

## **Zależność czas-koszt w przewidywaniu czasu realizacji budowy**

**Agata Czarnigowska<sup>1</sup>, Anna Sobotka<sup>2</sup>**

<sup>1</sup> *Katedra Inżynierii Procesów Budowlanych, Wydział Budownictwa i Architektury, Politechnika Lubelska, e-mail: a.czarnigowska@pollub.pl*

<sup>2</sup> *Katedra geomechaniki, Budownictwa i Geotechniki, Wydział Górnictwa i Geoinżynierii, Akademia Górniczo-Hutnicza im. Staszica w Krakowie, e-mail: sobotka@agh.edu.pl*

**Streszczenie:** W artykule podjęto próbę stworzenia modelu czasu realizacji budowy w funkcji cech charakteryzujących przedsięwzięcie, w tym kosztu. Model oparto na analizie tych cech zrealizowanych przedsięwzięć, które są znane lub możliwe do założenia we wczesnych etapach planowania, lecz bez analizy technologii i organizacji robót. Model taki mógłby być przydatny inwestorom do szacowania czasu budowy na wczesnych etapach przygotowania inwestycji, szczególnie do analiz wykonalności. Model mógłby być również podstawą do porównań czasu lub tempa robót w zależności od cech przedsięwzięcia.

**Słowa kluczowe:** model Bromilowa, regresja wieloczynnikowa, CART.

### **1. Wprowadzenie**

Oszacowania czasu i kosztu realizacji budowy są niezbędne już w najwcześniejszych fazach przedsięwzięcia – służą jako dane wejściowe do analiz wykonalności, budżetu i harmonogramu całości przedsięwzięcia. Ich wiarygodność jest bardzo istotna, dlatego ważna jest odpowiedź na pytanie: od czego zależą (lub chociaż z czym są skorelowane) koszty i czas realizacji przedsięwzięć? Odpowiedź może być oparta jedynie o doświadczenie – osobiste planisty, zgromadzone w formie baz danych lub zapisane w postaci modeli matematycznych. Artykuł porównuje modele związku czasu i kosztu przedsięwzięć: prosty model regresyjny, nieco bardziej rozbudowany wieloczynnikowy model regresyjny i model wykorzystujący technikę drzew regresyjnych. Modele zbudowano w oparciu o dane 100 polskich przedsięwzięć drogowych i przeanalizowano możliwość ich praktycznego wykorzystania.

### **2. Statystyczne modele czasu realizacji budowy**

Zapisywanie systematycznie gromadzonych doświadczeń w postaci modeli matematycznych jest przedmiotem zainteresowań wielu badaczy [1, 2, 3]. Statystyczne modele regresyjne, choć znacznie uproszczone i “uśredniające” analizowane zjawiska, ciągle uznawane są za użyteczne narzędzie, między innymi w planowaniu kosztów [4, 5], tym bardziej, że informatory cenowe i wewnętrzne, firmowe bazy danych stanowią podstawowe i powszechnie dostępne źródło danych wejściowych do takich modeli.

Koszt można uznać za uogólnioną miarę skali każdego przedsięwzięcia i jego podstawową cechę. Istnienie związku między kosztem, czasem realizacji i jakością przedsięwzięcia uznawane jest za oczywistość i opisywane w postaci trójkąta cech w większości klasycznej literatury dotyczącej zarządzania przedsięwzięciami [6], chociaż formalna postać tego związku nie jest zdefiniowana. Założenie o możliwości dokonania wiarygodnego oszacowania kosztu robót budowlanych na wczesnych etapach planowania przedsięwzięcia było podstawą wielu badań związanych z budową modelu określającego związek czasu i kosztu robót budowlanych. Pierwszy model regresyjny wiążący czas i koszt realizacji budowy przypisuje się Bromilowowi z zespołem [7]. Model ten, oparty na danych przedsięwzięć realizowanych w Australii w latach sześćdziesiątych dwudziestego wieku i wyznaczony metodą najmniejszych kwadratów, ma postać:

$$\ln L = \ln K + B \ln C, \quad (1)$$

gdzie  $L$  oznacza liczbę dni roboczych od chwili przejścia placu budowy przez wykonawcę do zakończenia robót,  $C$  – kwotę wynagrodzenia wypłaconą wykonawcy,  $K$  i  $B$  – stałe.

