

Ziemowit Suligowski

MODEL PROGNOZY ILOŚCIOWEJ DLA POTRZEB KANALIZACJI W OKRESIE PERSPEKTYWICZNYM

Oddziaływanie różnorodnych czynników powoduje, że w miarę upływu czasu rośnie znaczenie prognoz ilościowych w zakresie gospodarki wodno-ściekowej, wzrasta przy tym zarówno ich zakres, jak też zasięg. Wiarygodność prognoz jest szczególnie istotna dla systemów kanalizacyjnych, charakteryzujących się specyficzną „sztywnością” tradycyjnych rozwiązań [9]. W tym przypadku konieczne jest stosowanie odpowiedniego wyprzedzenia potrzeb w czasie, a w szczególnych sytuacjach kierowanie się wręcz całkowitą chłonnością danego terenu [1]. Dla mniej uciążliwych warunków horyzont prognoz dla kanalizacji powinien być głębszy od stosowanego dla potrzeb zaopatrzenia w wodę, zdarza się, że zaleca się nawet perspektywę 100 lat [7].

Założenia metodyczne

Stosowanie metod bilansowych do budowy prognozy ilościowej wymaga dysponowania sporządzonymi wcześniej planami zagospodarowania przestrzennego. Praktycznie plany te rzadko kiedy wykraczają poza perspektywę 20 lat, a więc bardzo krótką, z punktu widzenia budowy kanalizacji w dużym mieście lub dużego systemu grupowego. Pozostaje przy tym kwestia nierozstrzygniętą wiarygodność planów i ich realność [8] tym bardziej, że szereg rozwiązań nierzadko mieści się w kategoriach tzw. pobożnych życzeń, a wpływ na kształtowanie się procesów urbanistycznych jest bardzo ograniczony i wymaga dużego wysiłku [2].

W proponowanym niżej modelu wykorzystano charakterystyczne cechy urbanizacji [5] takie jak ciągłość (lub łwasiciągłość) zmian w wieloleciu oraz przebieg zgodny z zasadą inercji, to jest w warunkach, gdy dla wystąpienia istotnych zmian systematycznych w procesie konieczne jest upłygnięcie dłuższego czasu. Powyższe założenia są zgodne z ustaleniami szeregu prac [2, 5, 8]. Warunki te powodują konieczność prognozowania w oparciu o analizy szeregów czasowych, tworzonych przez badane wskaźniki uporządkowane w kolejności występowania.

Ogólną miarę zmian stanowi klasyczne równanie trendu empirycznego:

$$\hat{y} = a + b \cdot t \quad (1)$$

gdzie:

t — czas,

a, b — parametry równania.

Parametry równania (1) określane są zazwyczaj metodą najmniejszych kwadratów. Zagadnienie to, jak również weryfikacja ich istotności, przedstawione są w licznych pracach, m. in. [4, 6]. Trzeba przy tym zwrócić uwagę na konieczność jednoznacznej identyfikacji postaci równania (1) tym bardziej, że w licznych przypadkach przebiegi równania liniowego, potęgowego, czy też wykładniczego są w niektórych przedziałach bardzo podobne. W tym celu konieczne jest uwzględnienie dostatecznie obszernego materiału empirycznego — szereg wyjściowy powinien liczyć nie mniej niż 11 wyrazów. Kryterium decydującym jest minimalizacja wartości współczynnika zbieżności q^2 :

$$q^2 = \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (2)$$

gdzie:

y_i — wartość mierzona,

$\bar{y}$ — średnia arytmetyczna wartości mierzonych,

$\hat{y}_i$ — i -ta wartość obliczona wg (1),

$q^2 \in \langle 0, 1 \rangle$,

lub też maksymalizacja wartości współczynnika determinacji:

$$R^2 = 1 - q^2 \quad (3)$$

Spełnienie powyższego warunku nie wyklucza jednak w poszczególnych przypadkach konieczności wariantowania prognoz, czy też wprowadzenia dodatkowych kryteriów.

