

*Математична модель
як образ спостережуваного
об'єкта або процесу:
проблема оптимізації
в умовах неповноти інформації*

В.С.Степашко

План доповіді

1. Вступ
2. Постановка задачі моделювання
3. Властивості дисперсії помилки моделі
4. Метод критичних дисперсій
5. Аналітичне порівняння критеріїв
6. Окремі задачі теорії моделювання
7. Ілюстративний приклад
8. Алгоритм поліноміальної складності
9. Приклади порівняння МГУА та регресійного аналізу
10. Довжина вибірки і точність прогнозування (приклад)
11. Висновки

Вступ

Аналізується *проблема вибору найкращої моделі об'єкта в умовах неповноти інформації:*

- невеликий обсяг даних
- невідомий характер і рівень шуму
- неточний склад релевантних факторів
- відсутні обмеження на параметри моделі
- неоднозначний вибір ефективного критерію структурної ідентифікації

Постановка задачі

вибору найкращої моделі, лінійної за параметрами

1) вибірка $W=(X | y)$, $\dim X = n \times m$,
 $\dim y = n \times 1$, $\text{rank } X = m$

2) $y = \overset{o}{y} + \overset{o}{\xi}$, $\overset{o}{y} = E y = X \theta_0$

3) $E \xi = 0$, $E \xi \xi^T = \sigma^2 I_n$

4) Множина моделей: $\mathfrak{T}_m : y_s = X_s \hat{\theta}_s$, $s = \overline{1, m}$

5) “Ідеальний” критерій – дисперсія помилки моделі:

$$J(s) = E \left\| \overset{o}{y} - \hat{y}_s \right\|^2 = E \left\| \overset{o}{y} - X_s \hat{\theta}_s \right\|^2$$

6) Дисперсія помилки прогнозування:

$$J_F(s) \stackrel{\Delta}{=} J_F(s, X, X_F) = \mathbf{E} \left\| \overset{o}{y}_F - \hat{y}_{Fs} \right\|^2 = \mathbf{E} \left\| \overset{o}{y}_F - X_{Fs} \hat{\theta}_s \right\|^2$$

$$X_F \neq X, \dim X_F = n_F$$

7) Реальний критерій $CR(s) > 0$

Оптимальна структура для *відновлення*:

$$s^o = \arg \min_{s=1,m} J(s) \quad (1)$$

Оптимальна структура для *прогнозування*:

$$s_B^o = \arg \min_{s=1,m} J_F(s) \quad (2)$$

Залежність від характеристик невизначеності:

$$J(s) = \varphi(s, n, \overset{o}{\sigma^2}, \overset{o}{\theta_0}, \overset{o}{y}, \overset{o}{X}, \overset{o}{X}), \quad X = (\overset{?}{X} : \overset{o}{\tilde{X}}) \quad (3)$$

$$J_F(s) = \varphi_F(s, n, \overset{o}{\sigma^2}, \overset{o}{\theta_0}, \overset{o}{y}, \overset{o}{y}_F, \overset{o}{X}, \overset{o}{X}_F, \overset{o}{X}, \overset{o}{X}_F) \quad (4)$$

Властивості дисперсії помилки моделі

$$J(s) = J^b(s) + J^v(s) = \left\| \overset{o}{y} - X_s \bar{\theta}_s \right\|^2 + \sigma^2 s, \quad \bar{\theta}_s = E[\hat{\theta}_s] \quad (5)$$

$J^b(s)$ – структурна, $J^v(s)$ – шумова складові

$$J_F(s) = J_F^b(s) + J_F^v(s) = \left\| \overset{o}{y}_F - X_{Fs} \bar{\theta}_s \right\|^2 + \sigma^2 \text{tr} \left[\left(X_s^T X_s \right)^{-1} X_{Fs}^T X_{Fs} \right] \quad (6)$$