Ustalenia Bromilowa zostały potwierdzone przez innych badaczy na podstawie prób o różnej liczności i dotyczących rozmaitych przedsięwzięć budowlanych [7, 8, 9, 10]: funkcja (1) okazała się najlepiej dopasowana niż inne brane pod uwagę, choć współczynniki determinacji podawane przez autorów zwykle nie były wysokie (w przypadku dużych prób poniżej 0,75). W odniesieniu do przedsięwzięć tego samego rodzaju i realizowanych w tym samym kraju zauważono znaczne roczne wahania wartości stałych  $B$  i  $K$ , jednak bez wyraźnego trendu [11].

Statystyczna istotność związku czas-koszt stanowiła zachętę do poszukiwań modeli wieloczynnikowych – ujmujących zależność czasu od innych niż koszt cech przedsięwzięcia. Analizowano przede wszystkim przedsięwzięcia związane ze wznoszeniem budynków, a jedynie w nielicznych pracach analizowano budowę [7, 9, 12]. Z nielicznymi wyjątkami [13, 14, 15], koszt był traktowany jako podstawowa zmienna niezależna. W modelach najczęściej występował związek liniowy między logarytmem czasu i logarytmem kosztu, chociaż niektórzy autorzy proponowali inne przekształcenia [15, 16, 17]. Autorzy dochodzili również do rozbieżnych wniosków co do zestawu zmiennych statystycznie powiązanych z czasem realizacji. Na przykład, w niektórych pracach funkcja obiektu, jego wielkość i sektor inwestora okazywały się istotnie związane z czasem realizacji, w innych wykluczano je jako nieistotne. Wyjściowe zbiory cech również były rozbieżne, tym bardziej że modele miały służyć różnym celom: przewidywaniu czasu realizacji, poszukiwaniu cech skorelowanych z czasem, lub ocenie przewidywalności kosztu. Niektórzy badacze skupiali się na cechach związanych ze sposobem zarządzania przedsięwzięciem, inni wybierali czynniki techniczne, znane zarówno na wczesnych etapach planowania, jak i możliwe do określenia dopiero po zakończeniu prac. Mimo że modele na ogół pomyślnie przechodziły weryfikację okazując się wiarygodne i istotne, ich interpretacja w kategoriach związków przyczynowo-skutkowych w zasadzie nie jest możliwa. Na przykład, w modelu Walkera [18] większa efektywność komunikacji między projektantami a wykonawcą oraz większy stopień wykorzystania narzędzi informatycznych związane były z przedsięwzięciami o długim czasie realizacji, a budowy mające w swym zakresie pełne wykonanie trwały krócej, niż budowy o tej samej wartości, lecz obejmujące tylko stan surowy. W przytoczonych pracach rzadko poświęcano uwagę miarom jakości modeli, a informacje o przedziałach ufności i predykcji dla oszacowań dokonanych z ich użyciem można znaleźć tylko w [15, 16, 17].

### 3. Modele czasu realizacji przedsięwzięć drogowych

#### 3.1. Przedmiot analizy

W niniejszej pracy jako budowę traktuje się realizację robót zleconych w ramach jednego postępowania o zamówienie publiczne i jednej umowy, uzupełnionej ewentualnie o umowy o roboty dodatkowe. Populacją, do której odnoszą się analizy przedstawione w tej pracy, są budowy drogowe: prowadzone na drogach wojewódzkich oraz krajowych, z wykluczeniem autostrad, których investorem jest instytucja publiczna – odpowiednio jednostka samorządowa – Zarząd Dróg Wojewódzkich, albo jednostka administracji rządowej – oddział lub centrala Generalnej Dyrekcji Dróg Krajowych i Autostrad, w województwach: lubelskim, podkarpackim i małopolskim, zrealizowane w latach 2003-2009, mające charakter zadań inwestycyjnych – budowy lub przebudowy.

W trakcie analizy rynku inwestycji drogowych okazało się, że liczba budów nowych obiektów zakończonych między 2003 a 2009 rokiem na terenie objętym planowanym badaniem jest niewielka, a większość przedsięwzięć inwestycyjnych polegała na rozbudowie lub przebudowie istniejących obiektów. Włączono je do zakresu badań ze świadomością, że zbudowany w oparciu o nie model będzie miał bardzo ogólny charakter. Próba, na podstawie której zbudowano modele, obejmuje 100 budów, a 7 dodatkowych wykorzystano jako próbę służącą do oceny jakości predykcji.

