DATA WAREHOUSING IN THE OCEANOGRAPHIC SUPPORT OF MARITIME FORCES OF THE POLISH NAVY

Marian KOPCZEWSKI, Bartłomiej PĄCZEK, Marek TOBOLSKI

Summary: The paper presents selected results of the Research focused on the Oceanographic Support Information System of the Polish Navy, seen as its' computer subsystem aimed at the collection, storage and processing of oceanographic data in order to support military Commanders' decisive processes. The system has been modelled with utilization of both Conceptual and Logical Data Modelling.

Keywords: data warehouse, naval operations, oceanographic support.

1. Provision of military oceanographic support to Naval Operations and Exercises

Military Oceanography (MILOC) refers to oceanographic services, data and products that may be used to support NATO Operations. Examples of such services could include acoustic propagation information to support Anti-Submarine Warfare (ASW), water clarity information to support diving operations or surf forecasts to support amphibious or Mine Counter Measures (MCM) operations. Information may be obtained by in-situ or remote sensing measurements, together with advanced computer-based modelling.

Accurate, timely, relevant and reliable MILOC information provides NATO Commanders with the opportunity to plan and execute Maritime Operations. This applies especially to the Joint Force Commander (JFC) and the Maritime Component Commander (MCC).

MILOC data along with meteorology, geo-spatial and supporting remote sensing is collectively referred to as GEO/METOC information. GEO/METOC information can provide the Joint Force Commander with the knowledge necessary to anticipate and exploit the best window of opportunity to plan, execute, support and sustain specific operations. Exploiting GEO/METOC information to prepare the Joint Battlespace, to support planning and intelligence staff and to optimize employment of sensors, weapons, logistics, equipment and personnel is key to effective, efficient and safe conduct of Military Operations. In order to effectively exploit GEO/METOC information, various databases have been established.

The establishment of databases such as the Recognized Maritime Picture Database (RMP) and Recognized Air Picture Database (RAP) are NATO procedures that ensure unity of effort of national forces. The same doctrine is applied to GEO/METOC information through the principle of "one theatre, one forecast" and use of designated maps and datums. GEO/METOC information has spatial and temporal dimensions that permit the information to be databased. The database of physical environmental factors relevant to the particular Area of Operation (AOO) is called the Recognized Environmental Picture (REP).

Strategic, operational and tactical level forecasts and impacts can be derived from the REP and can be made available as a representative picture/graphic or may exist as data that operators may access or query.

Unfortunately, NATO has no organic data sources and is reliant on the nations to provide meteorological, oceanographic and geospatial data and imagery.

Thus, the Polish Navy faced a challenge to create organizational and functional framework of GEO/METOC service aimed to support both national and NATO Maritime Forces and equip it with an appropriate, tailored to specific needs, database system capable of creation of the Recognized Environmental Picture (REP) for the Baltic Sea basin.

1.1. Military Oceanographic Support of the Polish Navy

In order to adapt existing structures of the Polish Navy, responsible for oceanographic support of Maritime Forces operating in the Baltic Sea basin, to the requirements of NATO, the Commander of Naval Operations Centre ordered the organization of Oceanographic Support Information System (Figure 1) and determined the responsibilities of each institution involved.

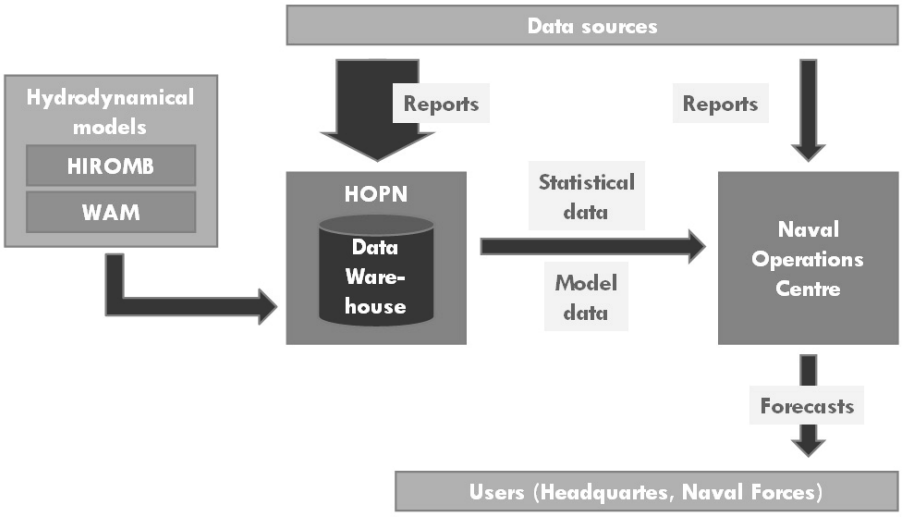

Fig. 1. Oceanographic Support Information System of the Polish Navy

The main organizational unit of the Polish Navy, responsible for analyzing and archiving oceanographic data is the Hydrographic Office of the Polish Navy (HOPN). One of its branches is the Meteorology and Oceanography Branch, tasked with sensing and collection, processing, dissemination and evaluation of oceanographic data.

