Analiza ryzyka kosztowego robót remontowo-budowlanych w warunkach niepełnej informacji

Mgr inż. Michał Bętkowski, dr inż. Andrzej Pownuk
Wydział Budownictwa
Politechnika Śląska w Gliwicach
Michal.Betkowski@polsl.pl, Andrzej.Pownuk@polsl.pl

Streszczenie

Ryzyko jest nierozłącznie związane z każdym przedsięwzięciem budowlanym. Szczególnie wysoki poziom towarzyszy prowadzeniu prac remontowych – z uwagi na trudny do określenia zakres robót w chwili planowania. W artykule autorzy przedstawili próbę rozwiązania tego problemu z zastosowaniem autorskiej koncepcji zmiennej losowej o parametrach rozmytych [3] oraz specjalistyczne narzędzie pozwalające uwzględnić niepewny charakter danych [4].

Słowa kluczowe: Ryzyko, algebra rozmyta, niepewność danych.

Risk is an immanent part of all civil engineering projects. Particularly one of the highest levels of risk accompanies repair work. In this case it is very difficult to predict a range of different tasks at the stage of planning. In this paper the authors try to resolve above described problems using a new concept of random variable with fuzzy parameters. [3] as well as a computer program which is able to take advantage of data uncertainty and new theoretical results [4].

Keywords: risk, fuzzy algebra, data uncertainty.

Wstęp

Prowadzenie prac remontowych wiąże się ze szczególnie dużym ryzykiem. Z uwagi na trudny do przewidzenia zakres [1] , stosunkowo rzadko mamy do czynienia z umowami ryczałtowymi – na ogół ceny jednostkowe mają charakter stały – natomiast rozliczenie wartości prac następuje na podstawie obmiarów robót i kosztorysu powykonawczego sporządzanego przez wykonaniu robót.
Dlatego istotną kwestią jest poszukiwanie narzędzi umożliwiających określenie przewidywanej ceny kosztorysowej robót i ryzyka z tym związanego.

$$C_K = \sum_{i=1}^{N} L_i \cdot Cj_i \tag{1}$$

Z jednej strony mamy stałe ceny jednostkowe Cj_i – wynikające z umowy – z drugiej trudną do przewidzenia na etapie planowania, ilość faktycznie wykonach jednostek L_i. Można to zjawisko potraktować jako w pewnej mierze problem losowy – błąd szacowania wielkości pomiaru. Typowym rozkładem charakterystycznym dla tego typu zjawisk jest rozkład normalny – jednak stykamy się tu z problemem braku wiarygodnego materiału statystycznego pozwalającego na oszacowanie parametrów charakterystycznych.

Rozwiązaniem tego problemu może być sięgniecie po autorską koncepcję zmiennej losowej o parametrach rozmytych [3] - pozwala ona połączyć zalety zmiennej losowej oraz zmiennej rozmytej eliminując w dużej mierze ograniczenia obu tych podejść – (zmienna losowa wymaga zbioru danych pozwalających na oszacowanie jej parametrów – zmienna rozmyta jest mało odporna na propagację błędu).

Matematyczny opis ryzyka

Z ryzykiem mamy do czynienia wtedy, gdy poniesione koszty realizacji projektu C_K będą większe od założonego poziomu ceny kosztorysowej C_{KZ}.

$$\begin{cases} C_K \le C_{KZ}, \text{ brak ryzyka} \\ C_K > C_{KZ}, \text{ ryzyko} \end{cases} \tag{2}$$

W przypadku, gdy dysponujemy odpowiednio dużą liczbą informacji, to możemy statystycznie opisać poniesione koszty realizacji C_K przy pomocy zmiennej losowej.

$$C_K : \Omega \ni \omega \to C_K(\omega) \in R \tag{3}$$

Teraz można zdefiniować ryzyko jako prawdopodobieństwo przekroczenia założonego poziomu cen.

$$R(C_{KZ}) = P\{\omega : C_K(\omega) > C_{KZ}\} \tag{4}$$

Jak widać ryzyko jest funkcją założonego poziomu ceny kosztorysowej C_{KZ}. Jeśli znamy dystrybuantę zmiennej losowej C_K,

$$\Phi_{K_C}(C_{KZ}) = P\{\omega : C_K(\omega) < C_{KZ}\} = \int_{-\infty}^{C_{KZ}} f_{C_K}(x)\,dx \tag{5}$$

to ryzyko może zostać określone następująco

$$R\left(C_{KZ}\right)=P\left\{\omega: C_K\left(\omega\right)>C_{KZ}\right\}=1-P\left\{\omega: C_K\left(\omega\right)\leq C_{KZ}\right\}=1-\Phi_{C_K}\left(C_{KZ}\right) \quad (6)$$

