Available online at www.sciencedirect.com

ScienceDirect

Procedia Engineering 00 (2017) 000–000

Procedia Engineering

www.elsevier.com/locate/procedia

Creative Construction Conference 2017, CCC 2017, 19-22 June 2017, Primosten, Croatia

# Prognostic and diagnostic capabilities of OOBN in assessing investment risk of complex construction projects

Magdalena Apollo*

*Faculty of Civil and Environmental Engineering, Gdansk University of Technology, Narutowicza 11/12 Str. 80-233 Gdansk, Poland*

**Abstract**

Modelling decision problems using Bayesian networks is extremely valuable especially in case of issues related to uncertainty; it is also very helpful in constructing and understanding visual representation of the elements and their relations. This approach facilitates subsequent application of Bayesian networks, however there can be situations where using simple Bayesian networks is impractical or even ineffective. The aim of this article is to present object-oriented Bayesian networks (OOBN) in the context of modeling investment risk. OOBN not only allow decomposition of a complex model into individual objects reflecting different groups of issues (for example risk areas) but also allow modeling time dependencies between those objects.

The use of object-oriented Bayesian networks is presented using an example of urban regeneration project. On the basis of a complex construction project the author presents both advantages and disadvantages of OOBN in terms of diagnostic and prognostic efficiency.

In course of the research it has been observed that during the construction of large Bayesian networks the possibility to automatically generate node probability tables is very useful, as it significantly accelerates construction of this type of models. The author also indicates additional recommendations in the field of defining object-oriented Bayesian networks instrumental in assessing investment risk of complex construction projects.



Peer-review under responsibility of the scientific committee of the Creative Construction Conference 2017.

*Keywords:* Bayesian Networks; complex investment projects; OOBN; risk assessment; urban regeneration.

* Corresponding author. Tel.: +48-58-347-16-37
*E-mail address:* magdalena.apollo@wilis.pg.gda.pl

1877-7058 © 2017 The Authors. Published by Elsevier Ltd.
Peer-review under responsibility of the scientific committee of the Creative Construction Conference 2017.

## 1. Introduction

Risk management as integral part of investments process takes a special meaning in the multidimensional process of preparation and realization of construction projects. The characteristics of a final product make this type of projects unique, therefore risk management and making investment decisions involves inherent uncertainty.

Bayesian networks (BN) (described in more detail in publication [7,10]) are useful in modelling investment risk, as they help to describe and understand the situation, are a visual representation of its elements and relations between them, simplifying the application of built structures at a later stage. There are however two types of situations where use of a simple network (i.e. direct representation of all elements) is impractical or even ineffective. Such situations include complex problems, where the model has too many variables (nodes), as well as situations where it is necessary to represent time sequence. The solution to the above difficulties is decomposition of a complex model down to its single objects, which represent different groups of issues, and can also represent time dependencies between objects [8]. This paper aims at presenting the first aspect of object-oriented Bayesian networks (OOBN), which is assessment of the investment risk of complex construction projects using an example of urban regeneration (excluding the time aspect of OOBN models).

## 2. Application of OOBN in the risk assessment of urban regeneration projects

Modelling investment risk using OOBN can be especially useful in the initiation and preparation phases of the project. Special types of risks occurring in such projects (described in detail in publications [2], [4], [5], [6], [9]) make the investment risk of urban regeneration projects very difficult to assess, especially at the initial stages of investment process. "Due to lack of historical data allowing to assess the probability of occurrence of specific circumstances (risk factors), as well as high level of complexity of urban regeneration, it can be stated that such projects involve not only risk but an inherent uncertainty. In case of this type of projects the standard tools and methods of quantitative risk analysis (for example: assessment of the risk factors register, event tree analysis, simulations using the Monte Carlo method) do not apply" [2]. Those methods are based usually on simple, regression statistical models, which require specific historical data. What is more, such methods may not be able to model the relations between the system elements.