Założenia modelowe

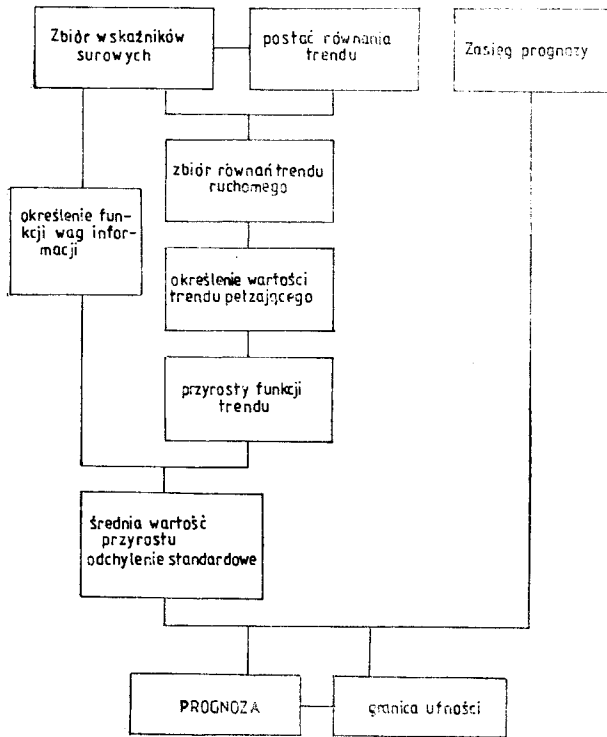
Decydujące znaczenie, z punktu widzenia struktury modelu prognostycznego, ma dopuszczalność założeń w zakresie wag informacji pochodzących z przeszłości. W warunkach regularności zmian występujących w czasie oraz zgodności punktów empirycznych i obliczonych wg (1) można dopuścić prognozowanie z równania trendu:

$$y_{t+\tau}^p = a + b(t + \tau) \quad (4)$$

gdzie:

τ — horyzont prognozy.

Na ogół jednak na prognozowane zjawisko oddziałują różne czynniki (w tym losowe), a w celu ich wyeliminowania materiał empiryczny (rys. 1) przekształca się do postaci tzw. trendu pełzającego, stosując wygładzenie trendem ruchomym. Ten ostatni obliczany jest dla 5 ko-



Rys. 1. Ogólny schemat budowy modelu prognozy.

lejszych okresów: 1, 2, 3, 4, 5; 2, 3, 4, 5, 6; n-4, n-3, ..., n (tab. 1). Podstawę dalszych operacji stanowią analizy przyrostów wartości trendu pełzającego:

$$W_{t+1} = \hat{Y}_{t+1} - \hat{Y}_t \quad (5)$$

$$\hat{Y}_t = [\hat{y}_1, \hat{y}_2, \dots, \hat{y}_n] \quad (6)$$

stanowiącego relatywnie najbardziej dokładne przybliżenie rezultatów obliczeń do wielkości mierzonych. Prognozę w ogólnej postaci opisują związki:

$$Y_{t+1}^p = Y_t^p + W_{t+1}; \quad Y_n^p = \hat{Y}_n \quad (7)$$

czyli zakłada się, że wartość trendu pełzającego bezpośrednio poprzedzająca okres prognozy, odpowiada pierwszej wartości prognozowanej.

Prognozowanie może być realizowane dwojako, przy założeniu stałych lub zmiennych w czasie wag informacji wyjściowych. W pierwszym przypadku podstawą jest średnia wartość przyrostu:

$$\bar{W} = \frac{1}{n-1} \sum_{t=1}^{n-1} W_{t+1} \quad (8)$$

Wartość wariancji wynosi:

$$S^2_1 = \frac{1}{n-1} \sum_{t=1}^{n-1} (W_{t+1} - \bar{W})^2 \quad (9)$$

Nie zawsze jednak uzasadnione jest założenie braku uzależnienia wagi informacji od okresu, z którego ona pochodzi. W tym przypadku średnia wartość przyrostu wynosi:

$$\bar{\bar{W}} = \sum_{t=1}^{n-1} C_{t+1}^n \cdot W_{t+1} \quad (10)$$

gdzie C_{t+1}^n jest współczynnikiem wag harmonicznych:

$$C_{t+1}^n > 0, \text{ dla } t=1, 2, 3, \dots, n-1$$

$$\sum_{t=1}^{n-1} C_{t+1}^n = 1 \quad (11)$$

Wprowadzenie tego współczynnika opiera się na założeniu, że informacje pochodzące z okresów najbardziej oddalonych od prognozy mają mniejszą wagę, a przyrosty wag są odwrotnie proporcjonalne do czasu dzielącego daną informację od najświeższej. Najdawniejszej informacji przypisuje się wagę:

$$m_2 = \frac{1}{n-1} \quad (12)$$

a następnej:

$$m_3 = m_2 + \frac{1}{n-2} = \frac{1}{n-1} + \frac{1}{n-2} \quad (13)$$

ZESTAWIENIE WARTOŚCI TRENDU

Tabela 1

t	Wartości trendu ruchomego numer										Wartości trendu pełzającego	
	1	2	3	4	5	...	n-8	n-7	n-6	n-5	n-4	
1	$\hat{y}_{1,1}$											$\hat{y}_{1,1}$
2	$\hat{y}_{1,2}$	$\hat{y}_{2,1}$										$(\hat{y}_{1,2} + \hat{y}_{2,1}) : 2$
3	$\hat{y}_{1,3}$	$\hat{y}_{2,2}$	$\hat{y}_{3,1}$									$(\hat{y}_{1,3} + \hat{y}_{2,2} + \hat{y}_{3,1}) : 3$
4	$\hat{y}_{1,4}$	$\hat{y}_{2,3}$	$\hat{y}_{3,2}$	$\hat{y}_{4,1}$								$(\hat{y}_{1,4} + \hat{y}_{2,3} + \hat{y}_{3,2} + \hat{y}_{4,1}) : 4$
5	$\hat{y}_{1,5}$	$\hat{y}_{2,4}$	$\hat{y}_{3,3}$	$\hat{y}_{4,2}$	$\hat{y}_{5,1}$							$(\hat{y}_{1,5} + \hat{y}_{2,4} + \hat{y}_{3,3} + \hat{y}_{4,2} + \hat{y}_{5,1}) : 5$
⋮												
n-4							$\hat{y}_{n-8,5}$	$\hat{y}_{n-7,4}$	$\hat{y}_{n-6,3}$	$\hat{y}_{n-5,2}$	$\hat{y}_{n-4,1}$	$(\hat{y}_{n-8,5} + \hat{y}_{n-7,4} + \hat{y}_{n-6,3} + \hat{y}_{n-5,2} + \hat{y}_{n-4,1}) : 5$
n-3								$\hat{y}_{n-7,5}$	$\hat{y}_{n-6,4}$	$\hat{y}_{n-5,3}$	$\hat{y}_{n-4,2}$	$(\hat{y}_{n-7,5} + \hat{y}_{n-6,4} + \hat{y}_{n-5,3} + \hat{y}_{n-4,2}) : 4$
n-2									$\hat{y}_{n-6,5}$	$\hat{y}_{n-5,4}$	$\hat{y}_{n-4,3}$	$(\hat{y}_{n-6,5} + \hat{y}_{n-5,4} + \hat{y}_{n-4,3}) : 3$
n-1										$\hat{y}_{n-5,5}$	$\hat{y}_{n-4,4}$	$(\hat{y}_{n-5,5} + \hat{y}_{n-4,4}) : 2$
n											$\hat{y}_{n-4,5}$	$\hat{y}_{n-4,5}$

Ogólnie:

$$m_{t+1} = m_t + \frac{1}{n-t} \quad (14)$$

gdzie $t=1, 2, 3, \dots, n-1$

$$\sum_{t=1}^{n-1} m_{t+1} = n-1 \quad (15)$$

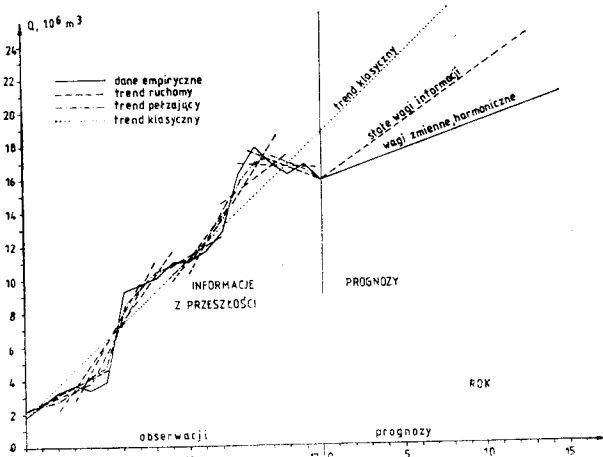
Wariancję określa zależność:

$$S^2_2 = \sum_{t=1}^{n-1} C^n_{t+1} \cdot (W_{t+1} - \bar{W})^2 \quad (16)$$

Założenie zróżnicowania wag jest słuszne, jeżeli:

$$\bar{W} > W; \quad S_2 < S_1 \quad (17)$$

Założenia w tym zakresie wpływają bardzo istotnie na rezultat prognozowania (rys. 2).



Rys. 2. Przykłady budowy prognoz wg proponowanego sposobu, przy założeniu liniowej postaci równania linii trendu empirycznego.