### 3.2. Zakres i metoda analizy

Etapy badań obejmowały kolejno: przegląd literatury, wywiad z przedstawicielami inwestorów w celu ustalenia stosowanych przez nich metod planowania czasu i kosztu budowy (poza zakresem tego artykułu), zgromadzenie danych wejściowych poprzez analizę źródłowej dokumentacji przedsięwzięć udostępnionej przez inwestorów (nie istniały gotowe bazy danych), wstępną selekcję potencjalnych zmiennych niezależnych modelu z wykorzystaniem techniki drzew regresyjnych, oraz budowę i porównanie następujących modeli:

- prostego modelu regresyjnego wiążącego funkcję rzeczywistego czasu realizacji budowy ( $L$ ) z funkcją rzeczywistego kosztu budowy ( $C$ ) o postaci  $f(L) = b_0 + b_1 f(C)$  – odpowiednika modelu Bromilowa, o parametrach wyznaczonych metodą najmniejszych kwadratów;
- drzewa regresyjnego CART zbudowanego na podstawie 25 potencjalnych zmiennych niezależnych, w tym kosztu budowy;
- wieloczynnikowego modelu regresji o parametrach wyznaczonych metodą najmniejszych kwadratów, wiążącego funkcję czasu realizacji ( $L$ ) z funkcjami zmiennych wybranych jako najistotniejsze w trakcie tworzenia modelu CART, o postaci  $f(L) = b_0 + b_1 f(x_1) + b_2 f(x_2) + \dots + b_n f(x_n)$ .

Obliczenia wykonano w programie Statistica 8.0.

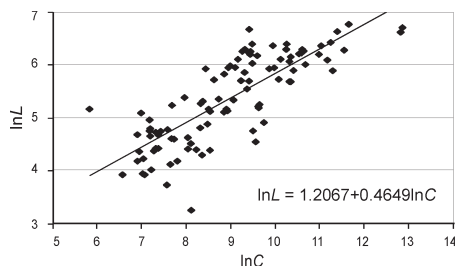
### 3.3. Prosty model regresyjny

Analiza wykresów rozrzutu czas-koszt (rysunek 1a) i próby z kilkoma postaciami zależności między czasem a kosztem potwierdziły, że model o postaci (1) jest najlepiej dopasowany do próby, spełnia założenia metody najmniejszych kwadratów, jest istotny i o istotnych parametrach. Współczynnik determinacji (0,64) i skorygowany współczynnik determinacji (0,63) nie są jednak wysokie. Równanie regresji ma postać:

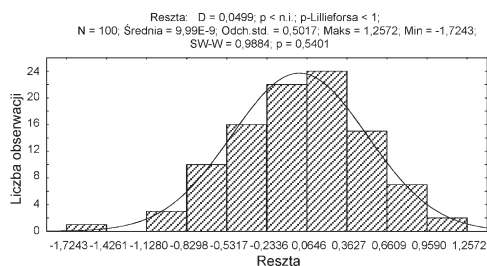
$$\ln L = 1,2067 + 0,4649 \ln C \pm 0,5042. \quad (2)$$

(0,3207)      (0,0352)

Średnia reszt jest równa 0, a wykres normalności reszt wskazuje, że rozkład reszt jest zbliżony do normalnego. Potwierdzają to wartości testów umieszczone na histogramie reszt (rysunek 1b): w teście Shapiro-Wilka wartość statystyki testowej  $W=0,9884$  nie jest istotna (poziom  $p=0,5401$  przekracza zakładany poziom istotności  $\alpha=0,05$ , więc nie ma podstaw do odrzucenia hipotezy o normalności rozkładu reszt. W teście Kolmogorowa-Smirnowa (Lillieforsa): wartość statystyki testowej  $D$  nie jest istotna, więc nie ma podstaw do odrzucenia hipotezy o normalności rozkładu. Zmiany rozrzutu reszt wraz ze wzrostem wartości przewidywanych nie były duże, a test Lagrange'a nie odrzucił hipotezy zerowej o równości wariancji w równolicznych podgrupach uporządkowanych według rosnącej wartości przewidywanej (poziom  $p=0,297$ ), mimo występowania obserwacji wyraźnie odstających. Usunięcie obserwacji odstających prowadzi do poprawy współczynnika determinacji i zmniejszenia błędu standardowego przy spełnieniu założeń metody najmniejszych kwadratów. Zdecydowano jednak na pozostawienie wszystkich obserwacji.