In the risk assessment the commonly used efficiency indicators (such as NPV, IRR) are insufficient, as the research has proved, the project may be financially unprofitable and have a high risk level due to that (conclusion based on analysis of four urban regeneration projects completed in Gdansk [11], [12], [13], [14]), and despite that authorities still decide to conduct this types of projects. This is caused by the fact that in decision making process other, non-financial factors are considered.

A holistic approach to complex projects is necessary; it allows to assess the level of risk taking into account relations between specific elements of the process (risk factors), creating a network of interdependencies. This approach is made possible using Bayesian networks. Their use allows to effectively apply existing knowledge base (expert knowledge) to build the network's structure (i.e. to identify the significant elements and relations between them), as well as its parameterization (defining the strength of those relations) through node probability tables (NPT). Bayesian networks allow the updates of the initial probabilities of all elements of the network through addition of new information from tests and observation. This in turn reduces the uncertainty and allows to build updated risk scenarios, resulting in more informed decision making.

Simple, direct representation of all risk factors in a complex project is not an optimal solution, to alleviate this situation in the course of research an object-oriented Bayesian network has been created.

## 3. Research model

The Bayesian network model, as a representation of risk factors and their groups, generating urban regeneration investment risk was built in specialized software AgenaRisk [1]. This model in local approach accounts for risk factors identified during expert polling as having significant impact on potential increase of cost, due to additional works required to complete the project (the total number of risk factors included in the risk register was therefore reduced from 98 to 42). Apart from risk factors, in line with methodology of constructing a Bayesian network, the model includes also nodes (variables) allowing to model the condition of the project. The following variables were defined: background, symptom, problem and intermediate ones. The background variables (green in Fig. 1) and symptom variables (pink in Fig. 1) when populated with information about actual project conditions allow to update assessment of real risk level in the project. The problem variables relate to diagnosis or prognosis (light blue and dark blue in Fig. 1) and the intermediate variables (orange in Fig. 1) are those which states cannot be directly observed. They play an important role in the process of obtaining the correct conditional dependence and independence and / or for the effective inference.

