboard the ship selected as the case study are discussed.

THERMODYNAMIC RESULTS

Although ARU consists of many components, only its three components: condenser, absorber and evaporator, play a key role in its applicability onboard ships. It is noteworthy that the effects of the condenser and the absorber on refrigeration cycle performance reveal nearly similar trends, which also coincide with those recorded in the published paper [35]. In this section, the effect of condenser and evaporator temperatures on the water-LiBr refrigeration cycle is studied as follows:

Effect of ARU condenser temperature

The effect of the condenser temperature on the performance of the water-lithium bromide refrigeration cycle is examined by changing \( T_c \) and \( T_g \), while simultaneously maintaining \( T_a = 30°C \), and \( T_e = 6.7°C \). Fig. 3 shows the effect of condensation temperature on the cycle COP and the required generator power for both exhaust gas and jacket cooling water heat sources. The required generator power increases significantly with the increasing condensation temperature. Its maximum values change from 325 kW for the generator operated on jacket cooling water, to 315 kW for the generator operated on the exhaust gas. In contrast, the cycle COP decreases with the increase of condenser temperature. This COP trend coincides with the experimental results presented in [36].
Variations of absorber power and strong solution mass fraction at different output condensation and generation temperatures are illustrated in Fig. 4. As the condensation temperature increases, the strong solution mass fraction decreases, while the absorber power increases. As a result of the increase of condensation temperature by 20°C, the strong solution mass fraction increases by an average of 6.8% when the generator operates on the exhaust gas heat source, compared to the jacket cooling heat source. This is due to the increased exit generator temperature from 80°C to 90.5°C. On the other hand, the absorber power decreases by approximately 0.7% when the solution mass fraction increases by 6.8% at the same condensation temperature. This is due to the increased strong solution mass flow rate from 0.237 kg/s to 0.30 kg/s.

In case of the evaporation temperature is increased, while maintaining both T_c and T_a at 30°C, the mass fraction of the absorber power and the weak solution mass fraction decrease. At the same time, the strong solution mass fractions are kept at the same levels, as shown in Fig. 6. The decreasing power and mass fraction take their lowest values when the evaporation temperature reaches its condensation temperature. In addition, the mass flow rates of both the strong and the weak solution start with high values, which then decrease by 55.77% and 67.09%, respectively, at the end of the considered evaporator temperature range, for the generator exit temperature of 80°C. These reductions are equal to 46.65% and 60%, respectively, when the exit generator temperature is increased to 90.5°C. These results also indicate that the absorber power for the jacket cooling water and the exhaust gas heat sources is nearly the same at high evaporator temperatures, starting from 15°C.
ENVIRONMENTAL AND ECONOMIC RESULTS

Practically, using ARU will lead to savings in the required generator power, and consequently to the reduction of both fuel consumption and emissions. For the examined case study, using ARU will save fuel consumption by 104 ton/year, at the fuel price of 720 $/ton [37]. This saving will result in dispensing with one of the diesel generators during ship cruise. The emission factors for the high-speed diesel engines fueled with marine diesel oil (MDO), with 1.0% sulfur, are respectively equal to 10.81 g/kWh, 4.1 g/kWh, 0.3 g/kWh, 645 g/kWh, and 0.2 g/kWh for NOx, SOx, PM, CO2, and HC emissions, during cruise [38, 39]. Based on these factors, the emission rates in kg/min during trip for each main engine can be calculated as shown in Fig. 7. For the case study, the SOx and NOx emission rates are 0.3936 kg/min and 1.038 kg/min, respectively, during cruise mode for each main engine. These rates should be compared with IMO 2020 and IMO 2016 (Tier III) rates of 0.24 kg/min and 0.287 kg/min, respectively. The same can be done for the maneuvering and stand-by modes. From Fig. 7, both NOx and SOx emission rates of the high-speed diesel engine will not be combatable with the new IMO emission limits during ship cruise. This highlights the importance of using ARU to meet the stringent IMO regulations.

![Fig. 7. Comparing emissions of case study vessel with IMO limits](image)

Environmental benefits of the use of ARU unit become clear when comparing the yearly emission reduction in ton/year after applying the ARU unit, as shown in Fig. 8. The yearly emissions from the main engine are 102.7 ton/year, 38.87 ton/year, 2,981 ton/year, 6115 ton/year, and 1,987 ton/year for NOx, SOx, PM, CO2, and HC, respectively. Applying the ARU unit during cruise mode will reduce these emissions down to 4.32 ton/year, 1.64 ton/year, 0.12 ton/year, 285 ton/year, and 0.08 ton/year, respectively.

![Fig. 8. Yearly emission reduction after applying ARU](image)

Economically, the application of the ARU unit can be judged from the annual installation costs and its recovery period. The capital cost of the ARU unit ranges from 500 $/kW to 700 $/kW, with the installation cost amounting to 12% of this cost. The annual operating and maintenance costs are 8 $/kW and 0.008 $/kWh, respectively [13, 27, 40]. For the application of the carrier absorption unit (16LJ11), the capital and installation costs are $218,817 and $206,976, at yearly fuel saving of 52,000 $/year, for the exhaust gas and jacket cooling water operated ARU units, respectively. The heat exchanger cost is the added cost only for the exhaust gas operated ARU unit, and not for the jacket cooling water operated unit. It adds extra expenses to the total ARU cost, but it improves the performance of the generator, compared to direct use of exhaust gas [13]. The initial heat exchanger cost is $10,572 which presents 5.7% of the capital cost of the exhaust gas operated ARU unit. The total lifecycle cost (LCC) of the two ARU units depends on the ship age and the working years remaining after the installation. Fig. 9 shows the LCC for the exhaust gas and jacket cooling water operated ARU units over 18 years of operation, assuming the average ship age of 28 years [6, 41]. The capital cost of the jacket cooling water operated ARU unit is $184,800, while that of the exhaust gas operated ARU is $195,372, representing 57.38% of LCC for this unit. For the jacket cooling water operated ARU unit, the operating and maintenance costs are $38,016 and $76,032, respectively, representing 11.8% and 23.6% of LCC. Therefore, evaluating the total LCC will help making decision about purchasing an ARU unit. The decision will be based on the unit which minimizes energy and maintenance costs.

![Fig. 9. Total lifecycle cost elements for waste-heat operated ARU units over remaining 18 years of operation](image)
In addition, the economic decision for applying the ARU unit should consider the time required for money recovery. Fig. 10 shows the yearly fuel saving cost and the annual cost for capital recovery with payback periods for the exhaust gas operated ARU unit. The payback periods should be compared with the available economic life of the ship. For the case study, the annual costs for capital money recovery are 54,836 $/year and 51,870 $/year for the exhaust gas and jacket cooling water operated ARU units, respectively, at i = 10%, and the payback period of 6 years. Moreover, the fuel saving cost has to be considered in order to evaluate the total economic benefits of applying the ARU unit onboard the case study ship. The annual fuel saving cost at the end of 18 years of ship operation will amount to $127,478, at 3% fuel price increment.

Fig. 10. Annual capital recovery and fuel saving cost after applying the exhaust gas operated ARU unit

On the other hand, calculating the annualized cost-effectiveness for each reduction in emissions will show extra eco-environmental benefits of applying the ARU unit for the main engine. Fig. 11 evaluates the cost-effectiveness of each emission reduction. The lower the annual emission cost-effectiveness, the higher the economic benefit of reducing the emission parameter. The effectiveness costs of reducing SOx emissions are 20.39 $/kg and 19.25 $/kg $/kg for exhaust gas and jacket cooling water operated ARU units, respectively, as shown in Fig. 11. NOx emissions will be reduced with cost-effectiveness of 7.73 $/kg and 7.30 $/kg, respectively. The most economic cost-effectiveness option for applying the ARU unit is for the reduction of CO2 emissions, which are reduced by 258 ton/year, with cost-effectiveness of 0.13 $/kg.

Fig. 11. Annual NOx, SOx and CO2 emission reduction cost-effectiveness after applying ARU unit

CONCLUSIONS

The performance of a water-lithium bromide absorption cooling machine installed onboard a ship has been analyzed. As a case study, a Ro-Pax vessel operating in the Red Sea area was investigated. The above performance was discussed from thermodynamic, environmental, and economic points of view. The main conclusions can be summarized as follows:

- With reference to the thermodynamic analysis, both exhaust gas and jacket cooling water heat sources can provide the required heat for the ARU generator to cover the required cooling capacity of 200 kW.
- From the environmental point of view, the application of ARU machine will decrease fuel consumption by 104 ton/year. This will reduce the NOx and SOx emissions by 4.324 ton/year and 1.64 ton/year, respectively. The highest reduction will be in CO2 emissions, by 258 ton/year.
- According to the economic results, using exhaust gas and jacket cooling water waste-heat sources for ARU machines will annually save $127,478 at the end of the expected ship lifecycle, at 3% fuel price increment. Based on the total lifecycle cost analysis, the jacket cooling water operated ARU unit is more economical than that operated with exhaust gas, at capital and installation costs of $218,817 and $206,976, respectively. This will lead to the reduction in SOx, NOx, and CO2 emissions with cost-effectiveness of 20.39 $/kg, 7.73 $/kg, and 0.13 $/kg, respectively. Both analyzed systems can be considered an economic option for newly built ships or currently operated ships with available payback period of six years.

ACKNOWLEDGMENT

Special thanks are due to all the companies that provided essential information, especially MARIN-TEK, MTU. In addition, many appreciations are due to staff of Alkahera Company for Ferries and Maritime Transport for their help, especially Eng. Tharwat Assad, the technical manager for his help.

REFERENCES


**CONTACT WITH THE AUTHORS**

_Naderr Ammar_

_e-mail: eng.naderr@gmail.com_

Department of Naval Architecture and Marine Engineering
Faculty of Engineering
Alexandria University
21544 Alexandria
**EGYPT**

_Ibrahim Sediek_

_e-mail: prof.isbrahim@gmail.com_

Department of Marine Engineering Technology
Faculty of Maritime Transport & Technology
Arab Academy for Science, Technology & Maritime Transport
1029 Alexandria
**EGYPT**
CONCEPT OF COMPUTER-AIDED ASSESSMENT OF THE TECHNICAL CONDITION OF OPERATED GAS TURBINE VANES

Mariusz Bogdan¹
Marcin Derlatka¹
Józef Blachnio²
¹ Bialystok Technical University, Poland
² Air Force Institute of Technology, Poland

ABSTRACT

The article presents a multi-stage algorithm for automatic (without any human, user – diagnostician intervention) detection of vanes-blades (technical objects) and their surfaces on a digital image, combined with color analysis, aimed at determining the technical condition of the tested turbine elements. The images recorded with the use of a camera, containing previously dismantled from the turbine operated stator vanes, were used as the exemplary analysis material. The paper presents the algorithm for the detection of the vanes' airfoil surfaces with the impact of the applied techniques and methods of image processing and analysis on the final result (software localization of the vane's trailing and leading edge). Then, the obtained image data, including the structural changes of both the vane's coating and material (metallographic testing) were correlated with the surface colour scheme (colour segmentation based on the YCbrCr colour space model). Thanks to this approach, areas on the surface of the blade were distinguished, characteristic for proper, overheated and transient condition.

Keywords: gas turbine, condition assessment, vanes (blades), coating, surface colour

INTRODUCTION

While operating a turbine engine, no matter whether an aircraft, a marine or a traction one, various failures to engine assemblies occur. The predominant majority of failures to gas turbine vanes (blades) are effected with inappropriate operation (misadjustment) of subassemblies mating with the turbine, first of all, the combustion chamber [8] and, like with turbines of aircraft turbojet engines, the exhaust nozzle (in particular, the mechanism to adjust nozzle-mouth cross-section) [1, 12]. The detection of gas turbine subassembly damage during operation is possible during periodic inspection by way of assessing the condition of blade and vane rings' surfaces, their mounting elements and interstage turbine sealings etc. [7, 13, 22]. In multi-stage turbines, the visual assessment of the first turbine stage is relatively easy to carry out when the combustion chamber assembly is disassembled. The intermediary stages and the last stage are inspected without the need of disassembly, by using fibre optic viewing devices (videoscopes) [11, 15]. The damage of initial turbine stages are usually caused by the impact of high temperature exhaust fumes, whereas the damage of the last stages (with the longest blades) – by the impact of mechanical load (vibration, centrifugal force) [10, 22]. The frequent cause of damage is material overheating and thermal fatigue of the thermal fatigue of the vanes’ nozzle apparatus and rotor blades, resulting from both the excessive temperature and the time the blade (vane) is exposed to high temperature and the chemical activity of exhaust fumes (the entire operation process features a change in the colour of the vane-blade leaf) [19]. The minimisation of the safety hazard of engine operation requires periodical inspection of the vanes'/blades' technical condition. Overhauls and inspections feature blade checks using classic NDT (non-destructive testing) methods, i.e. visual inspection, fluorescent, magnetic and eddy current flaw detection [5, 16, 18], or rapidly developing measurement techniques related to thermography
and tomography [3, 14]. Each of the aforementioned NDT methods is executed in specific conditions and the information obtained regards the tested turbine element’s flaw and damage types specific (characteristic) for each method. Furthermore, their execution requires extensive theoretical knowledge and experience of the diagnostician, because the diagnostic information acquired by applying them is often encumbered with certain complications of result interpretation and limited reliability when compared to destructive methods [19].

The overheating of the blade material, especially of uncooled blades, constitutes damage that is most difficult to identify and technical condition most difficult to classify. Thus far, only the eddy current method gives certain capabilities of detecting the degree of blade material overheating, but it is strongly dependent on many factors, including, among others: type of blade material and protective coating, applied measurement instrumentation and parameters, and it has not yet been confirmed on an adequate statistical sample [14]. Therefore, there is still a lack of reliable method of detecting structural changes after the impact of high temperature exhaust fumes. Despite the fact that the destructive methods (e.g. metallographic method) provide much more information about the structure of the tested turbine element’s material, they make its further use impossible, thereby increasing the overall costs.

Progressive damage cause by material overheating leads to faulty gas turbine operation and sometimes, as in the case of aircraft jet motors, to tragic accidents. This type of damage is always removed by overhauling the motor, which in case of aviation results in very high costs of approx. PLN 2-5 million, depending on the motor type. Therefore, the costs of motor overhaul are 1.000-3.000 times higher when compared to the value of a turbine rotor blade, for example. The decision about the need of motor overhaul is currently made by the diagnostician, who can diagnose the condition of hard to reach turbine elements by applying the visual method with the use of a videoscope. The condition assessment is executed based on the recorded image of the diagnosed element’s surface and by comparing the image with the model surface images of fit and unfit analogous turbine blade elements. Such condition assessment criteria are very subjective, because they depend on the diagnostician’s state of knowledge and sight. The diagnostician’s decision can be verified using the destructive method. The tested element undergoes analysis of the micro-structure based on a metallographic sample. Mistakes of the diagnostician’s subjective assessment can lead to deeming an overheated blade as fit or not overheated as unfit. The first case results in an aerial accident within a short time of motor operation, whereas the second case – in very high motor overhaul costs.