Przedziałowe charakterystyki ryzyka

W przypadku robót remontowych rozliczanych według obmiarów – pomimo zryczałtowanych cen jednostkowych trudno przewidzieć końcową cenę kosztorysową na etapie planowania. Rozrzut ilości planowanych jednostek, a rzeczywiście wykonanych może być powodowany niepewnościami jakie należy uwzględnić w tym procesie. Niepewności można ogólnie podzielić na trzy grupy:

- niepewności przypadkowe – wynikające z losowego charakteru wielu zjawisk,
- niepewności systematyczne – wynikają z braku wiarygodnego źródła informacji,
- błędy grube – zwykle są to różnego rodzaju pomyłki.

W przypadku, gdy mamy wystarczającą ilość danych niepewności przypadkowe można opisywać przy pomocy teorii prawdopodobieństwa. W takim przypadku zakładamy, że dysponujemy pewnym zbiorem pomiarów danej ilości jednostek przedmiarowych L_i :

$$L_{i,1}, L_{i,2}, ..., L_{i,N_i} \quad (7)$$

Liczby (7) można uznać za realizacje pewnej zmiennej losowej. W takim przypadku można konstruować probabilistyczne charakterystyki (statystki) opisujące zmienną losową $L_i : \Omega \ni \omega \rightarrow L_i(\omega) \in N$. Najczęściej obliczamy wartość średnią i/lub wariancję:

$$\overline{L_i} = \frac{1}{N_i}\sum_{j=1}^{N_i} L_{i,j}\ , \quad s^2_{\overline{L_i}} = \frac{1}{N_i}\sum_{j=1}^{N_i}\left(L_{i,j} - \overline{L}_{i,j}\right)^2 \quad (8)$$

Przy pomocy obliczonych statystyk możemy dokonywać dalszej obróbki statystycznej otrzymanych danych.

Należy jednak podkreślić, że każdy pomiar $L_{i,1}, L_{i,2}, ..., L_{i,N_i}$ obarczony jest również błędami systematycznymi, a czasem również grubymi. Zatem zamiast punktowej wartości pomiaru $L_{i,j}$ należy rozważyć cały przedział $\hat{L}_i = \left[L_i^-, L_i^+\right]$ możliwych wartości, które może przyjmować:

$$L_i \in \hat{L}_i = \left[L_i^-, L_i^+\right] \quad (9)$$

Górną i dolną granice przedziału można zwykle oszacować na podstawie ogólnej wiedzy o danym koszcie L_i. Dowolna statystyka h (np. wartość średnia, wariancja itp.) zmiennej losowej L_i jest funkcją danych pomiarowych

$$h = h\left(L_{i,1}, L_{i,2}, ..., L_{i,N_i}\right) \tag{10}$$

Zatem jeśli informacje będą przedziałami $\hat{L}_{i,1}, \hat{L}_{i,2}, ..., \hat{L}_{i,N_i}$, to statystyka h też będzie przedziałem

$$\hat{h} = h\left(\hat{L}_{i,1}, \hat{L}_{i,2}, ..., \hat{L}_{i,N_i}\right) = \left\{h\left(\hat{L}_{i,1}, \hat{L}_{i,2}, ..., \hat{L}_{i,N_i}\right) : L_{i,j} \in \hat{L}_{i,j}\right\} \tag{11}$$

Przykładowo przedziałową wartość średnią można określić następująco:

$$\hat{\bar{L}}_i = \frac{1}{N}\sum_{j=1}^{N_i}\hat{L}_{i,j} = \left\{\frac{1}{N}\sum_{j=1}^{N_i}\hat{L}_{i,j} : L_{i,j} \in \hat{L}_{i,j}\right\} \tag{12}$$

To samo dotyczy dowolnych innych probabilistycznych charakterystyk takich jak np. funkcji gęstości rozkładu prawdopodobieństwa $f_{L_i}(x)$. W tym przypadku otrzymujemy przedziałową funkcję gęstości rozkładu prawdopodobieństwa $\hat{f}_{L_i}(x)$.

$$\hat{f}_{L_i}(x) = \left\{f_{L_i}\left(x, L_{i,1}, ..., L_{i,N_i}\right) : L_{i,j} \in \hat{L}_{i,j}\right\} \tag{13}$$

gdzie $f_{L_i}\left(x, L_{i,1}, ..., L_{i,N_i}\right)$ jest empiryczną funkcją gęstości otrzymaną na podstawie pomiarów $L_{i,1}, ..., L_{i,N_i}$.