Usability of METOC data required by Commanders, depends on the timely collection of data to the required temporal and spatial resolution. The foundation of effective METOC support is set by observations from space, air, land and sea by sensors and/or platforms. These observations are essential inputs to numerical models from which METOC services and products are derived.

Due to the rapidly changing nature of the METOC environment, these observations are very perishable and must be continuously updated and made available to METOC organizations. The METOC Officers should review the readily available METOC data in light of the requirements and develop the most up-to-date REP.

Processing METOC data into useable METOC information will enhance Commanders' decision-making cycles. This processing will take place at both national and NATO facilities to include quality control and analysis of METOC data to form a coherent picture of the current state of the METOC environment and also forecast of the future state of the METOC environment. METOC forecasts should be developed for a particular Operation and tailored products should be produced to meet the requirements of Forces.

Current information must be relayed to users in time to be of operational and planning value. Efficient and cost effective solutions that meet the military requirement should be implemented and shortfalls should be highlighted via the chain of command.

To accomplish the above, oceanographic services must have the necessary equipment and devices, including:

- stationary and mobile systems of measurements of the atmosphere and hydrosphere,
- automated data processing systems,
- meteorological and oceanographic databases,
- dedicated data exchange networks.

2. The research

Such an assertion brought forth the need to take Oceanographic Support of Polish Navy Forces, seen as an information process, into consideration and in its framework to examine both the theory and practical implementation of this support in order to examine opportunities to enhance it with the modern tools of computer science.

Nowadays, in fact, technological progress creates new opportunities of processing oceanographic data and the ability to use advanced information systems to conduct the process of oceanographic support.

The authors were inspired to undertake the research by personal experiences, gained during national and NATO exercises which showed, that the available information about the state of the Baltic Sea environment is not fully utilized by the Commanders to plan and conduct Maritime Operations of Naval Forces.

The situation may be changed, in the authors' opinion, by implementation of one of the modern tools of computer science, namely the Data Warehouse, which is defined in the broad sense as a decision support system based on the analytical database, i.e. the data warehouse in the strict sense.

2.1. The object and purpose of the Research

The object of the Research has been identified as: the theory and practice of oceanographic support of the Polish Navy Forces, particularly in terms of the Commanders' decisive processes.

The research process, focused on the object defined above, has been than narrowed.

Therefore:

- the geographical scope of the Research was limited to oceanographic information characterizing the area of the Baltic Sea basin only,

- the functional scope of the Research focused on informational needs of the Polish Navy Forces and their headquarters,
- the structural analyzes were limited to Polish national structures not omitting, however, their relationships with allied structures,
- the data modelling phase of the Research has been limited to the Conceptual and Logical Data Models, stating that the next step of data modelling, i.e. the Physical Model is possible to define with utilization of widely, commercially available CASE tools.

The main objective of the Research was therefore adopted as: to develop functional, organizational and technical recommendations to utilize the Data Warehouse in Oceanographic Support of the Polish Navy Forces in the Baltic Sea basin.

To achieve the main objective, it has been decomposed into detailed ones, adopted as:

Cognitive objectives:

- determination of the state of theory and practice of oceanographic information utilization in the support of the Polish Navy Forces, and
- determination of the state of theory and practice of data warehouses utilization in decision support systems.

Theoretical objectives:

- determination of functional and organizational possibilities of data warehouse implementation in Oceanographic Support of the Polish Navy Forces, and
- development of Conceptual and Logical Data Models of the planned Data Warehouse.

Practical objective:

- development of recommendations for Oceanographic Branches of the Polish Navy on the field of their duties in the Oceanographic Support Information System of the Polish Navy.

2.2. Research problems and hypothesis

The way to achieve the Cognitive research objective, was the answer to the question, which has become the first of three major research problems, namely:

Do theoretical assumptions and practical solutions currently used in Oceanographic Support of Polish Navy Forces lead to the conclusion, that the available oceanographic and environmental data of the Baltic Sea is being used effectively to support decision-making processes of the Polish Navy Forces in the basin?