This paper proposes a computer assisted method which allows a non-invasive assessment of the blade condition based on their surface colour. The first stage described in the article concerns the detection and positioning of the surface strictly related to the blade’s leaf area, whereas the second stage concerns the colour analysis correlated with the microstructure testing of blades in various technical conditions (degree of overheating).

SUBJECT OF TESTING

AND RECORDING OF IMAGE DATA

The digital images (photos) of gas turbine vane (blade) surfaces can be recorded in two ways:

- after the blade (vane) disassembly from the turbine; the recording device can include any digital camera with suitable repeatability and resolution of the obtained images;
- without blade (vane) disassembly; the recording device can include, e.g. a videoscope – devices for acquiring photos of hard to reach places.

The input images included photos of guide vanes disassembled from a gas turbine of an aircraft jet motor (Fig. 1). The vanes were made from the ZS-6K alloy. An important technological aspect in these type of alloys is the thermal processing, which is based mainly on homogenising annealing (unification of the structure, increase in strength and ductility) [20]. The purpose of the processing is also obtaining the adequate dispersion and shape of precipitation of phase γ’ as the main reinforcement phase. The coating is applying to allow increasing the working temperature, thereby additionally protecting the base material against harmful impact of the high temperature working medium (exhaust fumes). Figure 1a presents an exemplary set of recorded images of turbine vane surfaces with various degrees of overheating (acc. to the currently applied classification of their technical condition) with marked trailing and leading edges.

The input images were saved in the JPEG format with a low compression ratio preserving all necessary information and details about the recorded objects. In the case of this type of image data recording, it is possible to ensure high repeatability by determining fixed acquisition conditions (selection of a suitable light source and even lighting of the photographed surfaces, no reflections, etc.) – Fig. 1a. On the other hand, photos made using videoscopes (working conditions), due to the minimisation of the capture device’s size and spot light sources, are encumbered by geometric distortion and colour corruption. Furthermore, the viewing device is movable and unstable, thus preventing exposure repeatability and even lighting of the tested elements (Fig. 1b, c). The fact that the tested element does not require disassembly from the turbine is undoubtedly an advantage of diagnostics using videoscopes.

**Fig. 1. Changes in vane surface colour recorded using various devices:**

- a) image recorded at the measurement station using a camera;
- b) image recorded using the Olympus Iplex SA II videoscope;
- c) image recorded using the Everest XLG3™ VideoProbe videoscope
DETECTION OF VANES AND THEIR SURFACE ON A DIGITAL IMAGE

This part of the paper will present subsequent steps of executing the automatic object (vanes – Fig. 3a) detection and positioning of the detected vane leaf surfaces based on images presenting optionally placed vanes with various technical condition. The execution of this task involves no human (user – diagnostician) intervention. The block diagram of execution of subsequent stages of the selected image processing and analysis techniques and methods is presented in Fig. 2.

The input image (the example of such an image is presented in Fig. 3a) is firstly subjected to initial binarisation with the binarisation threshold selected using the iteration method, in accordance with the following steps [21]:

1. The extreme pixels (areas with the dimensions of 10x10 pixels selected on the photo’s edges) are adopted as the background, while the other part of the image is adopted as the object. This served as the basis for designating the brightness of the background \( \mu_t \) and object \( \mu_o \). The threshold value is adopted in accordance with the following formula:

\[
T^{(0)} = \frac{\mu_t^{(0)} + \mu_o^{(0)}}{2}
\]  

2. Designate the n average value for the background and object, while maintaining designated threshold, the pixel division to background and object
3. Designate a new threshold value:

\[
T^{(n+1)} = \frac{\mu_t^{(n)} + \mu_o^{(n)}}{2}
\]  

4. Stop the algorithm when \( T^{(n+1)} = T^{(0)} \), otherwise proceed to point (2).

where:
- \( T^{(0)} \) – threshold value designated based on the initial assumptions,
- \( \mu_t^{(0)} \) – average value of the background designated based on the initial assumptions,
- \( \mu_o^{(0)} \) – average value of the object designated based on the initial assumptions,
- \( T^{(0)}, \mu_t^{(n)}, \mu_o^{(n)} \) – values of the binarisation threshold, background average, object average, respectively, designated in subsequent algorithm iterations,
- \( n \) – number of algorithm iterations.

The initial binarisation allows for separating the object pixels from the background pixels, however using it caused a large number of artefacts to be left on the binary image (Fig. 3b). They were removed by using the morphological function of binary image opening (with the structural element in the form of a disk with the size of 10 pixels); the effects of this operation are presented in figure 3c. An additional advantage of executing the opening operation will be the separation of possibly combined or interfacing vane surfaces. The next step featured the indexation (labelling) of objects (vanes), thereby reducing the number of small objects that did not belong to the vane leaf’s surface (mainly elements of the vane’s lock) – Fig. 3d. As result of the executed operations (Fig. 3c and d), most artefacts, left from image binarisation (Fig. 3b), were removed, thereby preventing their impact on the dimensions of the designated surfaces (Fig. 3e). In figure 3e, the red colour was used to mark the smallest rectangles that completely surrounded particular detected objects (bounding box).
Due to the fact that the vanes in the digital image can be positioned in any configuration (position, number), the detected objects – surfaces (each rectangle from Fig. 3e is a new image cut from the original image in Fig. 3a) must be placed (positioned) in the same way (identical position of TE and LE).

Therefore, the image fragments that include the detected surfaces were rotated, depending on the vane's position, by a suitable (designated) angle, thereby bringing all of them to a common reference frame. The general algorithm of automatic rotation is presented in Fig. 4. The position of the TE and LE for each vane was checked in the first step. Then, certain (selected) geometric features of the tested turbine elements were used to check the width at two positions: at 1/4 and 3/4 of the vane leaf's length (Fig. 5a). The dimensions comparison operation was executed on a binarised image fragment (Fig. 5b) – the sum of the values is designated in the right (dimension d2) and left (dimension d1) column. If d2 < d1, the entire image is rotated by 180°. Thus, we obtained images on which vanes are positioned as assumed, i.e. TE upwards.

Then, the vanes were rotated towards the horizontal plane in relation to the TE (for this purpose, two points belonging to it were designated). This was done by applying the gradient method of designating the object's edges (with Sobel's operator) and using the standard Hough’s transform – SHT) [6]. The designated edge was marked on the photo fragment using a section (Fig. 6).

Despite the fact that the Sobel's operators (which allow for detecting vertical hS1 and horizontal hS2 lines – formula 3) are built based on Prewitt's operators, however in this case they strengthen the impact of the nearest pixel surroundings, for which the gradient is designated, resulting in obtaining more apparent contours than in the case of Prewitt's and Roberts' operators [9]:

\[
\begin{bmatrix}
-1 & 0 & 1 \\
-2 & 0 & 2 \\
-1 & 0 & 1 \\
\end{bmatrix}
\]

The automatic detection of the vane edges (straight lines) for particular fragments of the original image (new images that include only a single vane – object) was conducted using the SHT, based on the dependencies resulting from the SHT and on the points belonging to the section (Fig. 6), we designated the straight line's directional coefficient and the rotation angle α:

\[
a = \frac{y_2 - y_1}{x_2 - x_1}
\]

\[
α[°] = \frac{180 \cdot α[rad/s]}{π}
\]

After changing the vanes' configuration and rotation, we obtained images on which the vanes are positioned in an identical way. Then, while maintaining only the vane surface pixels (without the background pixels), it was necessary to cut the rectangles with the largest area possible from the images. Due to the fact that all fraction points were to be located within the area of the vane surfaces, we conducted an image binarisation with the threshold selected using the iteration method. Then, we used the morphological erosion method (with a square structural element with a 20 pixel size, thereby “reducing” the vane's surface) on the binary image. This operation allowed for further application of the gradient (Sobels) operators and Hough's transform in order to designate the coordinates of two points belonging to the top (position of the edge on the photo) vane edge (blue section in Fig. 7).

Thus, we were able to designate the straight line limiting the area later cut from the image for further analysis from the top (top red straight line in Fig. 7). The applied steps of
DIAGNOSING AND FORECASTING THE FAILURE OF OPERATED VANES

The turbine vane (blade) durability is a sum of many factors, in which the material plays the crucial role. In terms of the material, durability can be defined as the time of operation, during which the alloy properties imparted during the manufacturing process remain stable. The stability of the properties (assumed working resource) is determined during the designing stage by selecting suitably high features (in comparison to the predicted loads and by analysing their changes during operation). The gas turbine vane (blade) destruction process starts with the deterioration of the aluminium coating (presented in the surface images in the form of colour scheme changes). The latter part of the paper will demonstrate the relation between the surface colour changes and its technical condition defined as the material criterion – material overheating (condition of vane coating, material – phase γ') [4, 23, 24]. The method of procedure when developing the concept of computer-assisted assessment of the technical condition of used gas turbine vanes included several stages related to the recording of image data, selection of suitable colour representation (colour space), approach to notions related to colour analysis and metallographic testing. In the first step, various sets of used vanes were recorded with a digital camera on a specially designed station. The vane sets are a set of objects that differ in terms of their: number, position, technical condition (also resulting from their time of operation – vanes with increasingly longer times of operation). Then, in order to determine the impact of high temperature exhaust fumes during operation on the degradation of the vane surface and the micro-structure of the ZS-6K super alloy, we conducted the metallographic testing. During the testing, the vane was divided into over a dozen parts, we conducted several measurements in locations of their most severe wear and colour profiles were drawn along the vane cutting lines. The micro-structure was examined under the Neophot optical microscope and the (SEM) Hitachi S-3000N scanning microscope. We characterised the vane deterioration process based on the metallographic testing. In the case of vanes permitted for exploitation (new vanes), the aluminium-silicone surface layer is evenly distributed around the entire vane circumference. Furthermore, the layer consists of two sub-layers – external sub-layer with increased content of aluminium and intermediate layer with increased content of silicone, which also includes increased content of carbide formers (especially molybdenum and tungsten).

After analysing the technical condition of the coatings of used vanes from a turbine’s jet crown (Fig. 9), it was established that after some time of operation the vane’s coating does not degrade (Fig. 9a – vane no. 1 of fig. 8a) and its thickness does not differ from the thickness of a new vane’s coating. This condition of the coating is characterised by lack of any damage. After a longer time of operation, the coating undergoes expansion and features local surface damage (Fig. 9b – the condition was observed incidentally on vane no. 3 from figure 8c, its share increases in vanes no. 4, 5, 6). This condition of the coating is classified as partially fit for further use.
The expansion of the coating causes its delamination and cracking in the latter period of operation (Fig. 9c – the condition was incidentally observed on the leading edge of vane no. 5 from figure 8e, while the biggest share of the condition was ascertained in vane no. 6 from figure 8f). This type of damage causes local separation of the surface layer and cracking of the transitional layer diffused into the native material of the vane. In effect, within a short time of operation, the surface layer comes off and the crack penetrates deep into the vane’s superalloy (Fig. 9d – vane no. 6 from figure 8f, mainly the area affected by the highest temperature, i.e. 1/3 height of the vane’s leaf). Damage in the form of coating delamination, local cavities in the surface layer and development of cracks result in its classification as an element that is unfit for further use. In the case of the vane material’s micro-structure, it was ascertained that vane no. 6 is characterised by dispersive secondary separation of the phase γ’, resulting from the impact of high temperature exhaust fumes. The morphology of phase γ’ indicates that after exceeding the critical temperature, the alloy is overheated, and the turbine vane cannot be deemed as fit for further use [4, 23, 24].

Figure 10a presents the line that represents the averaged colour profile (considering the cutting thickness, cuts parallel to the leading edge; the dotted lines mark cuts that are perpendicular to the leading edge). Considering the relatively high correlation of data – particular constituents of the RGB model (Fig. 10b) [2], the first step was to seek for a suitable representation of the image data, which would clearly reflect the colour changes of the vane surfaces. Among many models, i.e. HSV, CIEXYZ, CIELab, CMYK, etc., the largest differentiation in the image data (no correlation between particular constituents) was obtained for the YCbCr colour space model (Fig. 10c), the particular constituents of which, i.e. Y – luminance constituent, Cb – Y-B chrominance differential constituent (constituting the difference between luminance and blue colour) and Cr – Y-R chrominance differential constituent (constituting the difference between luminance and red colour), were designated based on the following conversion [17]:

\[
\begin{bmatrix}
Y \\
Cb \\
Cr
\end{bmatrix} = \begin{bmatrix}
0.229 & 0.587 & 0.114 \\
-0.168 & -0.331 & -0.500 \\
0.500 & -0.418 & 0.082
\end{bmatrix} \begin{bmatrix}
R \\
G \\
B
\end{bmatrix}
\]

The parameters describing the constituents of the YCbCr colour space model were designated based on vane no. 1. This vane is characterised by uniform colour and during the metallographic testing it was deemed as fit for use (correct technical condition, slightly deviating from a new vane). The method of procedure in relation to other vanes was identical to the vane presented in figure 10 (Fig. 10a – dotted and continuous line – vane cutting, metallographic testing, drawing of colour profiles). Based on the exemplary results specified in figures 10b and 10c it was ascertained that the waveform of constituent Y (luminance), also in the case of vane no. 6 (multi-coloured surface), has a similar waveform (of changes) as constituents R, G, B presented in figure 10b.

On the other hand, the chrominance constituents: Cb, Cr demonstrate a different waveform characteristic. They were thus used to describe the colour changes taking place on the vanes’ surfaces. In the case of vane no. 1, except for constituent Y (average value of Y_śr equal to 136, standard deviation – SD equal to 10.1), the variability of Cb, Cr, drawn in a few colour profiles was minor (for Cb: Cb_śr1=116, SD=2.2, for Cr: Cr_śr1=140, SD=1.4). The sequences of obtained Cb, Cr data included no values that would disrupt the designation of the average values. Figure 10c additionally presents the position of the average values describing the surface of vane no. 1 (Cb_śr1, Cr_śr1).

Based on the values of constituents Cb, Cr for vane no. 1 and the nature of the waveform of these constituents in the colour profiles of vane no. 6, we designated the threshold values describing the non-overheated surface. The threshold value was designated for constituent Cr in the range from 140 to the maximum value, i.e. 255 (the other two constituents were not limited) – Fig. 11a

The result of the surface segmentation of vane no. 6 with limitation of constituent Cr is presented in figure 11b (the area differing from the correct surface was dimmed by applying an operator). On the other hand, the range of the constituent Cb was modified in figure 11c (value range from 0 to 115). The result of the segmentation (limitation of the range of constituent Cb) and marking of the overheated area...
is presented in figure 11d. On the other hand, figure 11e presents the resulting image with two limitations of the range of constituents Cb, Cr.

The surface with the correct condition in the case of vane no. 6: Fig. 11b amounts to 11.34% of the total surface, Fig. 11d – 9.89%, Fig. 11e – 9.02% (double limitation slightly changes the percentage share of the surface).

Figures 12a, c, e present the results of segmentation, including the surface classified as satisfactory for the exemplary set of vanes from figure 8.