Rys. 1a. Model 1 – jednoczynnikowy model regresji; wykres rozrzutu z prostą regresji,

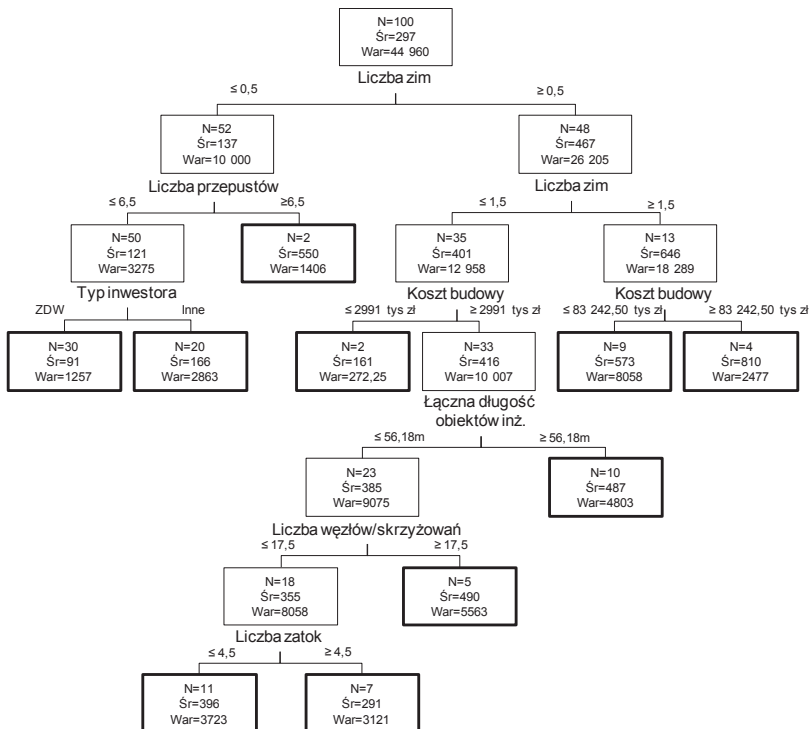


Rys. 1b. Model 1 – histogram reszt z wynikami testów normalności

### 3.4. Model CART

Metoda drzew regresyjnych jest techniką stopniowego podziału obserwacji (tu – przedsięwzięć drogowych) na rozłączne podzbiory (węzły) jednorodne ze względu na określoną cechę – zmienną zależną (w badanym przypadku – liczbę dni trwania budowy  $L$ ), w celu określenia wpływu zmiennych niezależnych na zmienną zależną. Algorytmy budowy drzew regresyjnych są rekurencyjne, oparte na lokalnej optymalizacji każdego kroku podziału. Metoda nie wymaga określenia rozkładów analizowanych zmiennych ani postaci zależności funkcyjnych między nimi. Można w niej operować zmiennymi ilościowymi i jakościowymi bez konieczności ich przekształceń. Drzewo nie opisuje charakteru zależności między zmiennymi, ale może służyć do predykcji wartości zmiennej zależnej lub do określenia, które ze zmiennych niezależnych są najsilniej związane ze zmienną zależną.

Przedsięwzięcia wchodzące w skład próby miały zróżnicowany charakter. Wybór zestawu wspólnych cech (potencjalnych zmiennych objaśniających) charakteryzujących wszystkie przedsięwzięcia był przez to utrudniony. Ponieważ model miał służyć predykcji czasu trwania budowy, a nie tylko opisowi zależności, cechy te miały być znane lub możliwe do założenia we wczesnych fazach przygotowania przedsięwzięcia. Zebrano dane dotyczące 25 cech: ilościowych ciągłych (np. koszt budowy, długość odcinka drogi), ilościowych dyskretnych (np. liczba skrzyżowań), oraz jakościowych (np. typ inwestora). W oparciu o metodykę budowy drzew binarnych CART, przyjmując procedurę Breimana do wyboru najlepszego drzewa [19] otrzymano model przedstawiony na rysunku 2.



Rys. 2. Struktura najlepszego drzewa regresyjnego (Model 2):  $N$  – liczba obserwacji w podzbiore,  $\bar{S}r$  – średni czas realizacji budowy w podzbiore (w dniach),  $War$  – wariancja czasu realizacji budowy w podzbiore.

Model wykorzystuje 7 predyktorów: zakładaną liczbę zim w trakcie budowy, koszt budowy, liczbę przepustów, typ inwestora, całkowitą długość obiektów inżynierskich, liczbę węzłów, liczbę zatok. Skorygowany współczynnik determinacji modelu [19] jest wysoki (wynosi 0,924). Model można więc uznać za dobrze dopasowany do próby. Biorąc jednak pod uwagę, że liczba obserwacji jest stosunkowo mała, i są one bardzo zróżnicowane, nie można tego automatycznie uznać za zaletę modelu, którego przeznaczeniem jest

predykcja. Gdyby używać go do przewidywania czasu trwania budowy, w zależności od cech przypisywałby przedsięwzięciom jeden z następujących dziesięciu czasów realizacji: 91, 161, 166, 291, 396, 487, 490, 550, 573, oraz 810 dni.