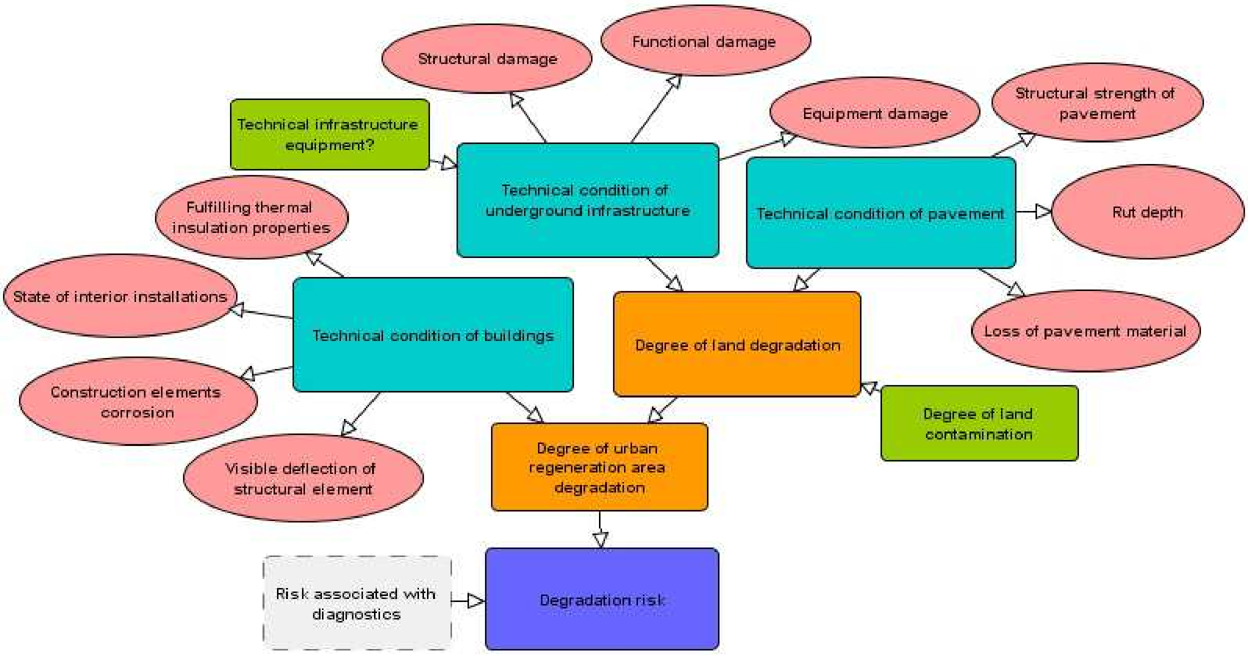


Fig.1. Example of a local structure used to assess degradation risk (input node marked in grey).

Data necessary for qualitative analysis (relations between specific objects and risk factors), as well as quantitative analysis of the network (*a priori* probability, node probability tables - NPT) was obtained during consultations with experts. In the context of probability *a priori* it has resulted in constructing histograms representing the frequencies of selected probability values of defined risks occurrence. NPT definition was based on expert’s and author’s knowledge, used also during their automated generation.

The goal of model verification is to check if the constructed model generates correct and logical answers in line with expert’s expectations. In order to achieve this several phases of verification have been conducted: based on scenario analysis and sensitivity analysis using tornado charts. The abovementioned analysis was used not only to verify the correctness of model outcomes, but mainly to identify the factors, which occurrence significantly affects urban regeneration investment risk.

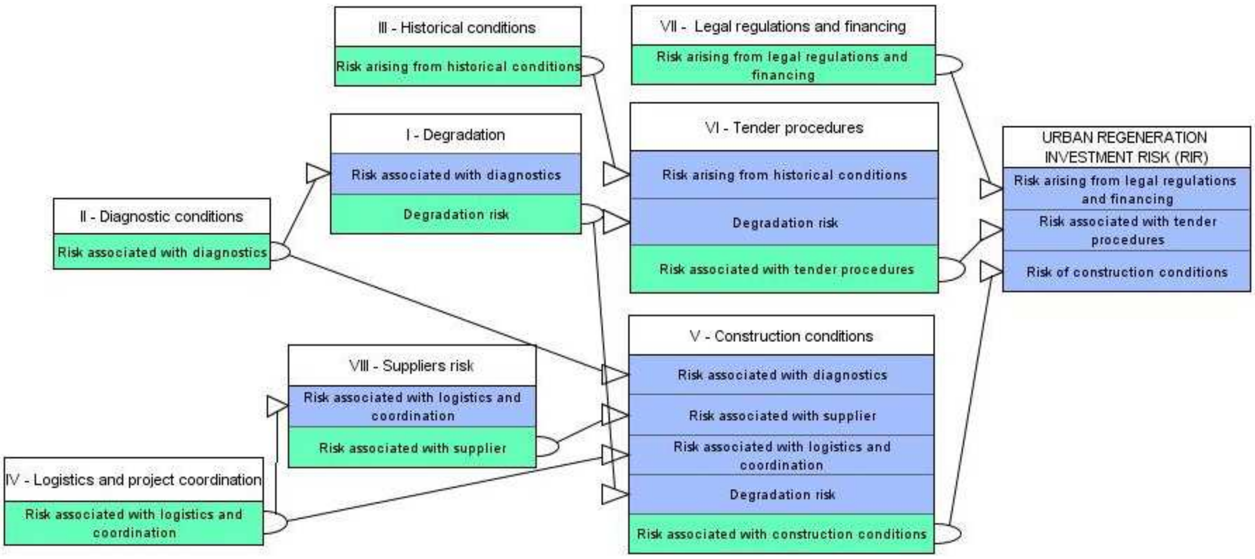


Fig.2. Object-oriented Bayesian network for the urban regeneration investment risk structured in AgenaRisk software (global approach).