Nearly identical results apply to figures 12d and 12f. The value fluctuations do not exceed 0.3%, whereas the largest difference of 4.03% applies to vane no. 5 (12d – 27%, 12f – 31.03%). The differences in figures 12a, c, e result both from the colour scheme of the tested surfaces (vane no. 1, 2 “shades of yellow”) and from the adopted method of determining the threshold values (average). The largest percentage share of surfaces classified as satisfactory, for particular vane numbers, was obtained with the limitation of constituent Cr (no. 1 – 91.3%, no. 2 – 94.25%, no. 3 – 85.93%, no. 4 – 35.54%, no. 5 – 37.94%, no. 6 – 11.34%) – Fig. 12b. The area that does not belong to the surface classified as satisfactory (black) also features the surface in transition (between satisfactory surface and overheated surface). It is characterised by a certain ambiguity due to the condition of the coating, the thickness of which increased (swelling of the surface layer and transitional layer), but the average thickness, due to the existence corrosion and erosion pitting, is lower than in the case of vanes no. 1, 2, but higher than in vane no. 6 (Fig. 9b). The structure (size, shape and distribution of separations of phase γ) deep into the vane material, only in extreme cases, at the transitional layer, slightly deviates from the model structure. The colour scheme of this surface is characterised by a different colour. Based on the metallographic testing combined with a colour analysis of the discussed type of surfaces, we determined the threshold values of segmentation for particular constituents, i.e. Y<58.160>, Cb<112.119>, Cr<0.138>. The result of the segmentation is presented in figure 13a (orange – areas of surface in transition).

The share of surfaces in transition does not exceed (vanes no. 1÷no. 6) 20% (Fig. 13b). The highest share concerns vane no. 5 and amounts to 16.5%, whereas for vanes no. 1 and no. 2, the share is miniscule and amounts to 3.25% and 2.5%, respectively. A similarly low value applies to vane no. 6 (4.1%). Figure 13c presents the share of the area (surface), the condition of which deviates from the satisfactory condition (the share of the surface in transition was also taken into consideration).
CONCLUSIONS

The paper proposes the concept of computer-aided, non-invasive method of assessment of the condition of operated vanes (blades) based on the colour of their surface. The described approach embraces several stages. The first stage concerns the detection of the surface (strictly determined and related to the area of the blade's leaf), which is then subjected to colour analysis in the next stage. The first stage concerns the detection of the surface (strictly determined and related to the area of the vane's leaf), which is then subjected to colour analysis in the next stage. The verification of the technical condition of used vanes was conducted using a destructive method on metallographic samples. The metallographic testing concerned the assessment of the coating's condition and the characteristics of the separation of phase γ' (both in the vane's surface layer and in its material). The material criterion was correlated with the colouristic changes on the vanes' surface. Based on the colour profiles, we were able to determine the threshold values describing particular colour ranges (histograms of constituents of the YCbCr model) representing an overheated, non-overheated vane and vane in transition. The most satisfactory results were obtained in the case of determining the range of values for constituent Cb (the determined range of constituent Cb includes the largest colouristic differentiation of the tested surfaces with simultaneous maintenance of the correctness of the obtained results – this thesis was confirmed by metallographic testing). Based on the results of the analysis of the vane surface images, it is possible to conclude that vanes no. 1, 2 and 3 are satisfactory, because the percentage share of the surface classified as satisfactory and in transition (with minimum share of this condition amounting to 3.25%, 2.5%, 6.25%, respectively) fluctuates around 95% (Fig. 13c). On the other hand, vanes no. 4, 5, 6 are overheated, because the percentage share of the overheated surface and surface in transition (in the case of vanes no. 4 and 5, with a substantial percentage share amounting to 13.87% and 16.5%) substantially differs from the aforementioned vanes, and the overheated surface for particular vanes is as follows: no. 4 – over 50%, no. 5 – over 45%, no. 6 – over 80% (Fig. 12a, 13b, c).

In some cases, the diagnostics of the condition of vanes (blades), aside from the strict classification of “element fit for use – element unfit for use”, features a third assessment of the condition, referred to as a “partially fit element”. This condition requires periodical diagnostics (the vane (blade) undergoes periodical diagnostics until it is unfit for further use). In the case of the proposed assessment of the technical condition, it is possible to deem a vane (blade) “with a certain share” of transition and miniscule share of surface deemed as overheated as a “partially fit element” (it, however, requires further testing; further specification). Furthermore, the results presented in figure 13b, providing information concerning the surface area in transition can contribute to the determination of the time within which the given object will conditionally maintain its properties (observation of the vane deterioration process during periodical diagnostics). Therefore, the development of a surface area in transition or its complete transition into an unfit condition can constitute prerequisites for estimating the assumed and actual conditions of vane (blade) operation, particularly since their time of operation increases, thereby leading to the vanes’ critical condition in the estimated range of load.

REFERENCES


CONTACT WITH THE AUTHOR

Mariusz Bogdan
e-mail: m.bogdan@pb.edu.pl
Bialystok Technical University
Wieska 45
15–333 Białystok
POLAND

Marcin Derlatka
e-mail: m.derlatka@pb.edu.pl
Bialystok Technical University
Wieska 45
15–333 Białystok
POLAND

Józef Blachnio
e-mail: jozef.blachnio@itwl.pl
Air Force Institute of Technology
Kościuszki 8
01–494 Warszawa
POLAND
MODEL OF A DUCTED AXIAL-FLOW HYDROKINETIC TURBINE – RESULTS OF EXPERIMENTAL AND NUMERICAL EXAMINATION

Adam Góralczyk
Adam Adamkowski
Institute of Fluid-Flow Machinery Polish Academy of Sciences, Gdansk, Poland

ABSTRACT

The article presents the numerical algorithm of the developed computer code which calculates performance characteristics of ducted axial-flow hydrokinetic turbines. The code makes use of the vortex lattice method (VLM), which has been developed and used in IMP PAN for years, to analyse the operation of various fluid-flow machines. To verify the developed software, a series of model tests have been performed in the cavitation tunnel being part of IMP PAN research equipment.

Keywords: Vortex Lattice Methods; Ducted hydrokinetic turbines; Cavitation tunnel; Performance characteristics

INTRODUCTION

The process of electric energy generation in hydrokinetic turbines assumes the use of kinetic energy collected in water currents of rivers or other inland waterways, as well as in sea currents. Unlike conventional water turbines, hydrokinetic turbines do not require artificial water damming systems, as a result of which the operation of these devices disturbs the natural flow of the water current to a minimal extent.

Studying possibilities of the use of hydrokinetic energy began in the 1970s, after the energy crisis caused by the embargo imposed by OPEC countries on the USA and Western European countries [24]. Dramatic price increase of crude oil in those times, followed by similar price increase of coal and natural gas, led to the stagnation and recession of the world economy. Countries importing energy-producing raw materials were forced to intensify the research oriented on modernisation of technology of energy production from conventional sources and search for new energy sources. At present, a basic motivation for taking actions which aims at reducing the percentage of mining resources in the entire energy production in favour of ecologically clean technologies based on renewable energy sources is environmental protection. That is why great interest towards hydrokinetic turbines is observed, and wide-ranging studies of devices of this type are performed worldwide. The overview of existing commercial designs of systems which utilise hydrokinetic energy to produce electric current, and their classification can be found in [18, 24, 29].

Research studies upon hydrokinetic turbines are performed in two ways: (1) as theoretical analyses, based on computational fluid dynamics and making use of relevant computer codes, and (2) as experiments, performed both on laboratory scale on research rigs, and in real, natural conditions. In the former case a number of calculation methods are used. One of the most popular methods used in the fluid-flow machine design and optimisation process is the Blade Element Momentum (BEM) method. It was initially developed by H. Glauert for airplane propellers [8], as a combination of the liquid flow theory (operating disc model) developed by W. Froude and W. J. Rankine [27] for marine propellers, and the model of blade segment developed, separately, by S. Drzewiecki and W. Froude [7]. A description of the BEM method applied to wind and hydrokinetic turbines can be found in [3, 13, 20, 23, 31].
Another group of methods used to analyse the operation of fluid-flow machines comprises methods which are based on surface distribution of singularities. Those methods apply the model of perfect fluid and assume the potential nature of the flow. As a consequence, the equations of motion are reduced to the Laplace equation, the solution of which is a superposition of elementary solutions representing the uniform flow and hydrodynamic singularities of vortex, source, and/or dipole type. A broad description of those methods, with their differentiation depending on the type of applied singularities and boundary conditions, can be found in [28]. This group of methods includes the Vortex Lattice Method (VLM), used in this publication, which is widely applied to analyse the operation of such fluid-flow machines as: marine propellers [21,22], centrifugal pumps [17], wind turbines [6,14], and hydrokinetic turbines [12,25,32].

In recent years, the increasing computing potential of newly built computers has provided good opportunities for the use of advanced and demanding, in terms of the required computer parameters, methods to study the flow through hydrokinetic turbines. These methods consist in solving the Reynolds Averaged Navier-Stokes (RANS) equations [16,32]. Very popular solvers, such as ANSYS/FLUENT, ANSYS/CFX, or NUMECA/Fine, which are used for this purpose make use of the finite volume method. These widely applicable solvers are used to perform comprehensive numerical analyses of issues related to hydrodynamics and compressible fluid mechanics, taking into account multiphase and transonic flows.

Along with theoretical-numerical analyses of the flow through hydrokinetic turbines, experimental investigations are also performed, mainly in laboratory conditions. Based on the available literature, two types of laboratory research rigs can be named. The first type is comprised of closed-circuit water tunnels [2,11]. In those tunnels, the water flow is generated by circulation pumps and the water flows in a closed space in which the examined model turbines are placed. In the other research rig type, the model turbines are placed in open water tunnels, or towing tanks (channels) with free water surface [4,15,30]. Experiments are also performed in real, natural conditions [5].

The object of numerical and experimental examination, the results of which are presented in this article, is a model axial-flow hydrokinetic turbine equipped with a duct as part of the flow system. The duct improves the energy efficiency of the turbine by increasing the velocity of the water flowing through the rotor. These solutions are sometimes used in hydrokinetic turbines [19,30] and wind turbines [1,26].

As compared to the earlier work by the authors [12], this article presents new elements introduced to computing procedures of the VLM based software developed in IMP PAN in recent years. The introduced elements mainly consist of:

- extending the VLM by a procedure which calculates the shape of the vortex wake downstream of the rotor [9],
- distributing the vortex lattice on both the suction and pressure sides of the rotor blade [9], instead of the average surface of the blade, a method which was used in earlier versions,
- introducing a numerical procedure which calculates the effect of walls bounding the measuring space of the cavitation tunnel on the obtained results of calculations [10].

As a result of the above actions, better compliance between the theoretically predicted performance characteristics of the examined model hydrokinetic turbines and those experimentally recorded in the IMP PAN cavitation tunnel was obtained.

**MODEL TESTS**

The examined physical model of hydrokinetic turbine consists of a 5-blade rotor with a hub, closed in a duct – see Fig. 1. According to the adopted assumption, both the rotor blades and the duct are 3D shaped.

![Fig. 1. The adopted physical model of hydrokinetic turbine](image)

The model turbine was examined in the IMP PAN water tunnel. The research rig is schematically shown in Fig. 2. The cross section of the measuring chamber has the square shape of size 0.425 m.

The measuring instrumentation of the rig enables to measure: the average velocity $V_\omega$ of the water flow in the tunnel chamber, the turbine rotor revolutions $n$, the torque $M$ on the turbine shaft, and the axial thrust $T$ generated by the turbine rotor. The average water velocity $V_\omega$ in the chamber is measured indirectly, using a calibrated tunnel confusor. A quantity which is measured directly for this purpose is the water head difference ($\Delta H$) between the confusor inlet and exit cross sections. The measurement is done using a precise differential manometer or pressure transducer, and the average water velocity is calculated using the formula $V_\omega = 4.677 \sqrt{\Delta H}$ obtained from calibration.

The turbine rotor rotational speed $n$ is measured using the pulse frequency meter which counts pulses generated by a disc with properly machined grooves.

The shaft torque and the hydraulic thrust of the examined model turbine were measured using dynamometers, installed in the rig, which were equipped with special balancing scales. In order to eliminate the effect of shaft resistance on the torque measured during the tests, an additional measurement was performed, which aimed at experimental determination of the dependence of the shaft line resistance torque on...
the rotational speed, in the case when the turbine rotor is dismantled. The model turbine torque which was then measured in the basic tests was corrected by the value of the shaft line resistance torque corresponding to the current rotational speed.

In case I the rotor was placed exactly in the inlet section of the duct, while in case II it was shifted by 5 cm towards the duct exit – Fig. 3.

The tests were performed for comparable parameters of water flow in the tunnel, with the average water flow velocity approximately equal to \( V_\infty = 3.4 \) m/s. The required rotational speed of the examined turbine model rotor was set using the electric current generator load control system.

![Fig. 2. Scheme of the research rig (cavitation tunnel) [11], 1 – turbine model, 2 – torque measurement, 3 – thrust measurement, 4 – measurement of pressure difference at measuring confusor inlet and exit, 5 – rotational speed measurement, 6 – circulation pump](image1.png)

The object selected for tests described in the article was a 5-blade rotor of \( D = 148 \) mm in diameter – detailed data in Tab. 1.

Tab. 1. Geometry of the examined rotor. Symbols: \( r/R \) – relative position of radial section, \( D = 2R \) – outer diameter of the rotor, \( c \) – profile chord length, \( p \) – blade pitch, \( t \) – maximal profile thickness, \( f \) – maximal profile camber

<table>
<thead>
<tr>
<th>( r/R )</th>
<th>( c/D )</th>
<th>( p/D )</th>
<th>( t/D )</th>
<th>( f/C )</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.175</td>
<td>0.1595</td>
<td>1.05</td>
<td>0.0392</td>
<td>-0.1207</td>
</tr>
<tr>
<td>0.3</td>
<td>0.1689</td>
<td>1.05</td>
<td>0.0351</td>
<td>-0.102</td>
</tr>
<tr>
<td>0.4</td>
<td>0.1797</td>
<td>1.05</td>
<td>0.0311</td>
<td>-0.0846</td>
</tr>
<tr>
<td>0.5</td>
<td>0.1926</td>
<td>1.05</td>
<td>0.027</td>
<td>-0.0682</td>
</tr>
<tr>
<td>0.6</td>
<td>0.2054</td>
<td>1.05</td>
<td>0.023</td>
<td>-0.0543</td>
</tr>
<tr>
<td>0.7</td>
<td>0.2155</td>
<td>1.05</td>
<td>0.0189</td>
<td>-0.0423</td>
</tr>
<tr>
<td>0.8</td>
<td>0.2257</td>
<td>1.05</td>
<td>0.0149</td>
<td>-0.0314</td>
</tr>
<tr>
<td>0.9</td>
<td>0.2318</td>
<td>1.05</td>
<td>0.0108</td>
<td>-0.0218</td>
</tr>
<tr>
<td>1</td>
<td>0.2338</td>
<td>1.05</td>
<td>0.0054</td>
<td>-0.0101</td>
</tr>
</tbody>
</table>

![Fig. 3. Duct geometry with marked sections I and II of rotor positions for two examined cases](image2.png)

The performed tests made the basis for calculating the below defined dimensionless coefficients:

**Water energy utilisation factor:**

\[
C_E = \frac{2\pi n M}{0.5 \rho V^2 \omega 0.25 \pi D^2}
\] (1)

**Thrust coefficient:**

\[
C_T = \frac{T}{0.5 \rho V^2 \omega 0.25 \pi D^2}
\] (2)

**TSR – tip speed ratio:**

\[
TSR = \frac{n D}{V_\infty}
\] (3)

The results of the performed tests, along with their uncertainty, are shown in the latter part of the paper (Figs. 9 and 10), and compared to the results of calculations obtained using the developed code.