Рекурентні вирази:

$$J(s+1) = J(s) - \bar{\mathcal{G}}_{s+1}^2 \beta_{s+1} + \sigma^2, \quad X_{s+1} = [X_s \dot{:} x] \quad (7)$$

$$\bar{\mathcal{G}}_{s+1} = \frac{\alpha_{s+1}}{\beta_{s+1}}, \quad \alpha_{s+1} = x^T D_s \overset{o}{y}, \quad \beta_{s+1} = x^T D_s x$$

$$J_F(s+1) = J_F(s) - 2\bar{\mathcal{G}}_{s+1} a_F^T b_F +$$

$$+ \bar{\mathcal{G}}_{s+1}^2 b_F^T b_F + \sigma^2 \frac{b_F^T b_F}{\beta_{s+1}} \quad (8)$$

$$X_{F,s+1} = [X_{Fs} \dot{:} x_{Fs}]; \quad \bar{\theta}_{s+1} = \left[\bar{\theta}_s^T \bar{\mathcal{G}}_{s+1} \right]$$

$$a_F \equiv a_F(s) = \overset{o}{y}_F - X_{Fs} \bar{\theta}_s,$$

$$b_F \equiv b_F(s) = x_F - X_{Fs} (X_s^T X_s)^{-1} X_s^T x$$

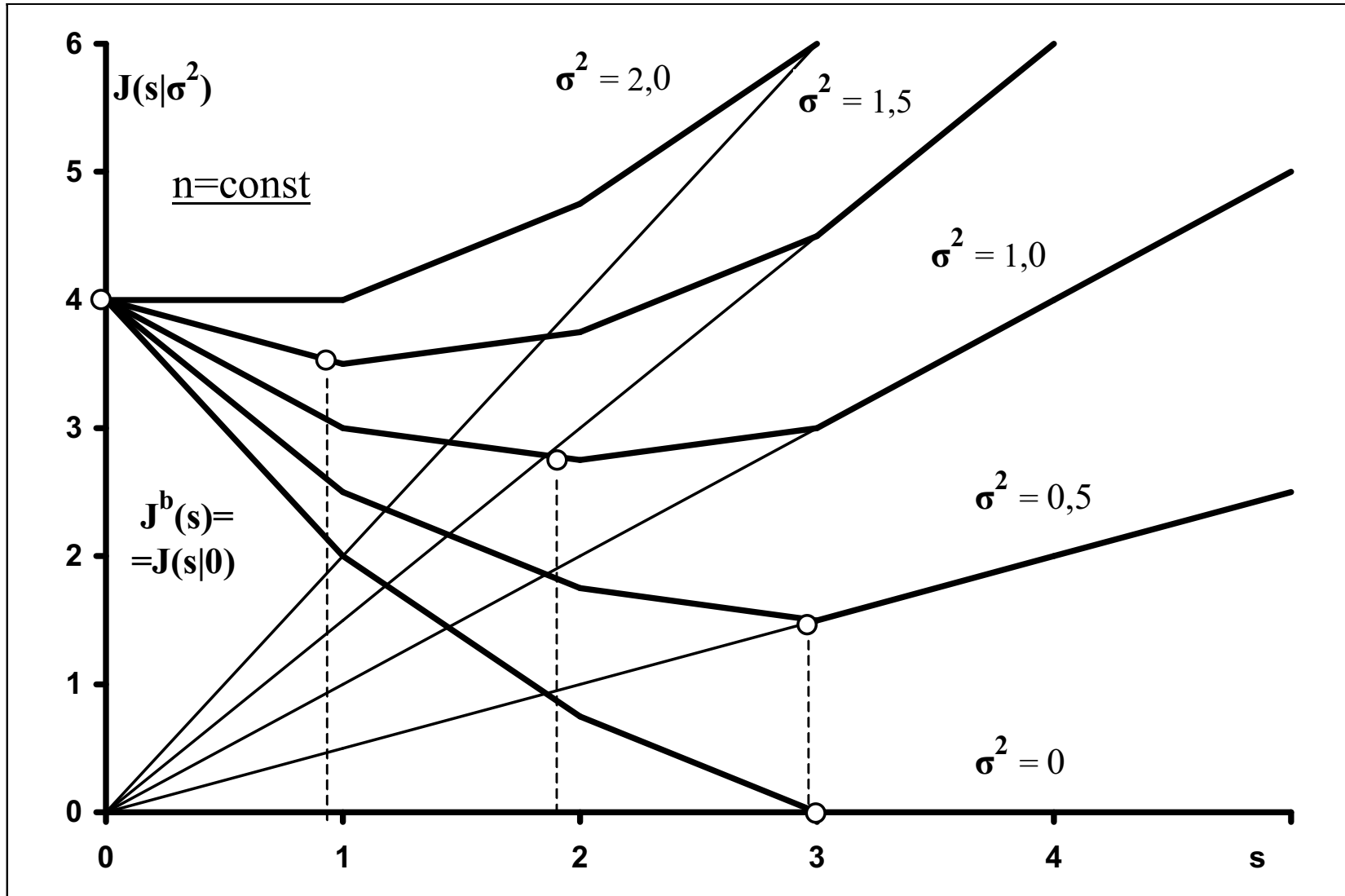


Рис.1. Дисперсія помилки моделі як функція складності структури s за різних дисперсій шуму (кружечками вказано положення мінімумів критерію $J(s)$, істинна структура відповідає складності $s_0 = 3$, загальна кількість регресорів $m = 5$)