Theoretical objective could be achieved by solving the next research problem, articulated with a question:

Do theoretical assumptions and experiences gained from the utilization of Data Warehouse to analytical data processing and decision making in enterprises lead to the conclusion, that it is possible to use the Data Warehouse as an element of Oceanographic Support Information System of the Polish Navy?

The achievement of Practical objective of the Research depends on achievement of both Cognitive and Theoretical ones and answering the question, which became the last of the major research problems, namely:

If it is possible to implement the Data Warehouse as a part of Oceanographic Support Information System of the Polish Navy, what actions should be taken by Oceanographic Branches to implement such a system, and whether or not such implementation will bring the expected benefits?

The major research problems became the basis to formulate a hypothesis, which diagnoses under-utilization of analytical capabilities of oceanographic data processing to satisfy needs of decision-makers, and forecasts that implementation of Data Warehouse will desirably change the state of affairs:

Currently applied theories, methods and organizational arrangements of oceanographic support of Polish Navy forces, as well as technical standards, do not take full advantage of currently available analytical tools of computer science to analyze, derived from various sources, oceanographic data, and the remedy to fill that gap is an implementation of information system based on the Data Warehouse.

The hypothesis is twofold, both quantitative and qualitative. Quantitative, because its diagnostic part identifies difficulties in processing rapidly growing amount of oceanographic data, gathered from various sources and qualitative, because it is expected, that implementation of the tool of modern computer science, namely the data warehouse, will integrate available environmental data in one coherent analytical processing system, thus affecting the quality of information obtained from oceanographic data.

2.3. Conduct and results of the research

During theoretical studies, methods of system analysis, synthesis, generalization, and inference have been used.

Empirical research has been conducted with utilization of following research tools and techniques:

- scientific observation of facts, events and processes (during military exercises and trainings of Naval Staffs);
- analysis of experts' opinions and judgments (Naval officers experienced in command posts, Naval Staffs); surveys have been conducted according to the scheme drawn from the Delphi method;
- Conceptual and Logical Data Modelling.

The analysis of the state of oceanographic support of Polish Navy Forces allowed to answer the question, what environmental data is actually used to support decision-makers (Military Commanders of tactical and operational command level) with oceanographic information.

Scientific observation of activities and exercises of the Polish Navy over 4 years may be concluded and synthesized as follows:

Currently, the Polish Navy demands three types of environmental information:

- **current**, reflecting the state of the maritime environment, at the moment of forces' operations,
- **statistical**, reflecting a long-term analyzes of historical data, and
- **modelled**, which is derived from hydrodynamic models of the Baltic Sea.

Research has shown, that currently operated by the Hydrographic Office of the Polish Navy relational database system has been designed to process source data with an assumption of relatively small inflow of incoming data, and existing procedures make it difficult to import a huge amount of data that became recently available from external sources (cooperating Institutions).

The existing Relational Data Base System and, more significantly, procedures greatly extend the workload of Oceanographic Data Base Operator. Thus, there is a considerable delay in the availability of information for the end user.

Another important organizational aspect was the fact, that computer system performing hydrodynamic models calculations was not owned by the military sector. Moreover, the system itself was, as a result of a misunderstandings between the Hydrographic Office of the Polish Navy and the system operator – the Maritime Institute in Gdynia, at the end of 2010, switched off.

Thus, the gap in taking full advantage of currently available analytical tools of computer science to analyze oceanographic data, as diagnosed in the hypothesis, has increased.

The prognostic part of scientific hypothesis suggested replacement of Oceanographic Data Base of the Hydrographic Office of the Polish Navy wit the Data Warehouse, forecasting thereby improvement of the processes of collection, processing, dissemination and evaluation of oceanographic data, both real and modelled.

The forecasted improvement was supposed to be achieved through the utilization of specific features of the data warehouse, which distinguish it from the currently used, operational Oceanographic Data Base.

To prepare a Data Warehouse implementation project in the proposed area, an analysis of the general requirements of data warehouse implementation has been conducted. These requirements were collected in form of universal stages, forming the so-called "Decision Support System life cycle" or "Data Warehouse life cycle".

The following stages have become the key for the Data Warehouse implementation: collection of users' data requirements, development of the database model and creation of the initial set of predefined reports.