## CALCULATION METHOD

The calculation method developed to analyse the flow through the examined hydrokinetic turbine model (Fig. 1), bases on the VLM, which, as mentioned above, is one of surface singularity distribution methods. It assumes that in the entire analysed domain the flowing fluid is incompressible and inviscid, except some singular fragments. After assuming that the flow is steady and at the same time without vorticity, \( \text{rot} \vec{V} = 0 \), the equations of motion are reduced to the Laplace equation:

\[
\Delta \phi(x, y, z) = 0
\] (4)

where \( \phi \) is the scalar function which bears the name of velocity potential and describes the velocity field as:

\[
\vec{V} = \text{grad} \phi
\] (5)
In the case of liquid flow through the examined turbine, the areas where singular points are located are: (1) boundary layers of solid elements (duct, rotor blades, hub) coming into contact with the fluid, and (2) vortex wakes forming downstream of these elements. In those areas, dominated by viscous forces, strong vorticity concentration is observed [17]. The adopted method consists in discrete distribution of the so-called horseshoe vortices in the abovementioned singular areas [22]. The horseshoe vortex consists of the bound vortex, free vortices, and trailing vortices modelling the vortex wake forming the downstream of the washed turbine elements due to the motion of the boundary layer – Fig. 4. This way, the continuous vorticity distribution in the flow is substituted by a discrete distribution of vortex filaments.

As mentioned in the introduction, certain modifications have been introduced to the initial version of the developed in-home code the results of which had been presented in [12]. Firstly, the vortex wake, modelled by trailing vortices, had initially been approximated by helices with the assumed pitch angle. Proper selection of this angle was essential for the correctness of the obtained results, and it required huge experience from the code user. Therefore a decision was made to determine the shapes of the trailing vortices in a numerical, iterative way. Justification for this approach is given further in the article, in the form of a comparison analysis.

The next innovation introduced to the earlier code version referred to the way of distribution of the vortex lattice over the rotor blades. Initially, the vortices had been distributed on the average blade surfaces. However, during attempts to modify the code it was found, based on comparison of numerical and experimental results of ship propeller tests [9], that the distribution of vortices over the suction and pressure side produce more realistic pressure distributions. This opinion became the motivation for introducing the above change.

The third essential modification was extending the algorithm with a correction taking into account the effect of the measuring chamber walls on the characteristics of the turbine model placed in this chamber during laboratory tests. This correction was introduced in the way described further in the article. Its introduction was necessary, as neglecting this effect had led to remarkable overstatement of model turbine performance, compared to the machine operating in unbounded space.

The developed calculation method neglects flow turbulence and some other phenomena, in particular those resulting from fluid viscosity, such as flow separation for instance.

In the presented calculation task, first the circulations \( \Gamma \) of the vortex filaments distributed over the surfaces of fluid-flow elements of the examined machine are to be determined. For this purpose the Neumann condition was used, which says that normal velocities to the surface of the washed body are zeroed. This condition can be written as follows:

\[
\frac{\partial \phi}{\partial n} = \vec{n} \cdot \text{grad} \phi = \vec{n} \cdot \vec{V} = V_n = 0 \quad (6)
\]

In the examined case this conditions is expressed as the sum of projections of the undisturbed flow velocity and the velocities induced by vortices of circulation \( \Gamma \) onto the direction normal to the wall at control points situated between the bound vortices:

\[
V_{ni} = \vec{V}_\infty \cdot \vec{\nu}_i + \vec{V}_T \cdot \vec{\nu}_i = 0 \quad (7)
\]

where:

- \( \vec{\nu}_i \) – normal vector at the control point on the wall,
- \( \vec{V}_\infty \) – relative velocity:
  - for non-rotating turbine elements (duct): \( \vec{V}_\infty = \vec{V}_0 \),
  - for rotating elements (rotor, hub): \( \vec{V}_\infty = \vec{V}_0 - \vec{\omega} \times \vec{R}_i \),
- \( \vec{V}_0 \) – fluid inflow velocity,
- \( \vec{\omega} \) – angular speed of the rotor,
- \( \vec{R}_i \) – vector connecting the \( i \)-th control point with the rotor rotation axis,
- \( \vec{V}_T \) – velocity induced by vortices at the control point on the wall, calculated using the Biot-Savart law:

\[
\vec{V}_T = \sum_{j=1}^{N} \xi_j \left( \frac{1}{4\pi} \int_{T_j} \frac{\vec{r}_{ij} \times \text{d}l_j}{r_{ij}^3} \right)
\]

The index \( i \) represents the sequential number of the control point, while the index \( j \) stands for the number of the horseshoe vortex. \( L \cdot W \) is the number of all vortices in the computational domain. The above equation is the starting point for constructing the equation system used to calculate circulations \( \Gamma \) of the horseshoe vortices. In the matrix form, this equation system can be written as:

\[
[Kn][\Gamma] = [W] \quad (8)
\]

where \([Kn]\) represents the impact factor matrix, \([\Gamma]\) is the vector of the searched vortex circulations, and \([W]\) stands for the free term vector, determined from formula: \( W_i = -\vec{V}_\infty \cdot \vec{\nu}_i \). The impact factor \( Kn_{ij} \) is interpreted as a projection of the velocity induced by the \( j \)-th horseshoe vortex with unit circulation onto the normal direction at the \( i \)-th control point. Its formula is:

\[
Kn_{ij} = \left( \frac{1}{4\pi} \int_{T_j} \frac{\vec{r}_{ij} \times \text{d}l_j}{r_{ij}^3} \right) \cdot \vec{\nu}_i \quad (9)
\]
The code assumes constant and uniform velocity of the fluid flowing into the turbine. As a consequence, the computational domain can be divided into characteristic segments, in which the circulations of the vortices situated at the same (corresponding) points are the same. This assumption leads to a remarkable reduction of the equation system and shortens the computational time. In this case, one main segment of the computational mesh is selected, and control points are placed within this segment. A sample single segment of the turbine flow system is shown in Fig. 5.

When applying this simplification, the impact factor formula in the equation system is to be modified to the following form:

\[
Kn_{i'j'} = \sum_{n=0}^{Z} \left( \frac{1}{4\pi} \int_{I_{n}}^{r_{n}} \frac{r_{n}^{2} \times d\Omega_{n}}{r_{n}^{2} (r_{n}^{2} / f)} \right) \cdot \vec{n}_{i'}
\]  

(10)

where \( n \) is the number of the segment with a single blade, \( Z \) is the total number of blades, \( i' \) is the sequential number of the control point in the selected main segment, and \( j' \) is the sequential number of the surface element on which the bound vortex is situated.

Once the circulations of the horseshoe vortices are known, the velocities tangential to the walls of the analysed flow element are calculated from the formula:

\[
V_{S(i')} = KS_{i'j'} \Gamma_{j'} + \frac{\vec{V}_{co(i')} \cdot \vec{\tau}_{i'} \pm \Gamma_{i'} \Delta x_{i'}}{2Dx_{i'}}
\]  

(11)

where \( \vec{\tau}_{i'} \) represents the tangent vector to the surface of the washed element at the \( i' \)-th control point, \( \Delta x_{i'} \) is the width of the surface element on which the bound vortex is situated. The sign “+” refers to the suction sides of the rotor blade surface, hub, and duct while the sign “−” refers to the pressure side of the rotor blade surface. The matrix of \( KS \) coefficients is calculated from the formula:

\[
KS_{i'j'} = \sum_{n=0}^{Z} \left( \frac{1}{4\pi} \int_{I_{n}}^{r_{n}} \frac{r_{n}^{2} \times d\Omega_{n}}{r_{n}^{2} (r_{n}^{2} / f)} \right) \cdot \vec{\tau}_{i'}
\]

Based on the Bernoulli equation, for the obtained tangential velocities the dimensionless pressure coefficient values are calculated as:

\[
CP_i = \frac{p_{i \infty} - p_i}{\frac{1}{2} \rho V_{i \infty}^2} = \left( \frac{V_{in}^2}{V_{i \infty}} \right)
\]  

(12)

where \( p_{i \infty} \) is the pressure in the undisturbed flow, \( p_i \) is the pressure at \( i \)-th control point, and \( \rho \) is the density of the liquid.

The obtained distribution of the dimensionless pressure coefficient is used to calculate the torque \( M \) of the turbine rotor and its axial thrust \( T \). These values are calculated from the following integral equations:

\[
M = Z \frac{1}{2} \rho V_{i \infty}^2 \int_{S} \left[ CP_i (n_Y \cdot r_Z - n_Z \cdot r_Y) + C_f \cdot (t_Y \cdot r_Z - t_Z \cdot r_Y) \right] ds
\]  

(13)

\[
T = Z \frac{1}{2} \rho V_{i \infty}^2 \int_{S} \left( CP_i \cdot n_X + C_f \cdot t_X \right) ds
\]  

(14)

In these equations \( S \) represents the area of the rotor blade surface, \( r_z, r_y \), are components of the vector connecting the rotation axis with the control point, \( t_x, t_y, t_z \), and \( n_x, n_y, n_z \) are components of the tangent and normal vector to the rotor blade surface, respectively. \( C_P \) and \( C_P \) are the dimensionless pressure coefficients on the suction and pressure side of the rotor blade, respectively, and \( C_M \) is the viscous drag coefficient, defined, according to [21], as:

\[
C_f = \left( 1 + \frac{1}{C_T} \right) \frac{0.455}{(log RE)^{0.5}}
\]  

(15)

The obtained results of calculations enabled to determine the dimensionless values of the water energy utilisation factor \( C_T \), thrust coefficient \( C_M \), and tip speed ratio \( TSR \), defined by the formulas (1), (2) and (3), respectively.

The experience gained when developing the software here described revealed high sensitivity of the obtained values of the above coefficients to the assumed shape of the trailing wake, which, as already mentioned, in the initial versions of the code had been approximated by helices with the assumed pitch angle. In the recent version, a decision was made to determine the shape of the trailing wake in an iterative way. For this purpose a procedure was developed, the basic concept of which is given in Fig. 6.

![Fig. 6. Scheme illustrating how the shape of vortex wake is determined [31].](image-url)
The calculation process begins with introducing the vortex wake in the form of helices with an assumed pitch angle to the computational domain. The shapes of these lines are approximated by a sequence of rectilinear line segments connected at nodes. In further iterations, the shape of the vortex wake is modified by new vortices shed sequentially from trailing points situated on particular turbine elements (rotor blades, hub, and duct). The coordinates of the vortex generation points, being the first nodes of particular vortex filaments, are kept constant in subsequent iterations, while the locations of the following nodes are found based on local flow velocity vectors, calculated as the sum of the undisturbed flow velocity and the velocities induced by the system of vortices having the shape determined in the previous iteration:

\[
\vec{V} = \vec{V}_\infty + \vec{V}_1
\]  \hspace{1cm} (16)

The calculation process for a single mesh node is shown in Fig. 8. It consists of two main steps. In the first step, according to formula (16), the local velocity \( \vec{V}_{i-1}^* \) is calculated at the earlier node \( P_{i-1} \), making the basis for determining the point \( P_i^* \) from the following formula:

\[
P_i^* = P_{i-1} + \vec{V}_{i-1}^* \Delta t
\]  \hspace{1cm} (17)

At this step, the local flow velocity \( \vec{V}_i^* \) is also calculated from (16). In the second step, the coordinates of the trailing wake node \( P_i \), are calculated using the average flow velocity vector obtained from the earlier calculated vectors \( \vec{V}_{i-1}^* \) and \( \vec{V}_i^* \) as:

\[
P_i = P_{i-1} + \frac{1}{2} \left( \vec{V}_{i-1}^* + \vec{V}_i^* \right) \Delta t
\]  \hspace{1cm} (18)

In order to improve the stability of the iteration process, the calculation procedure has been complemented by the relaxation formula having the following form:

\[
P_i^{s+1} = \varepsilon P_i^{s+1} + (1 - \varepsilon) P_i^s
\]  \hspace{1cm} (19)

where \( i \) is the node number, \( s \) is the iteration number, and \( \varepsilon \) is the weight coefficient, with the value of 0.5 assumed in the reported calculations.

Each time, for the newly determined shape of the vortex wake, the dimensionless values of water energy utilisation factor and thrust coefficient are calculated. When the values of those coefficients differ by less than 0.2% or oscillate around a constant value in successive iterations, the calculation process is terminated. The calculation algorithm described above, and the resulting developed code, can be used to calculate performance parameters of hydrokinetic axial-flow turbines working in unbounded space. Since the experimental examination of the model turbine was performed in the bounded tunnel space, the calculation method had to be properly extended to improve the compliance of the numerical results with the experiment. The method extension by a correction which takes into account the earlier neglected effect of walls is described below.

The algorithm used to calculate this correction bases on the image method [7]. In Fluid Mechanics this method is used for modelling the flow past a body (here: through a hydrokinetic turbine) at the presence of surfaces bounding the flow area, when the flow is modelled using a system of hydromechanical singularities, vortices for instance. The method consists of substituting the Neumann boundary condition \( \vec{V}_n = 0 \) on immobile walls by the images of systems of singularities having proper strengths and situated at proper points outside the analysed flow area, to get the resultant normal velocity component induced by the system of singularities and their images equal to zero. The method is illustrated in Fig. 7, for a single vortex in the flow bounded by a single wall.

The correction taking into account the effect of the presence of walls bounding the measuring space of the tunnel is included in the code by adding the velocities induced by the images of the vortex systems modelling the flow through the turbine. This is done when the coefficient of the equation system (8) and the tangential velocities are calculated, and the new shape of the trailing wake is determined.