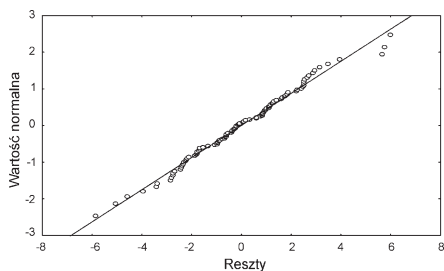
### 3.5. Model regresji wielorakiej

Liczba potencjalnych predyktorów czasu trwania budowy jest duża (25 cech). Poszukiwanie postaci wieloczynnikowego równania regresji jest przez to utrudnione, bo nie są znane najlepiej dopasowane postaci związków funkcyjnych – jedynie dla związku czasu i kosztu ustalono, że właściwe jest przekształcenie logarytmiczne. Może to ulec zmianie przy budowie modelu wieloczynnikowego. Wiele zmiennych ma charakter jakościowy, a ich zamiana na zmienne binarne powoduje przyrost liczby potencjalnych zmiennych niezależnych do umieszczenia w równaniu regresji do czterdziestu. Liczebność próby jest więc zbyt mała ( $n=100$ ), by w budowie modeli móc polegać na algorytmach regresji krokowej [20]. Dlatego zmienne poddano wstępnej selekcji za pomocą nieparametrycznej metody drzew regresyjnych typu CART [19] ograniczając ich liczbę do dziewięciu.

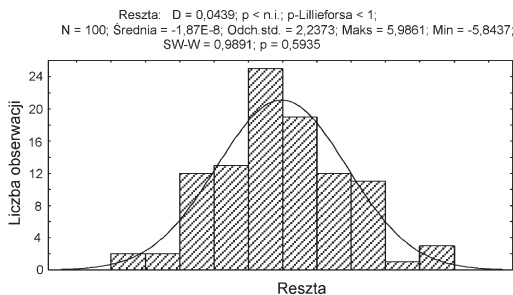
Prawdopodobne postaci modelu określono na podstawie analizy wykresów rozrzutu. Wypcyfikowano kilka modeli. Po określeniu ich parametrów przeprowadzono analizę istotności modelu (test  $F$  Fishera-Snedecora) i analizę istotności poszczególnych parametrów (test  $t$ -Studenta), eliminując krokowo składniki nieistotne. Sprawdzano założenia metody najmniejszych kwadratów [20]: graficznie na podstawie wykresów oraz testów normalności i stałości wariancji reszt (test Lagrange’a [20]). Zbadano również wpływ obserwacji odstających na parametry modelu i założenia metody najmniejszych kwadratów. Następnie dokonano porównania modeli w celu wyboru najlepszego. Modele opisujące tę samą postać zmiennej zależnej porównano na podstawie skorygowanego współczynnika determinacji, błędu standardowego, średniego procentowego błędu bezwzględnego i statystyki  $C_p$  Mallowa [20]. Za model najlepszy spośród analizowanych uznano model wykorzystujący cztery predyktory: koszt budowy  $C$ , sumę długości obiektów inżynierskich  $x_2$ , liczbę zim  $x_3$  i liczbę obiektów inżynierskich  $x_4$ , postaci:

$$\sqrt{L} = -4,4289 + 1,8925 \ln C + 0,5598 \ln x_2 + 4,4436 x_3 - 0,2770 x_4 \pm 2,2839 \quad (3)$$

(2,0498)
(0,2512)
(0,1477)
(0,3990)
(0,1101)