Ryzyko może zostać oszacowane jest na podstawie zbioru pewnych informacji dotyczących ilości jednostek przedmiarowych $L_{i,1}, ..., L_{i,N_i}$. Zbiór ten można przedstawić jako macierz o zmiennej liczbie elementów w każdym wierszu.

$$\mathbf{L} = \begin{bmatrix} L_{1,1}, ..., L_{1,N_1} \\ ... \\ L_{M,1}, ..., L_{M,N_M} \end{bmatrix}, \quad \hat{\mathbf{L}} = \begin{bmatrix} \hat{L}_{1,1}, ..., \hat{L}_{1,N_1} \\ ... \\ \hat{L}_{M,1}, ..., \hat{L}_{M,N_M} \end{bmatrix} \tag{14}$$

Więc ryzyko jest funkcją zbioru danych $\mathbf{L}$ oraz założonego poziomu kosztów.

$$R = R(C_{KZ}, \mathbf{L}) \tag{15}$$

gdzie C_{KZ} jest założonym poziomem kosztów oraz ryzyko określone jest przy pomocy wzoru (4). Zatem przedziałową wartość ryzyka $\hat{R}(C_{KZ})$ otrzymujemy następująco:

$$\hat{R}(C_{KZ}) = \{R(C_{KZ}, \mathbf{L}) : \mathbf{L} \in \hat{\mathbf{L}}\} \tag{16}$$

Rozmyte charakterystyki ryzyka

Przedziałowe charakterystyki ryzyka dają najbardziej pesymistyczne oszacowanie wartości ryzyka. Ponadto oszacowanie nie uwzględnia zależności pomiędzy różnymi wartościami $L_{i,j}$. Przykładowo błędy systematyczne zwykle, aczkolwiek nie zawsze, mają podobny wpływ na część pomiarów i nie trzeba uwzględniać całego zakresu zmienności przedziału $[L_{i,j}^-, L_{i,j}^+]$. Ponadto przedziały $[L_{i,j}^-, L_{i,j}^+]$ są czasem zbyt szerokie.

Częściowym rozwiązaniem tych problemu jest koncepcja zbiorów rozmytych [7]. W tym ujęciu zbór danych nie składa się tylko z przedziałów.

Składa się z zagnieżdżonej rodziny danych przedziałowych $\hat{\mathbf{L}}_\alpha$ (tzw. α-przekrojów).

$$\hat{\mathbf{L}}_{\alpha_0} \supseteq \hat{\mathbf{L}}_{\alpha_1} \supseteq ... \supseteq \hat{\mathbf{L}}_{\alpha_N} \tag{17}$$

Zakładamy przy tym, że przedziałowe dane $\hat{\mathbf{L}}_\alpha$ zostały utworzone na pewnym poziomie istotności α, czyli

$$P\{\hat{L}_i \cap (R \setminus \hat{L}_{i,\alpha})\} = \alpha \tag{18}$$

W tym przypadku dla każdego α-przekroju możemy obliczyć przedziałowe ryzyko.

$$\hat{R}_\alpha(C_{KZ}) = \{R(C_{KZ}, \mathbf{L}) : \mathbf{L} \in \hat{\mathbf{L}}_\alpha\} \tag{19}$$

Teraz można zbudować funkcję przynależności rozmytego ryzyka $R_F\left(C_{KZ}\right)$.

$$\mu_{R_F\left(C_{KZ}\right)}\left(x\right)=\sup\left\{\alpha : x\in \hat{R}_{\alpha}\left(C_{KZ}\right)\right\} \tag{20}$$

Zmienne losowe z rozmytymi parametrami

W dotychczasowych rozważaniach nie uwzględnialiśmy informacji na temat charakteru rozkładu prawdopodobieństwa. Jednak w pewnych zastosowaniach nie trzeba rozważać wszystkich możliwych funkcji gęstości i można się ograniczyć do pewnej klasy funkcji.
Typowym rozkładem charakterystycznym dla błędów pomiarów (oszacowania pewnych wielkości) jest rozkład normalny.
W takim przypadku znamy rozkład prawdopodobieństwa, a nie znamy jego parametrów. Parametry rozkładu normalnego określamy na podstawie wartości średniej z próby, oraz wariancji z próby.
Poprzez uwzględnienie dodatkowej informacji o charakterze rozkładu prawdopodobieństwa uzyskujemy zwykle węższe oszacowanie niż w przypadku zastosowania metody opisanej w poprzednim punkcie.
Jeśli mamy dany rozkład prawdopodobieństwa $f\left(x,\mathbf{h}\right)$ zależny od wektora parametrów $\mathbf{h}$, który charakteryzuje rozkład kosztu całkowitego inwestycji budowlanej, to ryzyko (zależne od parametru **h**) można obliczyć następująco

$$R\left(C_{KZ},\mathbf{h}\right)=\int_{C_{KZ}}^{\infty} f_{K_C}\left(x,\mathbf{h}\right)dx \tag{21}$$