Analyzes of the requirements of users (Commanders of Forces) in terms of the scope and timeliness of information revealed, on one hand, that the Commanders of Operational Level are practically not interested in current information, requiring generalized one - statistical or modelled.

On the other hand, Tactical Level Commanders demand reports primarily on current information, being ready in the lack of thereof to replace it with modelled information.

The users' requirements had to be reflected in the Data Warehouse model. Database modelling is considered as the centre of gravity of the implementation process of information system in any enterprise. The three phases of data modelling methodology have been determined - building Data Models respectively: Conceptual, Logical and Physical. They are created in sequence - a Conceptual Data Model is a prerequisite for a Logical Data Model, which is a prerequisite for a Physical Data Model. The physical database and table structures are based directly on the data structures outlined in the Physical Data Model.

The Conceptual Data Model (Figure 2) identifies both the main subject areas of a Data Warehouse and major entities related to that subjects. The main subject areas of Oceanographic Data Warehouse, identified during the research, are: Collection, Processing, Storage and Dissemination of data. The main entities interconnecting these areas are: Oceanographic Branches of the Polish Navy, Forces' Headquarters, Naval Forces and Partners (cooperating Institutions). There is one more, "virtual" entity named "Forecasts", that will be discussed later.

A Logical Data Model presents the entities and relationships of the enterprise. Logical Data Modelling uses Entity Relationship Diagram notation. An Entity Relationship Diagram visually displays the relations between the entities of a enterprise. A Logical Data Model achieves this visual display by focusing on each major subject area from the Conceptual Data Model. Each major subject area of a Conceptual Data Model, therefore,

will become a page by itself in a Logical Data Model. Each page will present all entities relevant to the major subject area.

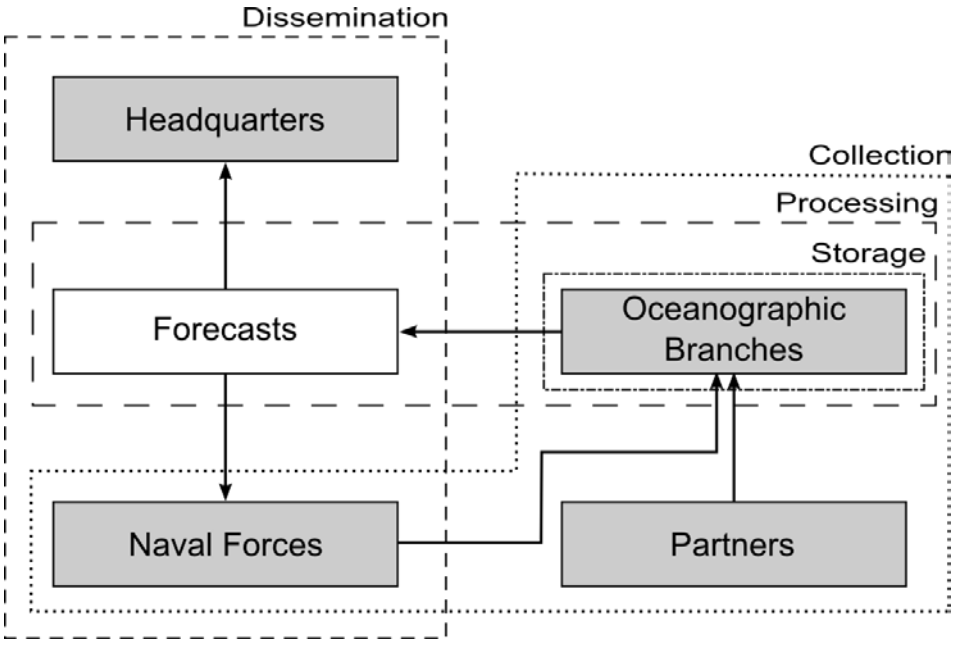

Fig. 2. Conceptual Data Model

Logical Model example page (fragment), reflecting the process of data collection from the Conceptual Model has been shown in the diagram below.

The starting point for defining a multidimensional data warehouse schema is the "Measurements" facts table.

The "Measurements" fact table, which represents the actually measured oceanographic parameters, has been associated with four tables of dimensions, in which the fact (i.e., the actual measurement of oceanographic parameters) took place.

The dimensions are:

- **geographical dimension**, indexing the geographical location of the measurement made, to the appropriate geographic region,
- **time dimension**, storing the time hierarchy of measurement in term of weeks, months, quarters, years and seasons,
- **type dimension**, categorizing measured oceanographic parameters in the sense of measured feature,
- **measuring institution dimension**, reflecting institutions performing the measurements.