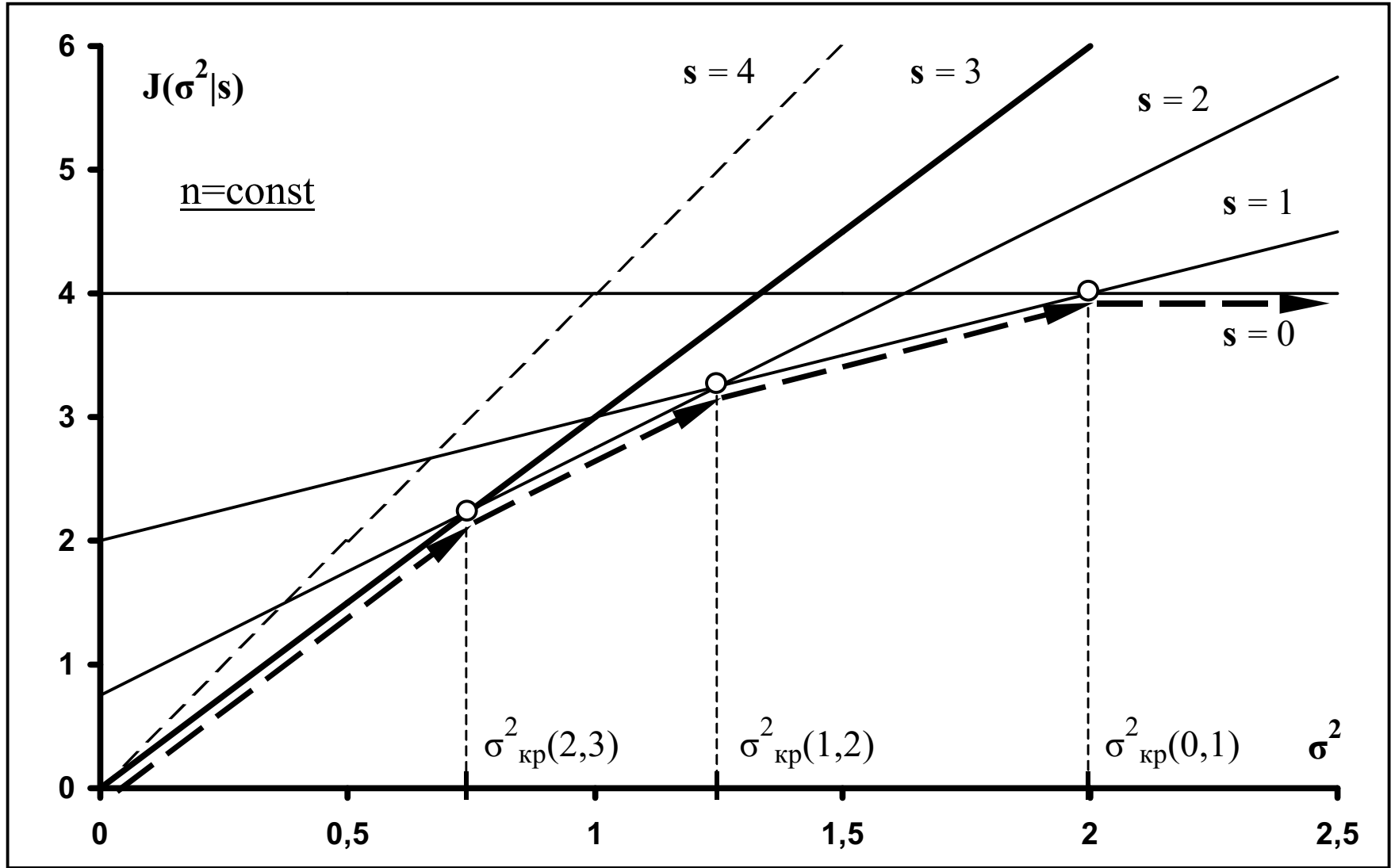