Teraz można obliczyć ryzyko jako liczbę rozmytą określoną przy pomocy wzoru (21) oraz metody α -przekrojów.

$$\hat{R}_{\alpha}\left(C_{KZ}\right)=\left\{R\left(C_{KZ},\mathbf{h}\right):\mathbf{h}\in\hat{\mathbf{h}}_{\alpha}\right\} \tag{22}$$

$$\mu_{R_F\left(C_{KZ}\right)}\left(x\right)=\sup\left\{\alpha : x\in \hat{R}_{\alpha}\left(C_{KZ}\right)\right\} \tag{23}$$

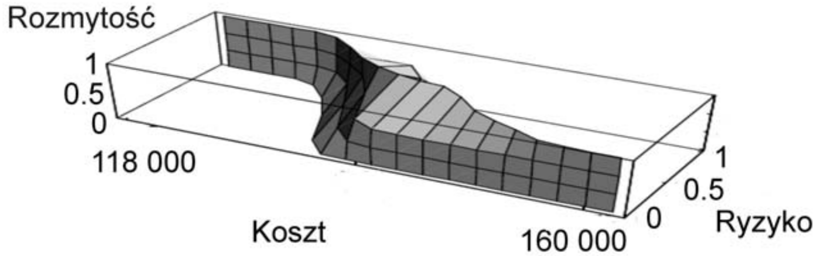
Tak określone rozmyte ryzyko $R_F\left(C_{KZ}\right)$ może zostać obliczone np. przy wykorzystaniu zmodyfikowanej metody Monte Carlo [2].
Dane potrzebne do określenia przedziałowej (rozmytej) charakterystyki parametrów rozkładu prawdopodobieństwa można zbierać przy pomocy formularzy takich jak tabela (patrz tabela 1). Można je wykorzystać do

zdefiniowania liczb rozmytych reprezentujących oceny dokładności oszacowanych ilości jednostek przedmiarowych.

Tabela 1

Tabela oceny dokładności oszacowania ilości jednostek: Nazwa procesu: koszt : osoba:									
	5	10	15	20	25	30	35	40	50
Dokładność +/- %									

Przykładowe wyniki obliczeń przedstawione są na rysunku 1. Obliczenia wykonano przy pomocy autorskiego programu [8] bazującego na zmodyfikowanej metodzie Monte Carlo [2, 6] oraz języku BPFPRAL [4].



Rys. 1

Zaprezentowana bryła obrazuje poziom ryzyka przy zakładanym poziomie kosztów maksymalnych w aspekcie poziomu zaufania do danych.

Podsumowanie

Wprowadzanie nadmiernej dokładności tam, gdzie nie mamy dostatecznej ilości danych, nie przekłada się na wiarygodność wyników. Prezentowana metoda pozwala, z jednej strony maksymalnie wykorzystać dostępne informacje, a z drugiej daje obraz niepewności danych.

Przy pomocy niepewnych informacji można w prosty sposób zdefiniować rozkłady zmiennych losowych z rozmytymi parametrami, które następnie można w efektywny sposób wykorzystać w obliczeniach rozmytego ryzyka.

Literatura

[1] Akintola, A.: Analysis of factors inuencing project cost estimating practice. Construction Management and Economics, 18(1) 2000.

[2] Bętkowski, M, Pownuk, A., Calculating Risk of Cost Using Monte Carlo Simulation with Fuzzy Parameters in Civil Engineering , NSF workshop on Reliable Engineering Computing, September 15-17, 2004, Savannah, Georgia, USA.

[3] Bętkowski M, Pownuk A., Koncepcja zmiennej losowej o parametrach rozmytych jako narzędzia do szacowania ryzyka procesów budowlanych w warunkach ograniczonej ilości danych. Konferencja naukowo-techniczna pt. BUDOWNICTWO POLSKIE W ROK PO WSTAPIENIU DO UNII EUROPEJSKIEJ Gdańsk 9 - 11 czerwca 2005

[4] Bętkowski M., Pownuk A., .: BPFPRAL ver. 1.8.2 – User manual, Gliwice, Poland, 2004.

[5] Neumaier A., Clouds, fuzzy sets and probability intervals, Reliable Computing 10: 2004.

[6] Zielińskim R. Metody Monte Carlo , WNT Warszawa 1970

[7] Zadeh L.A., 1968, Probability measure of fuzzy events. Journal of Mathematical Analysis and Applications, Vol.23, s.421-427

[8] http://pownuk.prv.pl/interval_web_applications.htm