Równania prognozy

Spśród licznych zależności [4] możliwych do prognozowania, przy zróżnicowanych sytuacjach [3] dla analizowanego przypadku, praktyczne zastosowanie mogą znaleźć trzy związki:

$$\text{— liniowy: } \hat{y} = a + b \cdot t \quad (18)$$

$$\text{— potęgowy: } \hat{y} = a \cdot (t)^b \quad (19)$$

$$\text{— wykładniczy: } \hat{y} = a \cdot (b)^t \quad (20)$$

spomiędzy których trzeba określić właściwą funkcję prognozy. Dla zależności liniowej związek ten przyjmuje postać:

$$Y^{p}_{t+\tau} = \hat{Y}_t + \tau \cdot W \quad (21)$$

Przy założeniu zróżnicowanych wag informacji:

$$Y^{p}_{t+\tau} = \hat{Y}_t + \tau \cdot \bar{W} \quad (22)$$

W przypadku związku potęgowego i stałych wag informacji:

$$Y^{p}_{t+\tau} = \hat{Y}_t \cdot \left(\frac{t+\tau}{t} \right)^{a_1} \quad (23)$$

gdzie:

a_1 — średnia elastyczność:

$$a_1 = \frac{1}{n-1} \sum_{t=1}^{n-1} a_{1,t} \quad (24)$$

$a_{1,t}$ — elastyczność cząstkowa:

$$a_{1,t} = \frac{1}{2} \frac{\hat{y}_{t+1} - \hat{y}_t}{\hat{x}_{t+1} - \hat{x}_t} \left[\frac{\hat{x}_t}{\hat{y}_t} + \frac{\hat{x}_{t+1}}{\hat{y}_{t+1}} \right] \quad (25)$$

Przy zróżnicowanych wagach informacji:

$$a_1 = \sum_{t=1}^{n-1} C^n_{t+1} \cdot a_{1,t} \quad (26)$$

Dla zależności wykładniczej:

$$Y^{p}_{t+\tau} = a_2 \cdot (a_1)^\tau \quad (27)$$

$$a_1 = 1 + \bar{W} \quad (28)$$

$$a_2 = \hat{Y}_t \quad (29)$$

przyrost określa związek:

$$W_{t+1} = \frac{\hat{y}_{t+1} - \hat{y}_t}{\hat{y}_t} \quad (30)$$

a w przypadku zróżnicowanych wag informacji:

$$a_1 = 1 + \bar{W} \quad (31)$$

Ocena jakości prognozy

Przyjmowane kryterium oceny jakości modelu musi uwzględnić brak znajomości przyszłego rozkładu zmiennej prognozowanej, co jest szczególnie istotne w warunkach prognoz perspektywicznych dla systemów kanalizacyjnych. Warunek ten spełnia przedział ufności skonstruowany w oparciu o nierówność Czebyszewa:

$$P(|X - \bar{X}| > r \cdot S) < \frac{1}{r^2} \quad (32)$$

gdzie:

X — wartość zmiennej losowej,

$\bar{X}$ — średnia wartość zmiennej losowej,

S — odchylenie standardowe,

r — współczynnik nierówności Czebyszewa.

Wartość współczynnika nierówności Czebyszewa można uzmiennić w czasie, przyjmując współczynnik wag harmonicznych C^n_{n-t} :

$$r(\tau) = r \cdot \sum_{t=\tau+1}^n C^n_{n-t} \quad (33)$$

przy założeniu narastania wielkości błędu w miarę pogłębiania horyzontu prognozy. Granice przedziału ufności można określić z zależności:

$$\Delta Y = r(\tau) \cdot S \quad (34)$$

przy czym współczynnik r zmierza asymptotycznie do wartości oczekiwanej. Przyjmując $r=4$ zakłada się prawdopodobieństwo realizacji zdarzenia na poziomie 95%.

Ostateczny sposób określenia granicy dolnej (d) i górnej (g) przedziału ufności zależy od postaci równania prognostycznego.

Dla predyktora liniowego i potęgowego jest on następujący:

$$g(d)Y_{n+t}^p = Y_{n+t}^p \pm r(\tau) \cdot S \quad (35)$$

a dla predyktora wykładniczego:

$$g(d)Y_{n+t}^p = Y_{n+t}^p [1 \pm r(\tau) \cdot S] \quad (36)$$

W warunkach prognozowania o bardzo głębokim horyzoncie i przy istotnej nieregularności danych wyjściowych, możliwe jest uzyskanie prognoz obarczonych znacznym błędem. Istotą problemu stanowią wówczas odstępstwa analizowanego zjawiska w minionym okresie od zasady inercji i niekiedy poznanie mechanizmów kształtujących je w przeszłości może pozwolić na pewne uściślenia. Generalnie też nie można wykluczyć wystąpienia w przyszłości nowych tendencji, czy też odmiennych powiązań pomiędzy zmiennymi, zaburzających dotychczasowe związki. W wyniku tego błąd prognozowania może przekroczyć wszelkie wcześniejsze oczekiwania [3], a jedynym sposobem przeciwdziałania jest zapewnienie ciągłości prognozowania. Korekty muszą być realizowane systematycznie — po upływie kolejnych okresów stanowiących podstawę (krok) prognozy. Generalnie przy tym wskazane jest, aby materiały wyjściowe do budowy prognozy obejmowały dłuższy okres od jej horyzontu.