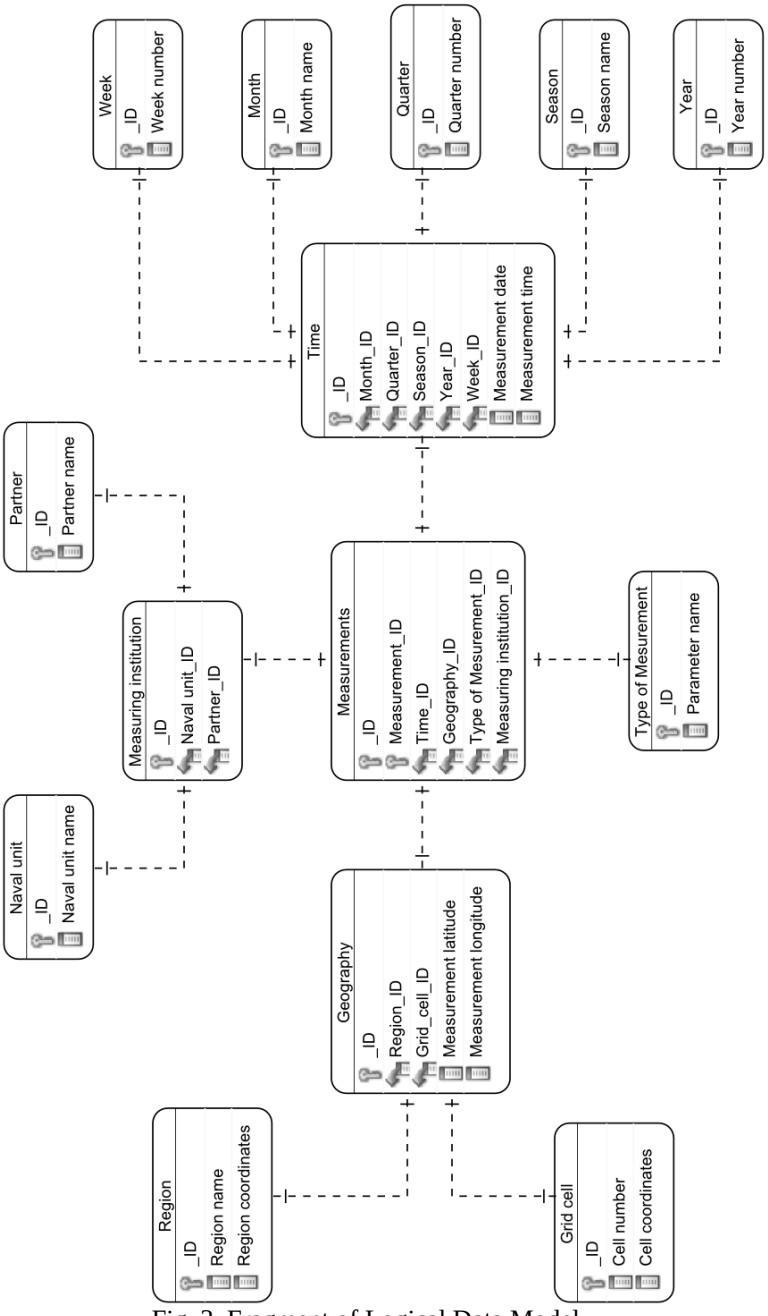

Fig. 3. Fragment of Logical Data Model

It is the Logical Data Model multidimensionality, that translated into the physical structure of the warehouse makes it possible to obtain immediate answers to questions asked by users, related to the previously defined dimensions.

Dimensions are in fact the basis for building, stored in a warehouse, so called, data cubes, i.e., multi-dimensional arrays containing the pre-processed, derived from the sources data.

After loading data into the warehouse, it is necessary to calculate the value of individual cube elements and define principles of the cube's content refreshment. Refresh must be made periodically to ensure a constant supply of current data into cubes underlying the users' analyzes.

Figure 4 shows an example of the data cube, defined on three dimensions: geographical, time and type of measurement dimensions, corresponding to the identified in the course of Research informational needs of users.