Rys. 3a. Model 3, wykres normalności reszt



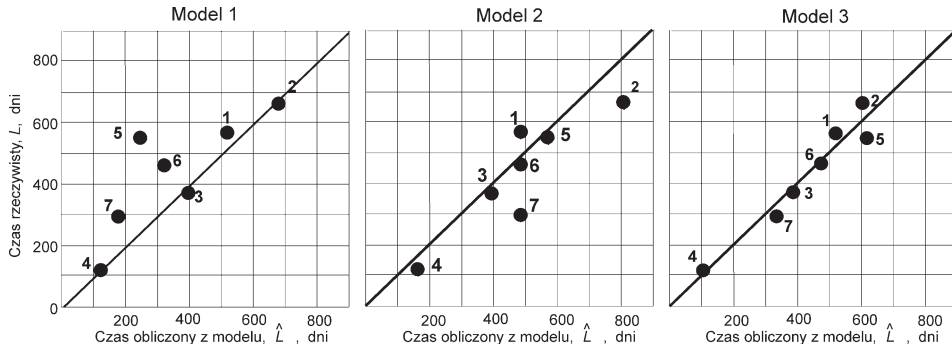
Rys. 3b. Model 3, histogram reszt

Rysunek 3 przedstawia wykresy związane z analizą reszt modelu. Nad histogramem (Rys. 3b) umieszczono wyniki testów normalności reszt.

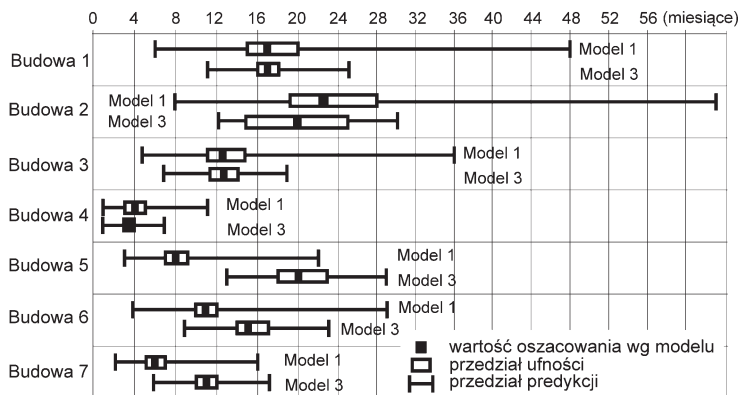
### 3.6. Porównanie modeli

Ponieważ trzy analizowane modele wykorzystują jako zmienne zależne różne funkcje czasu budowy, nie można ich porównać na podstawie podanych wcześniej skorygowanych współczynników determinacji, błędu standardowego modelu  $SEE$  ani średniego bezwzględnego błędu procentowego  $MAPE$ . Na bazie modeli ma być określany czas realizacji budowy  $L$ , więc do porównań obliczono błędy oszacowań wyrażonych w dniach, jako różnicę między obserwowanym czasem realizacji a czasem obliczonym na podstawie modelu. Można wtedy bezpośrednio porównać średnie procentowe błędy bezwzględne tych oszacowań. Największy średni procentowy błąd bezwzględny oszacowań w dniach występuje w Modelu 1 (45%), nieparametryczny model 2 ma najmniejszy błąd (23%),

jednak nie jest on o wiele mniejszy, niż w przypadku Modelu 3 (28%). Na tle błędów modeli opartych na zróżnicowanych próbach o podobnej wielkości [14, 15] rzędu 20-50%, błędy bezwzględne otrzymanych modeli są porównywalnej skali. Do porównania jakości przewidywań obliczono przewidywane czasy realizacji siedmiu budów z próby testowej – osobnej, lecz o cechach mieszczących się w zakresie zmienności cech przedsięwzięć użytych do budowy modeli – i obliczono błędy predykcji. Model 3 charakteryzuje się najmniejszymi rozbieżnościami między wartościami oczekiwanymi prognoz i faktycznym czasem realizacji budowy (rysunek 4).



Rys. 4. Wykresy rozrzutu wartości obserwowanych względem przewidywanych w próbie testowej - porównanie modeli



Rys. 5. Porównanie skali przedziałów ufności i predykcji czasu realizacji budowy w Modelu 1 i 3

Wniosek ten oparto jednak na małej próbie, a dobre dopasowanie testowych obserwacji do Modelu 3 może być przypadkowe. Dlatego dokonano porównania przedziałów ufności i predykcji – tylko Modelu 1 i Modelu 3 (parametrycznych), ponieważ nie ma podstaw do tego typu obliczeń w przypadku drzewa regresyjnego. Zakresy przedziałów przy poziomie ufności 95% i przeliczone na miesiące przedstawiono graficznie na rys. 5.