In order to take into account the effect of walls bounding the measuring space of the tunnel, the wall impermeability condition (7) was extended by adding the velocities induced by the images, to get the following form:

\[
\vec{V}_{ni} = \vec{V}_\infty \cdot \vec{n}_i + \vec{V}_f \cdot \vec{n}_i + \vec{V}_i^{\|} \cdot \vec{n}_i = 0
\]  \hspace{1cm} (20)

where \( \vec{V}_i^{\|} \) stands for velocities induced by images at the \( i \)-th control point. After applying the above equation to all control points, one obtains the equation system in the matrix form, which can be written as:

\[
([K_n] + [K_n^o]) [\Gamma] = [W]
\]  \hspace{1cm} (21)

where \([K_n^o]\) is the impact factor matrix from images.
The earlier code version assumed constant velocity of the fluid flowing into the turbine, which enabled to reduce significantly the number of unknown variables due to periodicity of turbine geometry. Although justified for flows in unbounded areas, this assumption is too simplistic for the turbine flow analysed here. Indeed, the examined turbine model was equipped with a 5-blade rotor and a duct being a body of revolution, but the test chamber had a square cross section, which made the entire system aperiodic. In order to preserve the earlier reduced equation system and, simultaneously, keep the accuracy of calculations at an acceptable level, a decision was made to average in radial direction the velocity field generated by the created images. The quantity which was averaged for this purpose was the projection of the velocity vector onto the normal direction to the surface of the washed element at the control points marked by the same sequential number in each system geometry segment. The impact factors from images were calculated using the following formula:

\[ K_{t_{p_{ij}}} = \frac{1}{Z} \sum_{k=0}^{Z} \sum_{\sigma}^{O} \sum_{n}^{O} \left( \frac{1}{4\pi} \int_{\sigma}^{O} \frac{r_{p_{ij}} \times n_{i}^{O}}{(v_{p_{ij}} \times n_{i}^{O})^2} \right) \cdot n_{i}^{O} \] (22)

where \( k \) is the sequential number of the segment in which the control point is situated, \( p \) is the sequential number of the image, \( n \) is the sequential number of the segment in which the vortex is situated, \( L \_O \) – is the number of images, and:

- \( r_{p_{ij}} \) is the vector connecting the control point \( i \) situated in segment \( k \) with the vortex \( j \) situated in segment \( n \) of image \( p \).
- \( n_{i}^{O} \) is the vector normal at the control point \( i \) of segment \( k \).

The tangential velocity calculation also includes velocities induced by the images, calculated from the formula:

\[ V_{s_{i}} = \sum_{i=0}^{W_{s}} (K_{t_{p_{ij}}} + K_{t_{p_{ij}}}) i_{j} + V_{\infty} \cdot \Delta_{i} \pm \frac{r_{i} \times r_{i+1}}{4\Delta X_{i}} \] (23)

in which \( L \_W \_S \) stands for the number of vortices in a single segment, \( i_{j} \) is the circulation of the vortex situated at control point \( i \), \( \Delta X_{i} \) is the distance between bound vortices, \( r_{i} \) is the tangent vector to the surface of the washed turbine element at control point \( i \) of segment \( k \). The values of the coefficient matrix \( K \) are calculated from the formula:

\[ K_{t_{p_{ij}}} = \frac{1}{Z} \sum_{k=0}^{Z} \sum_{i=0}^{O} \int_{\sigma}^{O} \frac{r_{p_{ij}} \times n_{i}^{O}}{(v_{p_{ij}} \times n_{i}^{O})^2} \right) \cdot n_{i}^{O} \]

Note: the sign in front of the last term in formula (23) is assumed in the following way: “-” for the suction side of the rotor blade surface, and the hub and duct surfaces, “+” for the pressure side of the rotor blade surface.

The velocities induced by the created images were also taken into account when calculating the vortex wake. In that case, the local velocity at mesh nodes in the area occupied by the wake was calculated using the following formula:

\[ \vec{V} = V_{\infty} + \vec{V}_{\Gamma} + \vec{V}_{\Gamma}^{D} \] (24)

As mentioned above, flow aperiodicity in the analysed case is the reason why the shape of the trailing wake shed from each turbine segment is not the same. In order to preserve the reduced equation system, the velocities induced by the images were calculated at wake node points in each geometry segment. Then, they were rotated by a given angle in such a way that the calculated vectors were situated at the mesh node of the main segment. Finally, the components of the calculated velocity vectors were averaged.

### COMPARISON OF EXPERIMENTAL AND NUMERICAL RESULTS

The results of experimental tests and numerical calculations are compared in Fig. 9 for case I, when the rotor is situated in the inlet plane of the duct, and in Fig. 10 for case II, when
it is shifted downstream of the inlet by 5 cm. The results are presented in dimensionless form, as water energy utilisation factors $C_E$ and thrust coefficients $C_T$, calculated for selected TSR values. Based on the experimentally recorded points, for which total measuring uncertainties were also marked, the approximating curves were constructed using the method of least squares. The distance between the numerically calculated points and this curve is considered the measure of the difference between calculations and experiment.

Based on the obtained results, a general conclusion can be formulated that for both examined turbine model cases good compliance between the measured and calculated values of the dimensionless energy utilisation factor, $C_E$, has been obtained. The largest discrepancies can be observed for extreme TSR values. In particular: for case I relatively large differences are observed at characteristics points representing maximal TSR values. For case II, on the other hand, the calculated values of both coefficients are slightly overestimated for high TSR values, as compared to the experiment based approximating curve, while the $C_E$ values are slightly underestimated for small TSR values.

At the present stage of research, only probable causes of the observed discrepancies can be indicated. A very likely cause, especially in the areas of lower $C_E$ values or those going beyond the optimal operating points of the examined models, is significant simplification of the developed calculation method, mainly resulting from neglecting fluid viscosity and flow separation phenomena in the model. What is also noteworthy is the vibration of the turbine model rotor observed during the experiment. This vibration was much higher than that the one observed when the rotor was removed to measure the dependence of the rotor shaft resistance on rotational speed of the turbine without rotor. Therefore, the real resistance torque of the vibrating shaft with rotor could be slightly higher than the values assumed in the calculations and, consequently, the resultant difference could be larger for higher TSR.

For the dimensionless thrust coefficient $C_T$ in case I, large discrepancies between the calculated and experimental values were observed for high TSR. In case II, relatively good compliance of $C_T$ values obtained from calculations and experiment was observed. A probable cause for the observed differences is disregarding separation and cavitation phenomena in calculations.

**BASIC CONCLUSIONS**

Comparing the results obtained from the experiment with those calculated numerically, one can observe:

- good compliance of values of the dimensionless water energy utilisation factor $C_E$ in the vicinity of the optimal point of turbine operation,
- relatively large differences of the dimensionless thrust coefficient for high TSR values in case II.

Good compliance between the calculated and experimentally recorded values of the dimensionless energy utilisation factor $C_E$ and thrust coefficient $C_T$ within the range of optimal operating conditions of the hydrokinetic turbine model suggests that the developed software can be effectively used in designing turbines of this type.

The presented results of computer simulations of the flow through the hydrokinetic turbine neglect such phenomena as flow separation and cavitation. Further steps of software development will include preparing relevant procedures which will enable to take into account these phenomena, which were likely to reduce the presently observed differences between experimentally recorded and calculated values of coefficients $C_E$ and $C_T$.

**NOMENCLATURE**

- $c$ – profile chord length
- $C_E$ – dimensionless water energy utilisation factor
- $C_T$ – viscous drag coefficient
- $C_p$ – dimensionless pressure coefficient
\( C_t \) – dimensionless thrust coefficient  
\( D \) – diameter of the rotor  
\( f \) – maximal profile camber  
\( K_n \) – vortex effect coefficient at the control point  
\( K_{n0} \) – vortex effect coefficient from images at the control point  
\( l \) – vortex line  
\( M \) – torque generated by the rotor  
\( n \) – rotational speed of the rotor  
\( \mathbf{p} \) – normal vector  
\( \mathbf{p} \) – blade pitch  
\( T \) – thrust acting on the rotor  
\( \mathbf{V} \) – tangent vector  
\( \text{TSR} \) – tip speed ratio  
\( V_n \) – liquid velocity vector projection onto the direction normal to the surface  
\( V_t \) – liquid velocity vector projection onto the direction tangential to the surface  
\( V_i \) – liquid velocity induced by vortices  
\( V_{i'} \) – liquid velocity induced by images  
\( V_{vo} \) – relative velocity of the liquid  
\( W \) – free term  
\( Z \) – number of blades  
\( \Gamma \) – horseshoe vortex circulation  
\( \rho \) – liquid density  
\( \omega \) – angular speed of the rotor

**INDEXES**

\( i, i' \) – numbers of a control point  
\( j, j' \) – numbers of a horseshoe vortex  
\( k, n \) – numbers of a segment  
\( p \) – number of an image

**BIBLIOGRAPHY**


CONTACT WITH THE AUTHORS

Adam Góralczyk
e-mail: adam.goralczyk@gmail.com

Adam Adamkowski
e-mail: aadam@imp.gda.pl

Institute of Fluid-Flow Machinery Polish Academy of Sciences
Fiszera 14
80-231 Gdańsk
POLAND
STUDIES OF RESISTANCES OF NATURAL LIQUID FLOW IN HELICAL AND CURVED PIPES

Michał Stosiak¹
Maciej Zawiślak¹
Bohdan Nishta²

¹ Wrocław University of Science and Technology
² Sumy State University, Ukraine

ABSTRACT

The main aim of this research was to determine in three ways, i.e. experimentally, analytically and by means of numerical modelling, the resistances of the flow of a natural liquid in a helical pipe and in curved pipes. The analyses were carried out for three pipes: one helical pipe and two curved pipes. Each of the pipes was 2 m long and its inside diameter was 4 mm. The experiment was carried out on a test stand making it possible to measure the rate of the flow of the liquid, the temperature at the pipe’s inlet and outlet and the pressure at the pipe’s inlet and outlet. The resistances of the flow of the liquid were calculated from analytical or empirical formulas found in the literature on the subject. Moreover, numerical modelling was performed using the finite volume element method.

Keywords: flow, resistance of flow, helical pipe, curved pipe, natural liquid, laminar flow

INTRODUCTION

There is practically no industry which to a greater or lesser extent does not need to move liquids or gases through pipes, ducts and other installations. The hydraulic or gas networks can vary widely in their complexity.

In some cases, these are backbone networks consisting of mainly very long straight conduits (oil and gas pipelines, water and steam pipes and ducts in air ventilation systems in large businesses).

In other cases, these are relatively short networks with a large number of differently shaped and branched parts and various obstructions in the form of throttling or control components. Hydraulic pumps, as well as motors, consist of consecutive, local obstacles of complex shapes in close proximity to each other [1]. In such elements besides volumetric losses [2], there are losses due to the flow of a medium through internal channels.

Hydraulic drive is also widely used in ships and other vessels. In marine applications is especially important for the whole system to be watertight due to the corrosive properties of seawater that can easily damage the elements. Furthermore, it is crucial from the standpoint of reliability and efficiency, to have reliable tools that can estimate the amount of flow losses in the system. During the design phase, it is convenient to use CFD modelling, however, acquired results should be validated experimentally.

In all the cases, the hydraulic (aerodynamic) resistances of the flow of a liquid or a gas need to be calculated. Furthermore, the modern design of structures with complicated forms of hydraulic and gas networks requires knowledge of the basic physical and mechanical processes which take place in complex network components and should take into account the recommendations for improving the flow conditions and reducing the hydraulic resistance of such components [3].
Losses of pressure always occur during the transport of liquids or gases through conduits [4]. The size of the losses affects the demand for power needed to transport the liquids or the gases and thereby has a bearing on the operating costs of the entire transmission system. Understanding of flow losses is also important in pumps, motors and hydraulic valves in which, due to the complex shapes of apertures and channels [5], known and simple relationships cannot be applied with satisfactory accuracy.

Therefore calculations of flow losses in hydraulic systems are of vital importance. The number of pipe configurations is huge and so not all of them can be presented here. Therefore three different pipes: a helical pipe and two curved pipes were selected for consideration in this paper.

EXPERIMENTAL DETERMINATION OF LIQUID FLOW RESISTANCES

DESCRIPTION OF STUDIED PIPES

As part of this research the resistances of the flow of a liquid in three rigid pipes which had been formed from steel pipes manufactured by Parker Hannifin were studied. Each of the pipes is 2 m long and its inside diameter amounts to 4 mm:

№1 – a helical pipe (fig. 1.1.(a)),
№2 – a curved pipe with two long straight sections and a bend having a radius of 32.5 mm (fig. 1.1.(b)),
№3 – a curved pipe with 9 bends each having a radius of 40 mm (fig. 1.1.(c)).

Measurements of the roughness of the pipe’s interior side (fig. 1.2) were carried out using a HOMMEL TESTER T1000. The results of the measurements are presented in table 1.1.

![Fig. 1.2. Interior side of pipe](image)

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Measurement №1</th>
<th>Measurement №2</th>
<th>Measurement №3</th>
<th>Average</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ra, μm</td>
<td>1.29</td>
<td>1.22</td>
<td>1.26</td>
<td>1.257</td>
</tr>
<tr>
<td>Rz, μm</td>
<td>7.78</td>
<td>7.54</td>
<td>8.12</td>
<td>7.813</td>
</tr>
<tr>
<td>Rt, μm</td>
<td>9.34</td>
<td>8.36</td>
<td>9.48</td>
<td>9.06</td>
</tr>
</tbody>
</table>

PROPERTIES OF WORKING LIQUID

Hydraulic oil FUCHS RENOLIN VG 46 was used as the working liquid. RENOLIN VG 46 is intended for hydraulic devices which can operate at high temperatures in equipment where the pumps require hydraulic oils reducing abrasion in mixed friction conditions. This oil is highly resistant to ageing and is characterized by high oxidation stability. Its foaming tendency is low and its air releasing power is high. The specifications of the oil are given in table 1.2 [6].

![Fig. 1.1. Studied pipes](image)

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Numerical value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Density at 15°C, kg/m³</td>
<td>877</td>
</tr>
<tr>
<td>Kinematic viscosity at 40°C, mm²/s</td>
<td>46</td>
</tr>
<tr>
<td>Kinematic viscosity at 100°C, mm²/s</td>
<td>6.8</td>
</tr>
<tr>
<td>Viscosity index</td>
<td>95</td>
</tr>
<tr>
<td>Fluidity loss temperature, °C</td>
<td>-21</td>
</tr>
</tbody>
</table>

DESCRIPTION OF LABORATORY TEST STAND

The test stand shown in fig. 1.3 was used to measure the resistances of the flow of the hydraulic oil. The hydraulic oil is delivered to the studied pipe’s inlet, where temperature T1 and pressure p1 are measured, by a pump via a pipe with a relief valve. Temperature was measured by a HYDROTECHNIK HySense TE100 ISDS temperature sensor. The pressure at the studied pipe’s inlet was measured by a STAUFF SPD-DIGI-B0100-B pressure gauge with a measuring range of 0-100 bar. The pressure at the pipe’s outlet was measured by a STAUFF SPD-DIGI-B0016-B pressure gauge with a measuring range of 1–16 bar. After leaving the studied pipe the oil was delivered to a HYDROTECHNIK HySense QG100 flow meter with a measuring range of 0.7-70 l/min. Then the oil returned to the tank. The data from the temperature sensors and the flow
meter were read off a HYDROTECHNIK Multi System 8050 multichannel measurement and control instrument.