Table 1. Concepts used by AgenaRisk software in OOBN modelling.

| Graphic representation for global view | Characteristics |
|---|---|
| | OBJECT |
|   | • specific dependencies structure, separated from global model,<br>• local network view of OOBN, grouping the most significant risks identified into thematic substructures, forming part of global model,<br>• initial object (without predecessors) has only output nodes, final (without successors) has only input nodes, intermediate includes both input and output nodes,<br>• more flexible than direct representation of all variables, which gives wide possibilities of constructing alternative relations and analysis of risk scenarios (replication of the basic feature of object modelling – abstract aspect) |
| | OUTPUT NODE |
|   | • in AgenaRisk software marked in green,<br>• has to be of the same type as input node (in the model nodes are marked as ranked[†]) and include the same set of possible states, so information can be passed (set of marginal probability values represented by the histogram),<br>• can be connected to input nodes of several objects,<br>• node allowing to pass information between objects, visible in local view of individual objects and in global OOBN view |
| | INPUT NODE |
|   | • in AgenaRisk software marked in blue,<br>• has to be of the same type as output node and include the same set of possible states, so information can be received (set of marginal probability values represented by the histogram),<br>• it is possible to connect only a single output node with a single input node,<br>• node allowing to pass information between objects, visible in local view of individual objects and in global OOBN view |

[†] Ranked variables – variables, which states are ordered on a scale, for example low, medium, high.

Keeping in mind the goal of building object-oriented Bayesian network [6], the identified risk factors (together with other abovementioned variables) have been grouped into 8 separate objects (thematic substructures) presenting different areas of urban regeneration investment risk (Fig.2).

Global representation of dependencies between specific objects, as defined during expert consultations and based on author's experience is shown on Fig.2. Sample structure of dependencies between each variable in the model (in their local approach) is shown in Fig. 1. It is important to note that relating specific model elements in their global approach, requires prior definition for each object of the input and output nodes within local structures (nodes visible in Fig,2). Those nodes are crucial in passing information, therefore their detailed characteristics has been presented in Table 1.

## 4. Sensitivity analysis

The network of dependencies, built as OOBN model, has been constructed in order to support the process of arriving at a conclusion in the circumstances of uncertainty. It is important not only to know the final level of urban regeneration investment risk, but also, possibly above other aspects, to know the factors (areas) that impact the risk level the most. To achieve this goal a sensitivity analysis has been conducted, which by default is used to verify the correctness of model outcomes, however can also be an element of risk management through identification of key risk areas.