Podsumowanie

Przedstawiona metoda pozwala budować prognozy dla procesów dynamicznych przy zachowaniu bardzo głębokich horyzontów oraz przy wykorzystaniu wyłącznie informacji ilościowych stwierdzonych w minionych okresach. Dokładność prognozowania rośnie przy tym w miarę wzrostu populacji danych wyjściowych. Wskazane jest, aby szereg wyjściowy był dłuższy o co najmniej jeden okres w stosunku do horyzontu prognozy.

Proponowane rozwiązanie pozwala ograniczyć do minimum liczbę wykorzystanych zmiennych objaśniających analizowane zjawisko. Możliwe jest w ogóle wyeliminowanie potrzeby korzystania z prognoz objaśniających, uzupełniających informacje wyjściowe i mogących bardzo niekorzystnie oddziaływać na jakość prognozy finalnej.

Powyższe cechy predestynują przedstawiony model do wykorzystania przy budowie prognoz ilościowych w kanalizacji dla okresu perspektywicznego. W razie potrzeby można też sporządzać niezależne od siebie prognozy dla poszczególnych okresów jednoimiennych. To ostatnie jest szczególnie ważne w warunkach miejscowości o charakterze sezonowym, a zwłaszcza rejonów nadmorskich stanowiących zaplecze turystyki. Wówczas w krótkim czasie (ok. 3 miesięcy w roku) ma miejsce znacznie większy (nawet wielokrotnie) pobór wody i odpływ ścieków sanitarnych w stosunku do całorocznego.

LITERATURA

1. K. ABRAMSKI, A. SADURSKI, Z. SULIGOWSKI: Uwarunkowania rozwoju Półwyspu Helskiego. Inżynieria Morska, nr 4, 1983, ss. 339—342.
2. A. J. CHORCHOT, G. J. LAWRIK, W. W. SAWCZENKO: Sowriemiennye tiendencii razwitiia krupnykh gorodow Ukrainy. Izwiestija WUZow, r. Stroitelstwo i Architektura, nr 5, NISI, Nowosibirsk 1980.
3. Z. CZERWINSKI, B. GUZIK: Prognozowanie ekonometryczne. PWE, Warszawa 1980.
4. Ekonometria. PWE, Warszawa 1984.
5. J. W. FORRESTER: Urban dynamics. MIT Press, za wyd. ros.: Dinamika razwitiia goroda, Progiess, Moskwa 1974.
6. M. KRZYSZTOFIK, D. URBANEK: Metody statystyczne. PWN, Warszawa 1977.
7. Nowyye napravleniia w projektirovanii i stroitelstwie kanalizacii bolszhix gorodow. Strojizdat, Moskwa 1975.
8. B. I. OGLY, W. S. CZUTONOW: Ob ewolucionnoj klassifikacii aglomeracionnykh sistem. Izwiestija WUZow, r. Stroitelstwo i Architektura, nr 12, NISI, Nowosibirsk 1983.
9. Oschiennosti planirowocznykh i konstruktivnykh rieszenij wodoprowoda i kanalizacii bolszhix gorodow. Strojizdat, Moskwa 1972.

MODEL OF A QUANTITATIVE LONG-RANGE FORECAST FOR THE MANAGEMENT OF A SEWER SYSTEM

Presented is a model of a quantitative forecast for an arbitrary horizon. The phenomena under analysis should be continuous (or quasi-continuous) in nature and inert with time. The forecast involves only information from the past, and makes no use of supporting forecasts, so the model is particularly appropriate for sewer systems. The methods applied con-

sist in eliminating incidental variations (moving trend method) and in approximating calculated data to those obtained by measurement (creeping trend method). The form of the forecast is derived by analyzing the creeping trend values. Forecasting may be carried out by assuming either constant or time variable weights of more or less recent information. The paper also gives the relationships for the basic forms (linear, powered and exponential) of the equation of the trend for many years' periods.