Fig. 1.3. Measuring stand: 1 – pipe in which flow resistance is investigated, 2 – digital pressure gauge at pipe’s inlet, 3 – digital pressure gauge at pipe’s outlet, 4 – thermometer at pipe’s outlet, 5 – thermometer at pipe’s outlet, 6 – flow meter, 7 – relief valve, 8 – tank, 9 – multichannel measurement and control instrument

DETERMINATION OF MEASUREMENT UNCERTAINTY

In order to determine the accuracy (by determining the maximum systematic uncertainty) of the pressure, flow rate and liquid temperature recording one should add up the individual uncertainties occurring in the whole measuring circuit. The maximum measurement uncertainty for the circuit measuring pressure \( p \), flow rate \( Q \) and temperature is expressed by the relation [7]:

\[
\delta_{\text{max}} = |\delta_{\text{sensor}}| + |\delta_{\text{amplifier}}| + |\delta_{\text{computer}}| \quad (1.1)
\]

The following were used in the experiments:
- a HYDROTECHNIK HySense TE100 ISDS temperature sensor with \( \delta_{\text{sensor}} \leq 0.5\% \) of the end value;
- a STAUFF SPD-DIGI-B0100-B pressure gauge with \( \delta_{p, \text{sensor}} \leq 0.5\% \);
- a STAUFF SPD-DIGI-B0016-B pressure gauge with \( \delta_{p, \text{sensor}} \leq 0.5\% \); a HYDROTECHNIK HySense QG100 flow meter with \( \delta_{Q, \text{sensor}} \leq 0.4\% \);
- the frequency input of the HYDROTECHNIK Multi System 8050 multichannel measurement and control instrument for flow rate recording, with \( \delta_{Q, \text{computer}} \leq 0.02\% \);
- the analogue inputs of the HYDROTECHNIK Multi System 8050 multichannel measurement and control instrument for temperature recording, with \( \delta_{T, \text{computer}} \leq 0.1\% \).

Taking into account the individual measurement uncertainties, the pressure recording uncertainty was evaluated to amount to \( \delta_{p, \text{max}} \leq 0.5\% \), the flow rate recording uncertainty to \( \delta_{Q, \text{max}} \leq 0.42\% \) and the temperature recording uncertainty to \( \delta_{T, \text{max}} \leq 1.1\% \).

Values of the exponents \( a \) after data approximation with an exponential function shows, that for pipe 1 flow is laminar for the whole range of \( Q \), and for pipes 2 and 3 the type of the flow changes to transitional state with a dominance of turbulent flow.

ANALYTICAL CALCULATION OF LIQUID FLOW RESISTANCES

INPUT DATA

The input data for the calculations are given in table 2.1.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>( \text{Pipe number} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \text{Pipe number} )</td>
<td>( #1 )</td>
</tr>
<tr>
<td>Inside diameter, mm</td>
<td>4</td>
</tr>
<tr>
<td>Initial oil temperature, ( \circ\text{C} )</td>
<td>39.4</td>
</tr>
<tr>
<td>Coefficient of kinematic viscosity, m/s</td>
<td>( 46 \times 10^{-4} )</td>
</tr>
<tr>
<td>Density, kg/m³</td>
<td>877</td>
</tr>
<tr>
<td>Flow rate, m³/s</td>
<td>( 1.354 \times 10^{-4} )</td>
</tr>
</tbody>
</table>

The Reynolds number for the studied pipes can be calculated from the formula [8]:

\[
Re = \frac{cd_h}{v} = \frac{4Q}{\pi d_h v} \quad (2.1)
\]
where: \( c \) – the average flow rate, \( \nu \) – the coefficient of the kinematic viscosity of the liquid, \( d_h \) – the hydraulic diameter of the pipe, \( Q \) – the rate of flow.

Substituting the appropriate values one gets \( Re_1 = 936.78 \), \( Re_2 = 1270 \), \( Re_3 = 1125 \). The calculations show that a laminar flow will occur in all the three pipes.

**CALCULATION OF LIQUID FLOW RESISTANCES FOR STUDIES PIPES**

In order to calculate pressure losses along the straight sections of the pipes for the laminar flow of the oil one can use the relation [8]:

\[
\Delta p = \frac{64}{Re} \frac{l}{d} \frac{c^2}{2} \tag{2.2}
\]

Since the studied pipes vary in the length of their straight section the overall length of the straight section of each of the pipes is calculated below and the pressure losses for the sections are determined. The results are presented in table 2.2.

**Table 2.2. Length of straight section for each pipe and respective pressure losses.**

<table>
<thead>
<tr>
<th>Calculated quantity</th>
<th>Pipe number</th>
<th>№1</th>
<th>№2</th>
<th>№3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overall length of straight section, m</td>
<td>0.1</td>
<td>1.898</td>
<td>1.15</td>
<td></td>
</tr>
<tr>
<td>Pressure loss, Pa</td>
<td>0.87×10^5</td>
<td>22.37×10^5</td>
<td>12.01×10^5</td>
<td></td>
</tr>
</tbody>
</table>

The pressure losses in the helical part of pipe №1 were calculated on the basis of [9, 10, 12].

Dean number \( De \) and critical Dean \( De_{cr} \) can be calculated from the formulas [9]:

\[
De = Re \sqrt{\frac{d}{R}} \tag{2.3}
\]

\[
De_{cr} = 2 \cdot 10^4 (0.5d/R)^{0.82} \tag{2.4}
\]

\[
R^* = R \left[ 1 + \frac{t_p}{(2\pi R)^2} \right] \tag{2.5}
\]

where: \( t_p \) – the pitch of the helix, \( R \) – the radius of the helix (fig. 2.1).

For the given helical part of pipe №1: \( t_p = 6 \) mm, \( R = 23.5 \) mm, \( De = 386 \) and \( De_{cr} = 2649 \).

If the Dean number is lower than the critical one can use the following formula for the coefficient of hydraulic friction, taking into consideration the circular cross section of the pipe and the laminar flow of the liquid [9]:

\[
\lambda = 0.1008(Re)^{1/2}[1 + 3.945(Re)^{1/2} + 7.082(Re)^{-1/2} + 9.097(Re)^{-1/2}]^{1/2} + 5.608(Re)^{-1/2} \tag{2.6}
\]

For the given helical part: \( \delta = 3870^\circ \), \( R = 0.0235 \) m, \( d = 0.004 \) m, \( c = 10.773 \) m/s and the coefficient of hydraulic friction calculated from formula (2.6) amounts to \( \lambda = 0.144 \). Substituting the data into formula (2.7) one gets \( \Delta p = 29.26 \times 10^5 \) Pa.

Since for all the curved sections of the pipes \( R/D > 3 \) and \( 50 < Re(2R/D)^{-0.5} < 600 \) the following formula was used to calculate the pressure drop in the curved sections of the pipe [9]:

\[
\Delta p = 0.0175 \frac{20}{Re^{0.65}} \left( \frac{D_0}{2R_0} \right)^{0.175} \frac{c^2}{\rho} \tag{2.8}
\]

The pressure drop in the helical part amounts to:

\[
\Delta p = \frac{0.0175 \times 20}{Re^{0.65}} \left( \frac{D_0}{2R_0} \right)^{0.175} \frac{c^2}{\rho} \tag{2.7}
\]

For the given helical part: \( \delta = 3870^\circ \), \( R = 0.0235 \) m, \( d = 0.004 \) m, \( c = 10.773 \) m/s and the coefficient of hydraulic friction calculated from formula (2.6) amounts to \( \lambda = 0.144 \). Substituting the data into formula (2.7) one gets \( \Delta p = 29.26 \times 10^5 \) Pa.

Since for all the curved sections of the pipes \( R/D > 3 \) and \( 50 < Re(2R/D)^{-0.5} < 600 \) the following formula was used to calculate the pressure drop in the curved sections of the pipe [9]:

\[
\Delta p = 0.0175 \frac{20}{Re^{0.65}} \left( \frac{D_0}{2R_0} \right)^{0.175} \frac{c^2}{\rho} \tag{2.8}
\]

Figure 2.3 shows the curved sections of pipe №3.
The results of the calculations of pressure losses in the curved sections of the studied pipes are presented in table 2.3.

Table 2.3. Pressure losses in curved sections

<table>
<thead>
<tr>
<th>Calculated quantity</th>
<th>Pipe number</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>№1</td>
</tr>
<tr>
<td>Pressure loss, Pa</td>
<td>1.48·10⁵</td>
</tr>
</tbody>
</table>

Since the studied pipes were connected to the pressure gauges via a 66 m long adapter with an inside diameter of 22 mm the pressure losses in the straight section of the adapters were calculated from formula (2.2) while the pressure losses due to an abrupt extension and contraction were calculated from the formula [8]:

\[ \Delta p = \frac{\zeta_M c^2}{2} \]  (2.9)

The coefficient of local losses due to an abrupt extension is equal to [9]:

\[ \zeta_M = \frac{8.44556 - 26.163(1 - F_0/F_1)^4 - 5.38086(1 - F_0/F_1)^4 + 4:\text{lg}Re[6.007 + 18.5372(1 - F_0/F_1)^4 + 3.9978(1 - F_0/F_1)^4] + (\text{lg}Re)^2[-1.02318 - 3.0916(1 - F_0/F_1)^2 - 0.680943(1 - F_0/F_1)^4]}{2} \]  (2.10)

where: \( F_0 \) – the cross-sectional area of the thinner pipe, \( F_1 \) – the cross-sectional area of the thicker pipe.

The coefficient of local losses due to an abrupt contraction is equal to [9]:

\[ \zeta_M = A \cdot B(1 - F_0/F_1) \]  (2.11)

where:

\[ A = \sum_{i=0}^{7} a_i(\text{lg}Re)^i; \quad a_0 = -25.12458; \quad a_1 = 118.5076; \]
\[ a_2 = -170.4147; \quad a_3 = 118.1949; \quad a_4 = -44.42141; \]
\[ a_5 = 9.09524; \quad a_6 = -0.9244027; \quad a_7 = 0.03408265; \]

\[ B = \sum_{i=0}^{2} \left( \sum_{j=0}^{2} a_{ij}\left(\frac{F_0}{F_1}\right)^i \right) (\text{lg}Re)^j \]  (2.12)

\[ a_{ij} = \begin{bmatrix} 1.07 & 1.22 & 2.9333 \\ 0.05 & -0.51668 & 0.8333 \end{bmatrix} \]  (2.13)

The calculated pressure losses in the straight section of the adapters and the pressure losses due to an abrupt extension and an abrupt contraction, and the total losses in the adapters are presented in table 2.4.

Table 2.4. Pressure losses in adapters

<table>
<thead>
<tr>
<th>Calculated quantity</th>
<th>Pipe number</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>№1</td>
</tr>
<tr>
<td>Pressure loss in straight section of adapters, Pa</td>
<td>125</td>
</tr>
<tr>
<td>Pressure loss due to abrupt extension, Pa</td>
<td>1.18·10⁵</td>
</tr>
</tbody>
</table>

The total pressure losses in pipes and the difference in comparison with the pressure losses determined experimentally are presented in table 2.5.

Tab. 2.5. Analytical and experimental pressure losses in pipes

<table>
<thead>
<tr>
<th>Calculated quantity</th>
<th>Pipe number</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>№1</td>
</tr>
<tr>
<td>Analytical total pressure losses in pipe Ðp2, Pa</td>
<td>33.15·10⁵</td>
</tr>
<tr>
<td>Experimental pressure losses Ðp1, Pa</td>
<td>40.98·10⁵</td>
</tr>
</tbody>
</table>

According to table 2.5, the theoretical pressure losses are lower than the experimental ones and the calculation error does not exceed 20%. Considering that the calculations were done for the isothermal flow of the oil and that in the literature on the subject [9] one can find an empirical correction factor of 75/64=1.17, the obtained results are satisfactory.

**CALCULATION OF LIQUID FLOW RESISTANCES BY MEANS OF NUMERICAL MODELLING**

**PROBLEM SETUP**

In this section the problem of the flow of the hydraulic oil in the studied pipes is solved by means of numerical modelling in ANSYS CFX. The solution was reached in several steps: the creation of the investigated object’s geometry of, the generation of a finite volume element mesh, the setting of proper liquid parameters and boundary conditions, computations, an analysis of the pressure distribution in the pipes and of the pressure losses.

The parameters of the liquid needed for the calculations are: density \( \rho = 877 \text{ kg/m}^3 \) and coefficient of dynamic viscosity \( \mu = 0.04 \text{ kg m}^{-1} \text{s}^{-1} \).

The pressure at the pipe’s inlet and the mass rate of flow at the pipe’s outlet, whose values are given in table 3.1, were set as the boundary conditions.

Tab. 3.1. Boundary conditions for numerical computations

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Pipe number</th>
</tr>
</thead>
<tbody>
<tr>
<td>Inlet pressure, Pa</td>
<td>№1</td>
</tr>
<tr>
<td></td>
<td>50·10⁵</td>
</tr>
<tr>
<td>Outlet mass rate of flow, kg/s</td>
<td>0.119</td>
</tr>
</tbody>
</table>

**COURSE OF NUMERICAL MODELLING**

First geometric models of the liquid for the studied pipes were created on the basis of the models of the pipes used in the experiments (fig. 3.1). Subsequently a mesh of volume elements was generated. Prismatic elements with a characteristic size of 0.8 mm
were used for this purpose. There were 5 boundary layers, the layer width amounted to 0.5 mm and the mesh growth rate to 1.2. A close-up of the volume element mesh is shown in fig. 3.2.

Then the boundary conditions were set according to table 3.1 (fig. 3.3). For all the models the inlet pressure, the outlet mass rate of flow and “hydraulically smooth” walls on all the surfaces, except for the pipe’s inlet and outlet, were set.

The laminar flow of the liquid was assumed in the computations.

**Fig. 3.1. Geometric models of liquid for studied pipes**

**Fig. 3.2. Close-up of volume element mesh for pipe №1**

**Fig. 3.3. Boundary conditions for computing oil flow in pip №1**

**Fig. 3.4. Distribution of pressure in pipes**

**Tab. 3.2. Total pressure losses in pipes, computed by means of numerical modelling**

<table>
<thead>
<tr>
<th>Parameter</th>
<th>№1</th>
<th>№2</th>
<th>№3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Inlet pressure, Pa</td>
<td>$50 \times 10^5$</td>
<td>$46 \times 10^5$</td>
<td>$54 \times 10^5$</td>
</tr>
<tr>
<td>Outlet pressure, Pa</td>
<td>$17.06 \times 10^5$</td>
<td>$15.39 \times 10^5$</td>
<td>$17.72 \times 10^5$</td>
</tr>
<tr>
<td>Pressure losses in pipe, Pa</td>
<td>$32.94 \times 10^5$</td>
<td>$30.61 \times 10^5$</td>
<td>$36.28 \times 10^5$</td>
</tr>
</tbody>
</table>
The total pressure losses in the pipes and the computation error in comparison with the experimentally determined pressure losses are presented in table 3.3.

<table>
<thead>
<tr>
<th>Computed quantity</th>
<th>Pipe number</th>
<th>№1</th>
<th>№2</th>
<th>№3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total power losses in pipe computed by means of numerical modelling</td>
<td>Δp3, Pa</td>
<td>32.94·10^5</td>
<td>30.61·10^5</td>
<td>36.28·10^5</td>
</tr>
<tr>
<td>Experimental power losses in pipe</td>
<td>Δp1, Pa</td>
<td>40.98·10^5</td>
<td>31.93·10^5</td>
<td>42.17·10^5</td>
</tr>
<tr>
<td>Error of computations by means of numerical modelling</td>
<td>[(Δp1- Δp3)/Δp1]·100, %</td>
<td>19.6</td>
<td>4.1</td>
<td>14</td>
</tr>
</tbody>
</table>

The results of the numerical modelling computations show that the pressure losses determined in this way are lower than the ones determined experimentally and the computation error does not exceed 20%, while for pipe №2 with a long straight section the error amounts to 4.1%, which can indicate good agreement between the experimental data and the numerical computations.