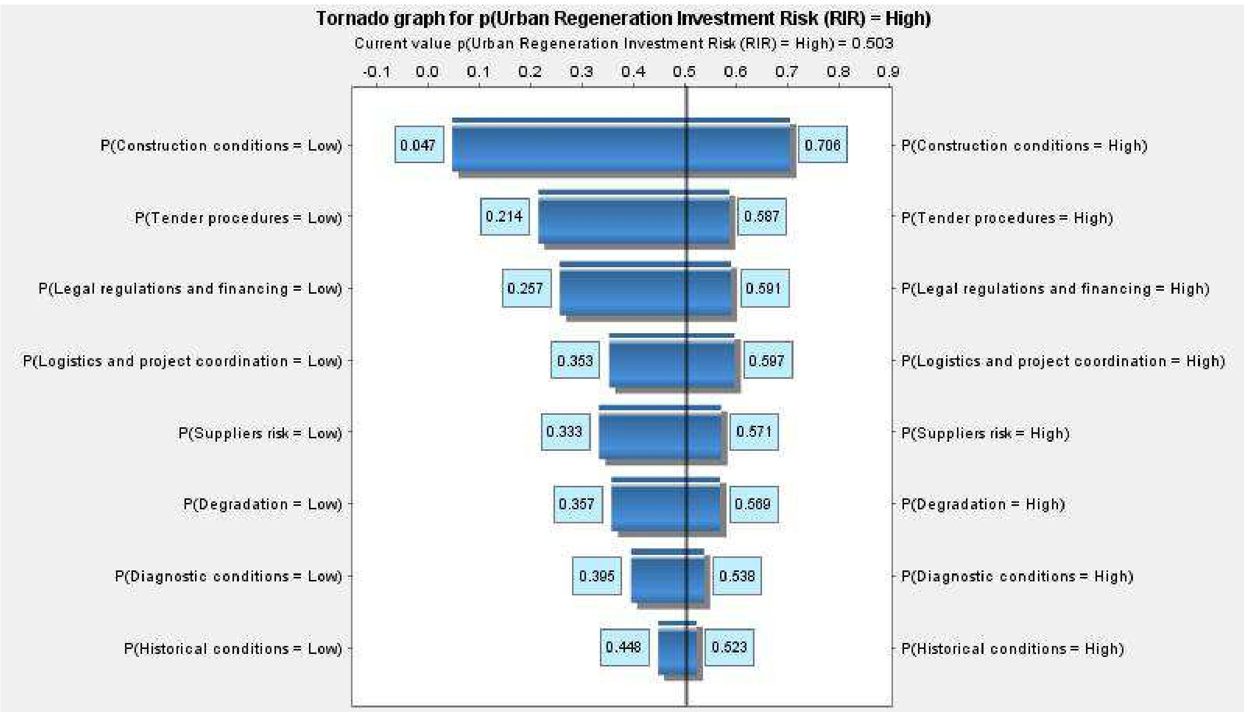


Fig. 3. Tornado graph for the target node Urban Regeneration Investment Risk (RIR) taking the state: High.

The purpose of the model sensitivity verification was to check if propagation of information generates outcomes that are in line with expert's expectations. To achieve this Tornado charts have been generated identifying the minimal and maximal values of *a posteriori* probability for each possible state of the target node if specific observations are input into the model. What is important, the OOBN model is taking into account pessimistic scenario representing extremely negative conditions of urban regeneration projects so the outcome of the sensitivity analysis can be used in for risk and uncertainty management.

Tornado analysis presented in Fig.3 is a summary of considerations regarding the object-oriented Bayesian network. It represents the impact of specific substructures (risk areas) on urban regeneration investment risk (target node) assuming unfavourable conditions of urban regeneration projects.

Taking into account those observations it can be concluded that the key impact on urban regeneration investment risk has the area related to the construction conditions (object no V on the Fig. 2; fluctuation of probability for the target node range from 4.7%÷70.6%). Such high impact is caused by the direct relation between the variables, however it should also be noted that the realization phase of construction investment process brings together the impact of four other risk areas (diagnostics conditions, the level of degradation, logistics and coordination as well as suppliers, what can be observed on the Fig. 2).

Slightly smaller, but equally significant impact on the target node has the area *VI – Tender procedures* and *VII – Legal regulation and financing*. Narrower range of variance of the probability value for the above areas is a result of indirect relation to the target node (a result of construction limitations of AgenaRisk software). The risk related to tender procedures brings together the impact of two other areas, therefore any change in conditions of object I or III can significantly increase risk level of tender procedures.

The following areas have an indirect and decreasing impact on the target node (urban regeneration investment risk): risk related to logistics and project coordination, supplier risk, level of degradation and diagnostics risk. Fluctuations of probability in the above areas are similar and relatively small, but any change in their conditions can have significant impact on the risk related to construction conditions.

Relatively the smallest impact on urban regeneration investment risk has area *III – Historical conditions*, its variance of probability for the target node fit in the range of 44.8%÷52.3%.

Results of this sensitivity analysis as well as number of simulations can be used in the next phase of research to help direct actions related to risk mitigation.