Można zauważyć, że przedziały predykcji Modelu 1 są szerokie. Na przykład, oczekiwany czas realizacji Budowy 2 można oszacować na 23 miesiące, przy czym przedział ufności dla tego przeciętnego oszacowania to od 19 do 28 miesięcy, a konkretne przedsięwzięcie o takich cechach mogłoby trwać od 8 do 63 miesięcy. Model 1 jest więc za mało dokładny do praktycznego użytku. Model 3, mimo znacznie węższych przedziałów ufności i predykcji również należy uznać za „mało dokładny”: szacując według niego czas realizacji tej samej Budowy 2 otrzymano by czas oczekiwany 20 miesięcy, z przedziałem ufności dla regresji od 15 do 25 miesięcy, oraz przedziałem predykcji od 12 do 30 miesięcy. Ze względu na postać zmiennej zależnej, oszacowania przeliczone na dni są tym mniej „dokładne” (tzn. mają tym szersze przedziały ufności i predykcji), im dłuższa jest budowa.

#### 4. Podsumowanie i wnioski

Próba przedsięwzięć drogowych analizowana w artykule była stosunkowo niewielka przy swym zróżnicowaniu, jednak ustalono zależności między czasem realizacji budowy a innymi cechami przedsięwzięć, które można uznać za istotne statystycznie.

Potwierdzono uniwersalny charakter prostego modelu regresyjnego wiążącego czas z kosztem realizacji przedsięwzięcia. Model ten jest jednak zbyt uproszczony, by znaleźć zastosowanie do szacowania czasu trwania budowy – jego błędy są duże, a przedziały ufności i predykcji niepraktycznie szerokie.

Model nieparametryczny (drzewo regresyjne typu CART dobierane według procedury Breimana) mimo dobrego dopasowania do próby trudno uznać za wiarygodne narzędzie szacowania czasu realizacji, ponieważ brakuje miar zaufania do wyników. Z uwagi na niewielką próbę testową (7 obserwacji) nie udało się dokonać wystarczającej walidacji, jednak w trakcie jego tworzenia ustalono, które z 25 analizowanych cech przedsięwzięcia mogą być potencjalnie najsilniej skorelowane z czasem realizacji budowy. Cztery cechy wskazywane w analizie drzew jako „najważniejsze”: koszt budowy, liczba zim w okresie budowy, łączna długość obiektów inżynierskich i liczba obiektów inżynierskich okazały się istotne w klasycznej analizie regresji wieloczynnikowej i wystąpiły w modelu uznanym za najlepszy. Co ciekawe, model CART wykorzystuje inny zestaw predyktorów, niż regresja wieloraka: zakładana liczba zim, koszt budowy i łączna długość obiektów inżynierskich są dla obu modeli wspólne, a zamiast liczby obiektów inżynierskich drzewo CART jako predyktory wykorzystuje liczbę przepustów, typ inwestora, liczbę węzłów i liczbę zatok.

Najlepiej dopasowany spośród branych pod uwagę modeli regresji wielorakiej o parametrach wyznaczanych metodą najmniejszych kwadratów ma ogólną postać:

$$\sqrt{L} = b_0 + b_1 \ln C + b_i \ln x_i + b_j x_j, \quad (4)$$

gdzie  $x_i$  oznacza zmienną ilościową ciągłą,  $x_j$  – zmienną dyskretną, a  $b_i$  – parametry. Postać ta jest inna niż podawane najczęściej w literaturze przedmiotu – wykorzystujące jako zmienną zależną logarytm czasu realizacji budowy. Zbliżone równanie regresji można znaleźć w [16] i komercyjnym „kalkulatorze czasu realizacji” [17], opartych na dużych próbach (tysiące obserwacji), lecz dotyczących budynków. Co ciekawe, skala przedziałów ufności i predykcji tamtych modeli jest porównywalna z bardzo szerokimi przecież przedziałami uzyskanymi w analizie przedstawionej w artykule. Niewątpliwie świadczy to o fakcie, że model parametryczny czasu trwania budowy oparty na wąskim zestawie cech jest bardzo zgrubny – nie uwzględnia mnóstwa czynników, które istotnie wpływają na czas, lecz mają indywidualny charakter. Można się spodziewać, że zwiększenie liczby obserwacji użytych do budowy modelu i poszerzenie zestawu predyktorów poprawiłoby dokładność modeli parametrycznych, oraz umożliwiło budowę wiarygodnych modeli nieparametrycznych, na przykład wykorzystujących sieci neuronowe lub drzewa regresyjne.