**CONCLUSION**

The experiment has shown that the highest pressure losses occur in the curved pipe. This can be due to the fact that most of the pressure drop in curved pipes is caused by the vortices arising on the interior wall as a result of the centrifugal force acting on the liquid.

The analytical calculations indicate that the calculation error in comparison with the experimental data amounts to 19%. However, it should be noted that the calculations were carried out for the isothermal flow of the liquid, which is a simplifying assumption. In the literature on the subject [8, 9, 11] one can find an empirical relation taking into account the non-isothermal character of the flow, in which the coefficient of linear losses is assumed to be equal to 75/Re instead of 64/Re. Using the coefficient of linear losses which takes the non-isothermal character of the flow into account, the analytically calculated pressure losses will be 17% higher and so the error in comparison with the experimental pressure losses will be less than 5%.

The numerical modelling showed very good agreement with the experimental data only in the case of pipe №2 with a long straight section. Whereas for the other two pipes the error amounts to respectively 19.6 and 14%.

**REFERENCES**

CONTACT WITH THE AUTHORS

Michał Stosiak  
e-mail: michal.stosiak@pwr.edu.pl
Wrocław University of Science and Technology  
Wybrzeże Stanisława Wyspiańskiego 27  
50-370 Wrocław  
**POLAND**

Maciej Zawiślak  
e-mail: maciej.zawislak@pwr.edu.pl
Wrocław University of Science and Technology  
Wybrzeże Stanisława Wyspiańskiego 27  
50-370 Wrocław  
**POLAND**

Bohdan Nishta  
e-mail: bohdan.nishta@gmail.com
Sumy State University  
Rymskogo-Korsakova 2  
40-007 Sumy  
**UKRAINE**
COLD CRACKING OF S460N STEEL WELDED IN WATER ENVIRONMENT

Jacek Tomków
Jerzy Łabanowski
Dariusz Fydrych
Grzegorz Rogalski
Gdansk University of Technology, Poland

ABSTRACT

This paper shows results of weldability testing of fine-grained high-strength low-alloy S460N steel welded in water environment by covered electrodes. The tests were carried out by using the CTS test specimens with fillet welds. Four specimens were welded under water and one specimen in air. Welded joints were subjected to non-destructive visual and penetration tests. The accepted joints were then subjected to macroscopic and microscopic inspection and Vickers hardness measurements as well. The experiments showed that S460N steel welded in water environment is characterized by a high susceptibility to cold cracking.

Keywords: underwater welding, wet welding, high-strength low-alloy steel, cold cracking, CTS test

INTRODUCTION

Development of underwater welding is a result of increasing demand on mineral raw materials. Searching for hydrocarbons resources on sea and ocean bed caused in extensive progress of diving techniques as well as making use of divers in building marine structures. From several dozen years an issue of conducting repair of ocean-engineering and hydro-technical structures has been appeared more and more often. The developing offshore industry demands to make the repairs in water environment, without disassembling the elements and moving them into air [1,2]. The most widely used repair method is wet welding which, for economic reasons, is carried out generally with covered electrodes [3,4]. However in wet welding FCAW method is also common used [5]. Despite its common application the method leads to many problems [6,7]. For structural elements working under water high strength steels are more and more widely used. Weldability of these steels in the air is a well investigated issue but transferring the process into water environment still produces many problems [3,4]. The joints made by using wet method show presence of welding imperfections such as gas porosity, shape failures and cracks [8,9]. The issue most often met is steel susceptibility to cold cracking [10-13]. Water environment significantly intensifies risk of producing such cracks due to the following factors appearing during welding process [1,14]:

• increased cooling rate resulting from fast transfer of heat from outer part of arc and welding area,
• hydrogen concentration resulting from presence of water vapour surrounding welding arc,
• presence of residual stresses,
• increased pressure values, leading to instability of welding arc.

The water environment makes it impossible or difficult to apply many traditional methods for preventing cold cracking, which are successfully used in making welded joints in the air. Out of the methods lowering risk of cold cracking in water environment, the following find application in practice [3,7,23]:

• the increasing of welding heat input,
• the application (for crucial elements only) of austenitic electrodes whose deposited metal is more ductile,
• the addiction of calcium fluoride to electrode coating which binds hydrogen into a durable compound in welding temperature, decreasing its influence on generating cold cracks.

There is a common opinion that water practically eliminates possible application of preheating to welded joints as well as their further thermal treatment to be performed in a traditional way, that is able to minimize steel susceptibility to cracking. Fast cooling rate unfavourably influences structural transformations, that impairs joint properties [15,16]. Water environment makes the cooling rate faster. However, the latest investigations show that application of heating induction is possible [17]. Research tests on ultrasonic enhancement the underwater welding processes [18,22] or application of temper bead welding technique [19,24] are also carried out. The minimizing of steel susceptibility to cold cracking is especially important in case of higher strength steels. Water contributes to increasing number of imperfections, resulting also in decreasing mechanical properties of welded joints. It results a. o. from diffusion of hydrogen which creates favourable conditions for producing cracks and micro-cracks lowering the strength of joints [7,14,20]. The recent investigations show that even a change in chemical content of electrodes does not eliminate the problem of cold cracking which appear in the joints [12].

For assessment of steel weldability in water environment the self-restraint CTS tests and Tekken tests are used [1,2,6]. As underwater welding operations, especially those connected with failure repair, require making fillet joints, the CTS test was selected for the assessment of susceptibility to cold cracking of the S460N steel.

**EXPERIMENTS AND RESULTS**

The research was aimed at the assessment of weldability of S460N fine-grained steel during wet welding by means of MMA method (111 acc. to ISO 4063). The research program covered conducting CTS tests on five joints in accordance with guidelines of EN ISO 17642-2 standard, four out of which were welded in water environment and one in air. The CTS test belongs to technological self-restraint tests and makes it possible to assess susceptibility to cold cracking in all zones of overlapping joints with fillet welds. Fig. 1 presents the schematic diagram of the specimen in question.

The test specimen was produced of S460N fine-grained high strength steel plate of 12 mm in thickness. Tab.1 shows the chemical composition of the steel according to a control analysis performed by using spark emission spectrometry. The mechanical properties of the examined steel are presented in Tab. 2.

---

**Tab. 1. Chemical composition of S460N steel**

<table>
<thead>
<tr>
<th>Element content, %</th>
<th>C</th>
<th>Si</th>
<th>Mn</th>
<th>P</th>
<th>Cr</th>
<th>Mo</th>
<th>Ni</th>
<th>Al</th>
<th>Cu</th>
<th>V</th>
<th>CeMIS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acc. control analysis</td>
<td>0.16</td>
<td>0.53</td>
<td>1.51</td>
<td>0.02</td>
<td>0.07</td>
<td>0.03</td>
<td>0.05</td>
<td>0.033</td>
<td>0.13</td>
<td>0.097</td>
<td>0.464</td>
</tr>
<tr>
<td>CeMIS = C + (Cr+Mo+V)/5 + (Cu+Ni)/15</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Tab. 2. Mechanical properties of S460N steel**

<table>
<thead>
<tr>
<th>Yield point Re, MPa</th>
<th>Ultimate tensile strength Rm, MPa</th>
<th>Elongation A, %</th>
</tr>
</thead>
<tbody>
<tr>
<td>511</td>
<td>626</td>
<td>27.3</td>
</tr>
</tbody>
</table>

The welding operation was carried out in water environment at the water depth of 150 mm by using the underwater test stand (Fig. 2) of the Laboratory of Welding Engineering in Gdańsk University of Technology.
The anchor welds and test welds were made with the use of OMNIA (ISO 2560-A: E 38 O R 11) commercial rutile electrodes of 4 mm diameter, intended for welding in all positions. The electrodes were selected based on the fact that they ensure good plastic properties of deposited material in order to minimize possible generation of cracks. The chemical composition and mechanical properties of the applied electrodes are given in Tab. 3.

<table>
<thead>
<tr>
<th>Chemical composition, wt %</th>
<th>Mechanical properties of deposited material</th>
</tr>
</thead>
<tbody>
<tr>
<td>C</td>
<td>Mn</td>
</tr>
<tr>
<td>0.07</td>
<td>0.55</td>
</tr>
</tbody>
</table>

The CTS specimens were marked by using the test name and subsequent arabic digits. The welding was carried out with the constant DC-polarity in compliance with producer’s recommendations. The welding process parameters are presented in Tab. 4.

<table>
<thead>
<tr>
<th>Weld I</th>
<th>Weld II</th>
</tr>
</thead>
<tbody>
<tr>
<td>No. of specimen</td>
<td>Welding current, I, A</td>
</tr>
<tr>
<td>W1</td>
<td>192</td>
</tr>
<tr>
<td>W2</td>
<td>192</td>
</tr>
<tr>
<td>W3</td>
<td>184</td>
</tr>
<tr>
<td>W4</td>
<td>188</td>
</tr>
<tr>
<td>P5</td>
<td>181</td>
</tr>
</tbody>
</table>

The prepared specimens were subjected to non-destructive tests after 72 h from the end of welding operation. The visual and penetration tests were conducted in compliance with the standards EN ISO 17637 (VT) and EN ISO 3452-1 (PT). In the case of the specimens welded under the water, weld imperfections were observed in one weld of the specimen 2 (insufficient weld thickness, incompletely filled groove, incomplete fusion) as well as in two welds of the specimen 4 (in the first - laid weld: insufficient weld thickness, lack of fusion; in the second-laid weld: undercuts, incomplete fusion, insufficient weld thickness, spattering). The joints were qualified unacceptable. In the case of the specimen welded in air, cracks were observed in crater pipe, which did not excluded the joints from further tests. In the visual tests (VT) the welds complying with the quality level B acc. EN ISO 7637 standard were acceptable. In the penetration tests the welds complying with the quality level 2X acc. EN ISO 3452-1 standard were considered acceptable. Fig. 3 presents example results of the non-destructive tests in question.

MACROSCOPIC METALLOGRAPHIC TESTS

The metallographic tests were carried out in acc.to guidelines of EN ISO 17639 standard – after etching the specimens with 4 % Nital. The macroscopic tests of CTS specimens welded in water environment revealed imperfections such as gas pores and cracks. The welds made in air did not have imperfections. Fig. 4 presents example results of the macroscopic tests.

Fig. 3. Results of non-destructive testing of S460N steel joints; a) VT, the specimen W2, the first-laid weld: incompletely filled groove, incomplete fusion; b) PT, the specimen W1: no imperfections.

Fig. 4. Pictures illustrating results of the macroscopic testing of S460N steel; a) the specimen W3.1: cracks along fusion line and in HAZ; b) the specimen P5.1: no imperfections. Etching: Nital
MICROSCOPIC METALLOGRAPHIC TESTS

As a result of the performed microscopic tests it was stated that the structure of S460N steel is composed of fine-grained pearlite as well as fine-grained ferrite with layers. HAZ has brittle structures characteristic for joints made in fast cooling rate conditions. Cracks was revealed in all the specimens welded in the water. They were placed in the welds, HAZ as well as fusion line along 65-75% of its length. During examining the specimen welded in air no cracks were found. Fig. 5 presents the results of microscopic tests of S460N steel.

HARDNESS MEASUREMENTS

The Vickers hardness measurements under 98 N load (HV10) along the line placed 2 mm below weld face were made in compliance with the requirements of PN-EN ISO 9015-1:2011 standard. Fig. 6 presents the schematic layout diagram of measurement points. Fig. 7 shows example results of hardness measurements for the S460N steel specimens. The maximum values of hardness in HAZ of the tested CTS joints welded in water environment were equal to around 500 HV10, reaching its very maximum of 508 HV10 in the specimen W3. In the case of the specimen P5 welded in air the maximum hardness reached 468 HV10. For S460N steel the limit hardness value of 380 HV10 was assumed according to the PN-EN ISO 15614-1:2008 standard. Tab. 5 shows results of hardness measurements in the form of maximum values found in HAZ of particular joints.

Fig. 5: Results of the microscopic testing of S460N steel joints welded under water; a) base material, magn. 400x; b) cracks in HAZ, magn. 50x; c) cracks in HAZ, magn. 500x; d) cracks in weld, magn. 500x. Etching: Nital

Fig. 6: Schematic layout of hardness testing points – CTS test

Fig. 7: Results of hardness measurements for S460N steel; a) hardness distribution across W3 (first-laid weld) joint welded underwater, HV10max=508; b) hardness distribution across P5 (first-laid weld) joint welded underwater, HV10max=414
Tab. 5. Results of hardness tests

<table>
<thead>
<tr>
<th>No. of specimen</th>
<th>Maximum value of HV10</th>
</tr>
</thead>
<tbody>
<tr>
<td>W1</td>
<td>503</td>
</tr>
<tr>
<td>W2</td>
<td>500</td>
</tr>
<tr>
<td>W3</td>
<td>508</td>
</tr>
<tr>
<td>W4</td>
<td>525</td>
</tr>
<tr>
<td>P5</td>
<td>433</td>
</tr>
</tbody>
</table>

**SUMMARY**

Weldability tests of S460N high strength steel 12 mm plate specimens welded with covered electrodes in water environment using wet method and in air were performed. The welding process involved many problems - welding arc was burning instable that made forming correct welds difficult. In joints made of the tested steel many cracks located both in the weld and HAZ were found. The microscopic tests revealed that the cracks run along 65-75 % fusion line. In the case of joints welded in air environment no imperfections were found. Results of Vickers hardness measurements (HV10) demonstrated that all the specimens welded under water had the hardness value reaching 525 HV10. It contributes to forming cold cracks such as those found during the other tests. None of the tested joints will satisfy the assumed hardness criterion according to EN ISO 15614-1 standard (380 HV10) for the tested steel. One of the ways of lowering the hardness of welded joints made of a high strength steel in water environment may be application of the temper bead welding technique [25]. The other way to improve the quality of joints made in water environment is friction taper welding [27-29].

Haring regard to the value of S460N steel carbon equivalent, welding conditions favouring the increase of steel tendency to cold cracking and experience [3,4,8,19,24] it should be recognized, that presented results are justified

**CONCLUSIONS**

1. Based on the results of the technological weldability tests of S460N high strength steel it may be stated that CTS welded joints made under water by using covered electrodes (rutile) in wet conditions are characterized by a high susceptibility to cold cracking.

2. The underwater wet welding using covered electrodes resulted in a significant increase of hardness of the HAZ of welded joints up to 525 HV10 (in the specimen W4). All of the specimens made under water and made in air do not satisfy the acceptance criterion acc. EN ISO 15614-1 standard.

3. It is recommended to perform weldability tests of the butt-welded joints (Tekken test) as well as to conduct a research aimed at the improvement of S460N steel weldability in water environment.