## 5. Conclusion

### *5.1. General observations*

Among the scientists and practitioners of risk management there are opinions that quantitative risk analysis makes sense only if the parameters required for its assessment can be precisely defined. It is hard to resist the impression that a completely objective and fact based risk analysis of construction investment project does not exist, taking into account that every construction project is considered unique. As a consequence it can be claimed that conducting this type of analysis for holistic projects, characterized by complexity, including a number of subjects – active and passive participants in the investment process, is pointless. The number of variables impacting the time and cost of the investment, as well as its scope and quality is so big that it is difficult to define the impact of specific risks on the basic project assumptions. Does that mean that such analysis should not be conducted?

In the opinion of the author urban regeneration projects require a holistic analysis both qualitative and quantitative. There are difficulties in precise analysis of potential time and cost of investment realization. The multiplicity of various cost factors and actions (not only construction related) that form parts of urban regeneration projects makes such analysis lose its practical value – conducting calculations and simulations would be extremely difficult. As was shown the complexity of urban regeneration projects requires taking into account causal relations between identified risk factors, which adds another layer of complication. “From the risk management standpoint, it is more important [than defining time and cost schedule of the investment] to identify critical risk factors, that threaten project realization in the within the given time and cost” [15]. Because of the abovementioned reasons in this paper only the level of risk related to a given urban regeneration project has been analyzed, as reflected in the level of probability of certain occurrences (threats). In such approach the prognostic and diagnostic aspect of OOBN is applied, as it allows to identify in the model the variables (risk factors) that have the most significant impact on urban regeneration investment risk.

The constructed Bayesian network is object oriented, and together with previously created risk register forms an integrated approach to risk management in complex construction investment process that has high level of uncertainty. This uncertainty is mostly caused by unique and social characteristics of urban regeneration projects, resulting in the need of unconventional approach to risk management.

Bayesian networks considered in this paper are described as object oriented, because they have some characteristics of this type of modelling. The main advantage of OOBN is the possibility of modelling objects that are more universal and inherently consistent, which gives wide possibilities in various risk scenarios analysis. Using AgenaRisk software it was also possible to create both local (Fig. 1) and global structures (Fig.2).

Using Bayesian network for modelling we have several ways of representing the uncertainty related to analyzed objects. Other than node probability tables in the proposed network the concept of variance is also used. Another, not applied in this paper, way of expressing the uncertainty is the inclusion in the model construction a utility node (or nodes), which allows to reflect the investor's approach to risk (approach used in decision models). The use of utility node is disputable. If the basis of making the decision using the utility node are real financial values (loses caused by risk occurrence, diagnostics costs, fines, costs of additional works etc.) it is difficult to argue such data (the topic described in more detail in [4]). However if the decision maker's preference are expressed in non-measurable numeric values they introduce into the network another element of subjectivity that is widely criticized in subject literature and by risk management practitioners. In this context the approach proposed in this paper is more credible and objective.

### *5.2. Conclusions and observations regarding OOBN modelling*

While constructing large networks the ability to automatically generate node probability tables significantly speeds up the process. After deciding the right mathematical function or function of probability distribution and establishing its parameters (mean and variance assessed by experts), it is possible to automatically generate the tables. In order to verify conditional probabilities created this way it is recommended to convert to manual mode of data input to "sharpen" the extreme cases of conditional probability. The use of automated generator results also in other risks. While inputting the parameters of the mathematical functions one defines, among others, the variance, being a reflection of our opinion regarding the average value of distribution (in AgenaRisk software the variance value is related to the level of certainty, it can be interpreted as the precision of inputted information). Defining this parameter's value at low level of certainty (higher than 0.1) causes the output to become vague or even insignificant, therefore in large models it is necessary to set this parameter close to zero, which results in higher model sensitivity to the simulations conducted. This observation is significant as object-oriented Bayesian networks generally have issues with information propagation. Connecting the objects into larger network causes the information input on the edges of OOBN can have insignificant impact on the final result. Due to that the correct definition of automatic distribution function on the high level of precision is even more significant

To conclude there is one more significant limitation of OOBN. Using object-oriented Bayesian networks allows conducting the complete calculation only in the context of prognostic, because information is passed only from predecessor object to the subsequent one. Diagnostic calculations are not possible, because the preceding object is not impacted by the observations populated in the subsequent object. In the global view the information input into the network is only passed forward, which provides us prognostic calculations. The option of backward propagation (diagnostics) is still possible however in the local structure of each object.