*Praca była finansowana ze środków statutowych Ministerstwa Nauki i Szkolnictwa Wyższego (S/63/2013).*

*Autorki dziękują panu Joe Martinowi, Dyrektorowi BCIS, za pomoc i udzielenie informacji, oraz krakowskim, rzeszowskim i lubelskim oddziałom GDDKiA oraz Podkarpackiemu, Małopolskiemu i Lubelskiemu Zarządowi Dróg Wojewódzkich za udostępnienie dokumentacji przedsięwzięć.*

#### Literatura

1. Lai Y-T., Wang W-C., Wang H-H. AHP- and simulation-based budget determination procedure for public building construction projects. *Automation in Construction* 17 (2008) 623-632.
2. Lee J-R, Hsueh S-L, Tseng H-P. Utilizing data mining to discover knowledge in construction enterprise performance records. *Journal of Civil Engineering and Management*, 14 (2008) 79-84.
3. Kaplinski O. *Modelling of Construction Processes: A Managerial Approach*. KILiW PAN, IPPT, Studia z Zakresu Inżynierii Nr. 43, Warszawa, 1997.
4. Chou, J.-S., Tseng, H.-S., 2011. Establishing expert system for prediction based on the project-oriented data warehouse. *Expert Systems with Applications*, 38 (2011) 640-651.

5. Jin RZ, Cho KM, Hyu CT, Son MJ: MRA-based revised CBR model for cost prediction in the early stage of construction projects. *Expert Systems with Applications*, 39 (2012), 5214-5222
6. Project management Body of Knowledge (PMBOK® Guide). Fourth Edition. Wydanie polskie. Management Training & Development Center, Warszawa, 2009.
7. Kaka A., Price A.D.F. Relationship between value and duration of construction projects. *Construction Management and Economics* 9 (1991) 383-400.
8. Chan A.P.C. Time-cost relationship of public sector projects in Malaysia. *International Journal of Project Management*, 19, 223-229.
9. Yousef G., Baccarini D. Developing a cost-time model for estimating construction durations of sewerage projects in Perth. *Journal of Construction Research* 2 (2001) 213-220.
10. Ogunsemi D.R., Jagboro G.O. Time-cost model for building projects in Nigeria. *Construction Management and Economics* 24 (2006) 253-258.
11. Skitmore R.M., Ng S.T. Note. Australian project time-cost analysis: statistical analysis of intertemporal trends. *Construction Management and Economics* 19 (2001) 455-458.
12. Irfan, M., Khursid, M.B., Anastasopoulos, P., Labi, S., Moavenzadeh, F. Planning-stage estimation of highway project duration on the basis of anticipated project cost, project type, and contract type. *International Journal of Project Management* 29 (1) (2011) 78-92.
13. Skitmore R.M., Ng S.T. Forecast models of actual construction duration and cost. *Building and Environment* 38(8) (2003) 1075-1083.
14. Love P. E. D., Tse R. Y. C., Edwards D.J. Time–Cost Relationships in Australian Building Construction Projects. *Journal of Construction Engineering and Management*, ASCE, 131(2) 2005 187-194.
15. Stoy Ch., Polalis S. Early estimation of building construction speed in Germany. *International Journal of Project Management* 25 (2007) 283-289.
16. Martin J., Burrows T.K., Pegg I. Predicting Construction Duration of Building Projects. Shaping the Change XXIII FIG Congress, Munich, Germany, October 2006, 8-13.
17. BCIS Construction Duration Calculator <http://www.rics.org>, dostęp w styczniu 2013
18. Walker D.H.T. An investigation into construction time performance. *Construction Management and Economics* 13 (1995) 263-274.
19. Gatnar E. Nieparametryczna metoda dyskryminacji i regresji. PWN, Warszawa, 2001.
20. Stanisław A. Przystępny kurs statystyki z zastosowaniem STATISTICA PL na przykładach z medycyny. Tom 2. Modele liniowe i nieliniowe. Statsoft Polska, Kraków, 2007.

## Time-Cost Relationship for Predicting Construction Duration

Agata Czarnigowska<sup>1</sup>, Anna Sobotka<sup>2</sup>

<sup>1</sup> *Department of Construction Management, Faculty of Civil Engineering and Architecture, Lublin University of Technology, e-mail: a.czarnigowska@pollub.pl*

<sup>2</sup> *Department of Geomechanics, Civil Engineering and Geotechnics, Faculty of Mining and Geoengineering, AGH University of Science and Technology, Kraków, e-mail: sobotka@agh.edu.pl*

**Abstract:** The paper aims at creating a model of road construction duration in the function of project qualities including the construction cost. The qualities considered are likely to be defined or possible to be estimated at early stages of project planning, giving no consideration to construction method or organisation of works. Potentially, the model might be applied by construction clients in their feasibility studies. It could also be used for comparing construction duration or construction rate with respect to project qualities.

**Keywords:** Bromilow's model, multiple regression, CART.