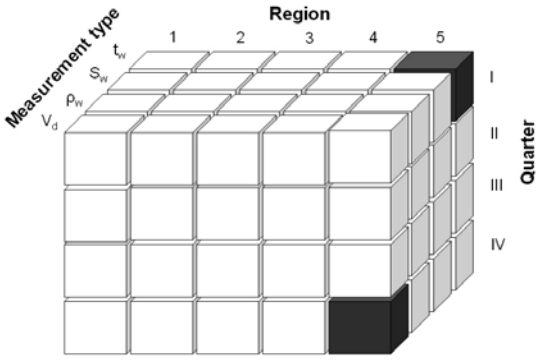

Fig. 4. Sample data cube

(t_w – water temperature, S_w – water salinity, ρ_w – water density, V_d – sound velocity)

For example, the highlighted elementary cells of data cube on Figure 4 hold answers to the questions, routinely asked by users:

- „What is the average water temperature distribution in the region number 2 (e.g. deep of Gdansk), in the first quarter of the year ?", or
- "What is the average sound velocity distribution in the region number 4 (e.g. Bornholm deep) in the second quarter of the year ?".

Appropriate pre-definition of cubes' dimensions allows users to get an immediate answers to routine questions that required, in case of currently used Relational Oceanographic Data Base, a vast number of records' matches, greatly extending the time required to get a reply.

3. Conclusions

Taking the result of the Research and successive stages of the data warehouse lifecycle into account, a hybrid solution has been chosen, as a combination of relational system (ROLAP), with a multidimensional one (MOLAP).

It resulted in storage of oceanographic data in a relational database structures of an existing server, introducing a snowflake schema, and additional utilization of MOLAP tools that operate on multidimensional structures, i.e. data cubes.

This enabled, on the one hand, the ability to integrate data from different sources in a single, coherent structure of the database, on the other hand, it has been prepared for

immediate provision of answers to frequently asked questions of repetitive nature, by building a data cubes with dimensions underpinning routine queries.

The measurable effect is, firstly, the integration of data improving the quality indicators, and secondly, reduction of time required to get answers to routine questions in the area of the Military Oceanography.

References

1. Dolecki A., Pączek B.: Podręcznik Normalizacji Obronnej PDNO-06-A070, Oceanografia wojskowa. Zabezpieczenie sił MW. Ministerstwo Obrony Narodowej, Warszawa, 2009.
2. Inmon W.H.: Building the Data Warehouse. Wyd. John Wiley & Sons Inc., ISBN 0-471-08130-2, USA, 2002.
3. Januszewski A.: Funkcjonalność informatycznych systemów zarządzania. Tom 2, Systemy Business Intelligence, Wyd. PWN, Warszawa, 2008.
4. Kimball R., Ross M.: The Data Warehouse Toolkit. Wyd. John Wiley & Sons Inc., ISBN 0-471-20024-7, USA, 2002.
5. Kopczewski M., Korczak J. Pączek B.: Wybrane modele hydroakustyczne Morza Bałtyckiego w aspekcie informatycznego wspomagania oceny stanu środowiska morskiego na potrzeby działań sił okrętowych MW RP. [w] Knosala R. (red.), Komputerowo Zintegrowane Zarządzanie, Tom I, Oficyna Wydawnicza PTZP, Opole 2008.
6. Kopczewski M., Pączek B.: Data Warehouse as a Part of Defence System. Polish Journal of Environmental Studies Vol. 17, No 3C, Olsztyn, 2008.
7. Kopczewski M., Pączek B.: Funkcjonowanie systemu informacyjnego oceanograficznego zabezpieczenia działań sił MW RP. [w] Kopczewski M. (red.), Modele Inżynierii Teleinformatyki Wybrane Zastosowania, nr 4, Wydawnictwo Uczelniane Politechniki Koszalińskiej, Koszalin, 2009.
8. Silvers F.: Building and Maintaining Data Warehouse. CRC Press, ISBN 978-1-4200-6462-9, USA, 2008.
9. Stackowiak R., Rayman J., Greenwald R.: Oracle Data Warehousing and Business Intelligence Solutions. Wyd. Wiley Publishing, Indianapolis, Indiana, 2007.
10. Szubrycht T. (red.): Leksykon bezpieczeństwa morskiego. AMW, Gdynia, 2008.

Dr hab. inż. Marian KOPCZEWSKI
Dr Bartłomiej PĄCZEK
Wydział Dowodzenia i Operacji Morskich
Akademia Marynarki Wojennej
81-103 Gdynia, ul. Śmidowicza 69
tel./fax: (58) 626 26 27
e-mail: b.paczek@amw.gdynia.pl

Mgr inż. Marek TOBOLSKI
Delegatura Wojskowej Ochrony Przeciwpożarowej w Gdyni
ul. Jana z Kolna 8A, 81-301 Gdynia
tel./fax.: (0-58) 6266178
e-mail: marektobolski@tlen.pl