## References


[1] AgenaRisk 7.0 Software, (2016).

[2] Apollo M., Brzezicka J.: Sources of investment risk in urban regeneration projects". Czasopismo Techniczne nr 1-B/2014

[3] Apollo M., Grzyl B.: Problemy koordynacji realizacji robót budowlanych na przykładzie rewitalizacji dzielnicy Gdańsk-Letnica. Inżynieria Morska i Geotechnika 5/2013. Gdańsk

[4] Apollo M., Kembłowski M.W.: Observation value analysis – integral part of Bayesian diagnostics. Procedia Engineering 123C (2015), pp. 24-31

[5] Apollo M.: Koncepcja modelu oceny ryzyka rewitalizacji obszarów miejskich. Materiały konferencyjne konferencji pt: „Wpływ Młodych Naukowców na Osiągnięcia Polskiej Nauki - Nowe trendy w naukach inżynieryjnych - IV Edycja". ISBN 978-83-63058-27-2. Gdańsk 12-14.04 2013

[6] Apollo M, Kristowski A.: Logistyka realizacji robót budowlanych podczas rewitalizacji dzielnicy Gdańsk-Letnica. Logistyka 3/2012

[7] Apollo M., Miszewska-Urbańska E.: Podejmowanie decyzji w warunkach niepewności przy użyciu sieci Bayesa – przykład zastosowania. Logistyka nr 6/2014, pp. 1496- 1504

[8] Fenton, N., Neil, M.: Risk Assessment and Decision Analysis with Bayesian Networks. CRC Press, Boca Raton (2012).

[9] Grzyl B., Kristowski A.: BIM jako narzędzie wspomagające zarządzanie ryzykiem przedsięwzięcia inwestycyjnego, Materiały budowlane 6/2016, pp. 52-54
[10] Siemaszko, A.; Kembłowski, M. W. Ocena efektywności monitoringu obiektów inżynierskich za pomocą sieci Bayesa. Materiały Budowlane, 2016.
[11] Studium Wykonalności dla Projektu „Rewitalizacja Dolnego Miasta w Gdańsku". Załącznik nr 1 do Wniosku o Dofinansowanie w Ramach Poddziałania 3.2.1. Kompleksowe Przedsięwzięcia Rewitalizacyjne RPO dla WP na lata 2007-2013, (2010).
[12] Studium Wykonalności dla Projektu „Rewitalizacja Dolnego Wrzeszcza w Gdańsku". Załącznik nr 1 do Wniosku o Dofinansowanie w Ramach Poddziałania 3.2.1. Kompleksowe Przedsięwzięcia Rewitalizacyjne RPO dla WP na lata 2007-2013. Gdańsk, (2009).
[13] Studium wykonalności dla projektu „Rewitalizacja Letnicy w Gdańsku", Załącznik nr 1 do wniosku o dofinansowanie w ramach poddziałania 3.2.1 Kompleksowe Przedsięwzięcia Rewitalizacyjne Regionalnego Programu Operacyjnego dla Województwa Pomorskiego na lata 2007-2013. Gdańsk, (2009).
[14] Studium Wykonalności dla Projektu „Rewitalizacja Nowego Portu w Gdańsku". Załącznik nr 1 do Wniosku o Dofinansowanie w Ramach Poddziałania 3.2.1. Kompleksowe Przedsięwzięcia Rewitalizacyjne RPO dla WP na lata 2007-2013. Gdańsk, (2009).
[15] Teixeira, J.C., Kulejewski, J., Krzemiński, M., Zawistowski, J.: Zarządzanie ryzykiem w budownictwie. , Guimaraes, Warszawa (2011).