**REFERENCES**


CONTACT WITH THE AUTHOR
Jacek Tomków
e-mail: jacek.tomkow@pg.edu.pl
Gdańsk University of Technology
11/12 Narutowicza St.
80 - 233 Gdańsk
POLAND
INFLUENCE OF GEOTECHNICAL CONDITIONS ON DAMAGE STATES OF GDANSK BAY COAST CLIFF FORMATIONS

Eligiusz Mieloszyk 1
Mariusz Wyroślak 2
1 Gdansk University of Technology, Faculty of Civil and Environmental Engineering, Department of Rail Transportation and Bridges, Poland
2 Gdansk University of Technology, Faculty of Civil and Environmental Engineering, Department of Geotechnics Geology and Marine Civil Engineering, Poland

ABSTRACT

A geotechnical aspect of destruction processes of seashores was identified based on the case of erosion of Gdansk Bay Coast cliff formations. Causative factors of landslide were described in the context of natural phenomena, land development and its anthropogenic transformations. Possible directions of theoretical analysis of changes in ground-water relations were indicated.

Keywords: coastal cliff, escarpment stability, underground water flow

INTRODUCTION

The Erosion of coastal cliffs is a natural process. Safe existence of cliff areas is not always associated with huge outlays on protecting structures. A right recognition of destructive factors may help in preparation of an effective shore protection project [9], but an erroneous recognition will not hold back cliff erosion processes [5].

This paper is aimed at indication of influence of geotechnical factors on forming ground sculpture in the region of South Baltic Sea Coast, taking into consideration the case of a landslide occurred at Kamienna Góra (“Rock Hill”), Gdynia. Unquestionably, sea wave impact has the greatest effect on sliding the cliff, if only it is located close to sea. If it is separated from sea by a wide beach the sea wave influence diminishes and the influence of water filtering through the ground from land side becomes the crucial destructive factor.

GEOLOGICAL CONFIGURATION

In the Gdansk Bay Coast region there are geological formations made as a result of accumulative action of continental glacier, accumulative and erosion action of thaw waters, river waters, denudation and sea deposits, plant-built forms as well as anthropogenic ones [1].

The area of Kamienna Góra is a moraine plateau (hill) of a very differentiated sculpture formed as a result of a non-uniform accumulation of glacier material.

The accumulative end moraine produced a differentiated ground stratification in this area. In the bed, up to 20 m in depth, there can be found mainly sands, clay sands and also layers or irregular dumping grounds consisting silts and ashes formed after the thawing of land glacier fragments remaining in terrain depressions. Geotechnical investigations conducted in the area of the moraine hill Kamienna Góra showed that there were present fine and coarse sands coming from Pleistocene as well as fine sands of Miocene. Sand series are separated by series of Pleistocene marginal lake ashes or piled-up glacial clays. The grounds are of very different filtration parameters.

For many centuries the primeval shore cliff of Kamienna Góra has overcome substantial transformations due to sea action. In present in some places the cliff has been changed by anthropogenic activity associated with infrastructure development, as well as by mining useful mineral materials. A wide beach separates the moraine plateau of Kamienna Góra from sea.
EFFECTS OF SEA AND LAND ACTION ON SEASHORE FORMING

Observations of cliff landslides indicate that the main natural factors causing erosion are the following:
- permanent or periodical attacks of storm sea waves against cliff base [4], as a result of which cliff slope retreats step by step back into the land (Fig. 1),
- effects of ground mass forces onto sheer slopes, that results in loss of stability and landslides,
- infiltration and flow of underground water from the side of land towards sea, which liquefies the ground and changes its state (Fig. 2).

Among the anthropogenic factors, many effects associated with development of urban and industrial infrastructure may be counted; these are e.g. embankment structures in the form of tight or cavity walls separating natural groundwater flows which concentrate water within determinate cliff areas or accelerate its filtration in ground due to change in water flow direction.

Infrastructure associated with ground surface hardening may also detrimentally affects water-ground relations since water is not capable of gentle filtering into ground and hence it forms fast running surface streams eroding the cliff.

Landside effects usually cause also destructions of shore protection structures, for instance such as the damage of cliff protection gabion structure shown in Fig. 3.

The deformations could not be caused by storm waves in view of that the beach is a dozen or so meters wide [11]; their causes should be searched in periodical changes in groundwater relations within the cliff itself.

In the case in question it may be assumed that sea effects are negligible. However, there are local areas where changes in landform features happen periodically as a result of loss of ground stability.

COURSE OF THE PHENOMENON

In the night from 23rd to 24th February 2017, in Gdynia, below Sienkiewicz Street, a sudden landslide and run-off of ground-water mass occurred from the escarpment. The slid-down and watered ground mass displaced forming a precipice. Simultaneously, the active tongue of the landslide locally deformed the lie of the land, rising the terrain level by 1,5 m in some places.

A 7 m fragment of the street pavement and roadway was damaged. A geotechnical consequence of the occurred situation is a new state of force equilibrium spontaneously settled in the ground. The geotechnical investigations of the terrain in question show [10] that this state is not stable and may be changed after subsequent intensive watering the ground. A week after the reported event some water flow sources and symptoms of the significant ground watering were found in the site of the landslide.

CAUSATIVE FACTORS

The course of the phenomenon indicates that causes of forming the landslide should be searched out in accumulation of water in the sandy ground covering a layer of clayey and ashy ground of a low filtration ratio. In the entire area of the moraine hill Kamienna Góra some quantities of water collected over dump clay inclusions may be expected.

As a result of the critical water pressure produced in pores of the ground, the hydraulic piping flow and its significant pressure occurred, that caused washing away the road embankment and liquefying the ground. This is a frequent cause of building damages associated with changes in ground-water relations, that is described in the geotechnical literature [2]. In the photographic documentation there are visible a hydraulic piping flow canal and a cavern formed as a result of the washing out of
The geotechnical cross-sections in the line of the landslide have confirmed possible water accumulation in ground lens (Fig. 6). On this basis it is possible to conclude that the geological structure of the base in which ground layers of a higher and lower water permeability alternately appear, helps in forming privileged ways of water filtration in the ground and underground water flows concentrated within a rather small space. In the base there are ashy ground layers which are characteristic of substantially decreased strength parameters in case of watering.

The causative factors can be subjected to a numerical analysis by using a mathematical description method applied in hydrology.

The one-dimensional differential equation describing water flow in the non-saturated ground area has the following form [3]:

\[ \frac{\partial}{\partial z} (D(u) \frac{\partial u}{\partial z}) + K(u) = \frac{\partial u}{\partial t}, \]

where \( u(z,t) \) stands for volumetric moisture content in the ground, \( t \) – time, \( z \) – a coordinate describing the depth in which the phenomenon is analyzed, \( D(u) \) – moisture diffusion coefficient in the ground, \( K(u) \) – filtration ratio along \( z \)-axis. It's worth adding that Eq. (1) describes also many other phenomena which contributed in the damage of the coastal cliff in question. Among them a special case of filtration, i.e. infiltration should be mentioned. In order to determine a solution of Eq. 1, at least one condition should be added. However, in view of that we are to consider the problem in the set \( \Omega = \{ 0, \infty \} \times \{ 0, \infty \} \), i.e. for \( (z, t) \in \Omega \), it is necessary to add to Eq. (1) initial-boundary conditions in the following form:

\[ u(z,0) = \varphi_0(z), \]
\[ u(0,t) = \psi_0(t), \]
\[ u(\infty,t) = \psi_1(t). \]

In the considered case for the qualitative analysis it is sufficient to assume that in Eq. (1), \( K \in \text{Ker} \frac{\partial}{\partial z} D(u) = \text{const} \) and, additionally, that the above given functions \( \varphi_0(z), \psi_0(t), \psi_1(t) \) constitute constants equal to \( \varphi_0, \psi_0, \psi_1 \), respectively. The so obtained initial-boundary problem can be solved with the use of the generalized calculus of operators [6, 7, 8]. Making use of two operations: \( S_1 = -D \frac{\partial}{\partial z} \) and \( S_2 = D \frac{\partial}{\partial z} \), we can write the modified Eq. (1) in an equivalent form as follows:

\[ S_1 u + S_2 u = 0. \]
The solution of the differential equation (4) with the conditions (2), (3) in the operators domain has the form as follows:

\[ u(x, t) = \frac{T_2\varphi_0 + T_1 s_2 u}{T_1 + T_2} = \frac{p_1\varphi_0 + p_2 s_2 u}{p_1 + p_2} \tag{5} \]

where the operations \( T_1, T_2, s \) are determined by the operations \( S_1, S_2, p_1, p_2 \) are the generalized Heaviside operators for the operations \( T_1, T_2 \) [6, 7].

As results from Eq. (5), the decisive (from the point of view of the qualitative analysis of the problem) impact on the course of the phenomenon had the underground water flow with the critical speed. The causative factor was rainfall water migrating within the ground, which was accumulated in more permeable ground layers supported by less permeable ones; such conclusion also results from the mathematical analysis of the problem in question.

**NATURAL PHENOMENA**

The analyzed area of the moraine hill is characteristic of occurrence of local depressions (de-levelling) of 15÷30 m in depth. The de-levelling is associated with occurrence of slopes with 30° steepness. Such large de-levelling results in susceptibility of this area to occurrence of ground mass movements manifested in local landslides of a natural character [4].

The geological data [10] confirm that in the area of Kamienna Góra there are formations stratified by ground layers of different filtration ratios. In this area no free surface pools of groundwater occur, but ground layers are supplied with rainfall and running-off waters from a significant part of the area of the moraine hill Kamienna Góra. If we correlate this fact with the occurrence of the formations stratified by grounds of different filtration ratios, we will get strong argument that the landslide in question was caused by the watering of the ground and, consequently, the loss of its strength properties associated with the watering.

According to the data of the Institute of Meteorology and Water Economy (IMiGW) in Gdynia, a rainfall of almost 14 mm happened 96 hours before occurrence of the landslide. The excessive, long-lasting rainfall water flow into the ground could result in watering the ashy ground layers, lowering their strength parameters as well as exceeding the critical water pressure in ground pores, resulting from the hydraulic piping flow as well as local liquefaction of ground masses.

**LAND DEVELOPMENT**

With land development there are associated anthropogenic factors of changes introduced in urban areas of Gdynia. The infrastructure built in the area of Kamienna Góra resulted in an increased destruction of its natural landform and natural elements protecting slopes against landslide. Among various conditions affecting escarpment stability, there is enumerated a.o. flora which acts also deep into ground due to its root system [12]. Trees and bushes present on the slopes affect water circulation conditions within the ground, causing its water permeability greater in the neighbourhood of the root system. This way, the root system is capable of decreasing excessive water pressure in ground pores by a controlled outflow of water from the ground. Moreover, water demand from the side of plants necessary for their biological processes contributes in lowering ground humidity in their neighbourhood. Additionally, due to transpiration, in the root system area there occur compressive stresses which increase effective stresses in the ground and its shear strength. The substantial influence of plants on escarpment stability is also associated with their root systems which play the role of the so called natural ground reinforcement as the roots are capable of transferring tension stresses, increasing this way ground shear strength that is associated with occurrence of apparent cohesion.

For several years before the landslide event, a devastation of the existing stand of trees has been observed. As a result, a lack of the natural factor holding the ground on the slope has been introduced. The attached photographs reveal a wastage of the stand of trees in this area (Fig. 7a, Fig. 7b).

![Fig. 7. Changes in the stand of trees within the landslide, which occurred between the years 2014 and 2016. (Aerial photographs shot by MGGP AERO)](image-url)
question of channelling the rainfall waters and protecting the slope against sliding. According to the data of the State Geological Institute, the area of the moraine hill Kamienna Góra is classified as a terrain endangered by landslides and exposed to periodical impacts of underground water coming from intensive rainfalls [10].

In the opinion of inhabitants of nearby houses, streams of water appear very fast during heavy rainfalls. It means that the drainage systems of the terrain are ineffective. There is a lack of collectors for the water which appears during heavy rainfall. Roadways are not fitted with channels and drain pipes to discharge rainfall water, and, as a result, the water flows over kerbs and pavements, flooding this way slopes and watering the ground locally. Also, the fact of substantial hardening of the nearby terrain surface is not without a significance as the water coming from a long-lasting rainfall is not able to seep gently into the ground, and in consequence, water streams flow along the streets.

ANTHROPOGENIC TRANSFORMATION OF THE TERRAIN

Beginning from the year 2009, in the landslide area a dwelling house has been built. Many earthworks impairing ground structure have been conducted, natural flow channels of underground water have been cut off or changed. After completing foundation trench, there were observed seepages of water from the ground. During building works a significant water inflow to the trench was noted, and in spring periods it was continuous.

As neither storage reservoirs nor drainage for discharging the ground water were made, the water concentrated at the foot of slope in the site of the damage. According to the project, around the house a French drainage system was provided for discharging rainfall water. This is a solution in which a mineral aggregate carrying water away is wrapped in a geo-unwoven cloth which serves as a filtration layer. The French drainage fulfils its role correctly if only it is connected to a reservoir of a sufficient capacity or retention. In case of an extensive ground water flow a French drainage, if not connected to an appropriate storage reservoir, may be inefficient due to its low retention capability and even hazardous in view of possible infiltration of stored rainfall water into earth, that results in decreasing the strength parameters of the ground, especially in the case of the ground profile in question where layers of ashy ground (ashes, clayey sands, ashy sands) especially exposed to a loss of strength caused a decrease in the ground strength parameters as a result of watering, occur.

In the building site a complete drainage system has not been made. Till the break in the building works only a part of band drainage has been worked out. No storage reservoir, the crucial element of the drainage system, intended for storing the rainfall water, has been built. The completion of only a part of the rainfall water intake and storage system had more detrimental effect onto stability of the slope than a complete lack of it. The drained rainfall water was directed straight into the ground, penetrating its deeper layers, directly.

CONCLUSIONS

Any of the above specified factors itself probably would not be able to cause the landslide. However their cumulative effect and coincidence of unfavourable events finally resulted in the building damage.

The removal of flora from the escarpment area of the plot in question and neighbouring plots exposed the terrain to intensive surface erosion and increased infiltration of rainfall water into the ground, that resulted in the ground strength lowering. Additionally, a lack of transpiration causes that water is not taken up by the plant root system and in the surrounding ground there do not exist respective compression stresses resulting from sucking the water by plants.

Lack of any intake system of rainfall water from the roadway and areas of hardened surface makes water infiltration rate into the ground greater.

The break in the building works, which was introduced without completing realization of the drainage system, resulted in penetration of the taken-in rainfall water deep into the slope and its accumulation over the damage site, that caused a decrease in the ground strength parameters as a consequence of watering the ground.

REFERENCES


CONTACT WITH THE AUTHOR

Eligiusz Mieloszyk
Gdansk University of Technology
Faculty of Civil and Environmental Engineering,
Department of Rail Transportation and Bridges
POLAND

Mariusz Wyroślak
Gdansk University of Technology
Faculty of Civil and Environmental Engineering,
Department of Geotechnics Geology and Marine Civil Engineering
POLAND