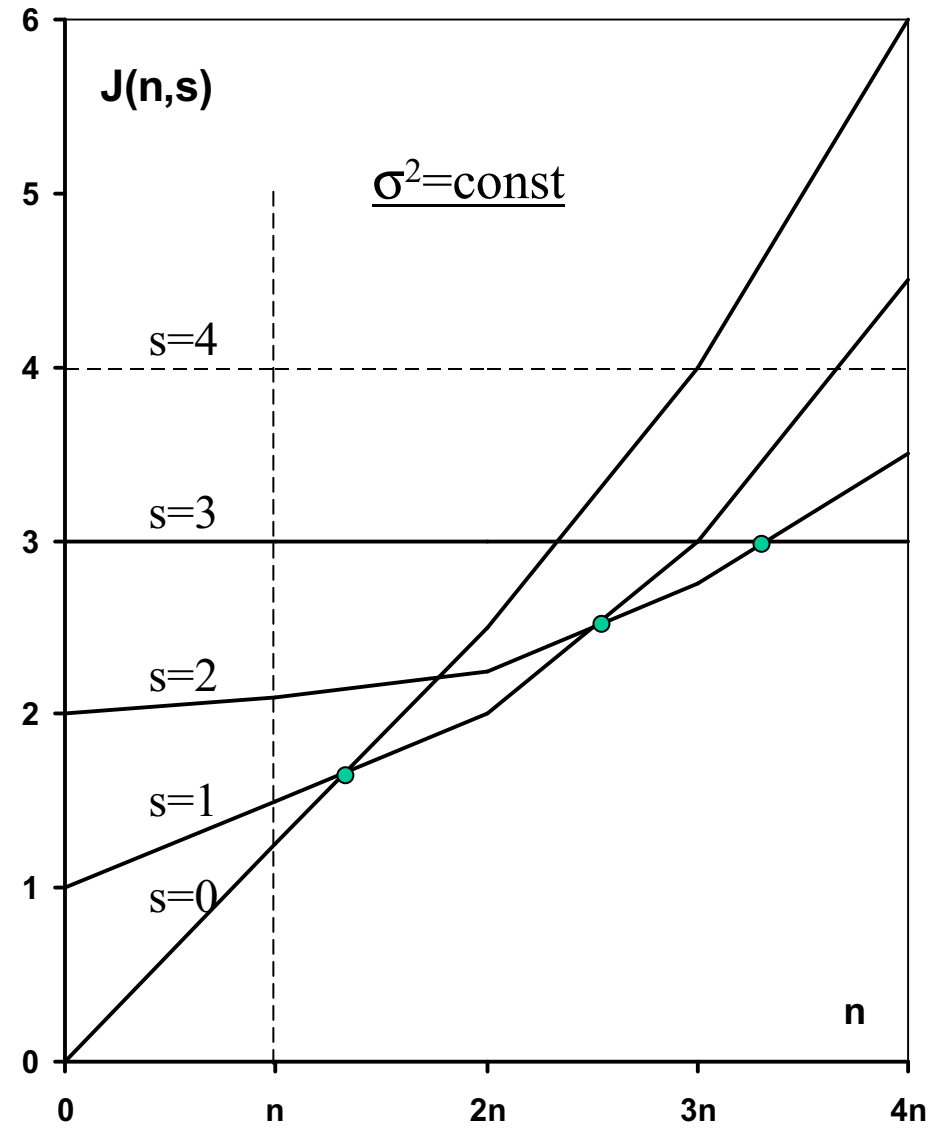
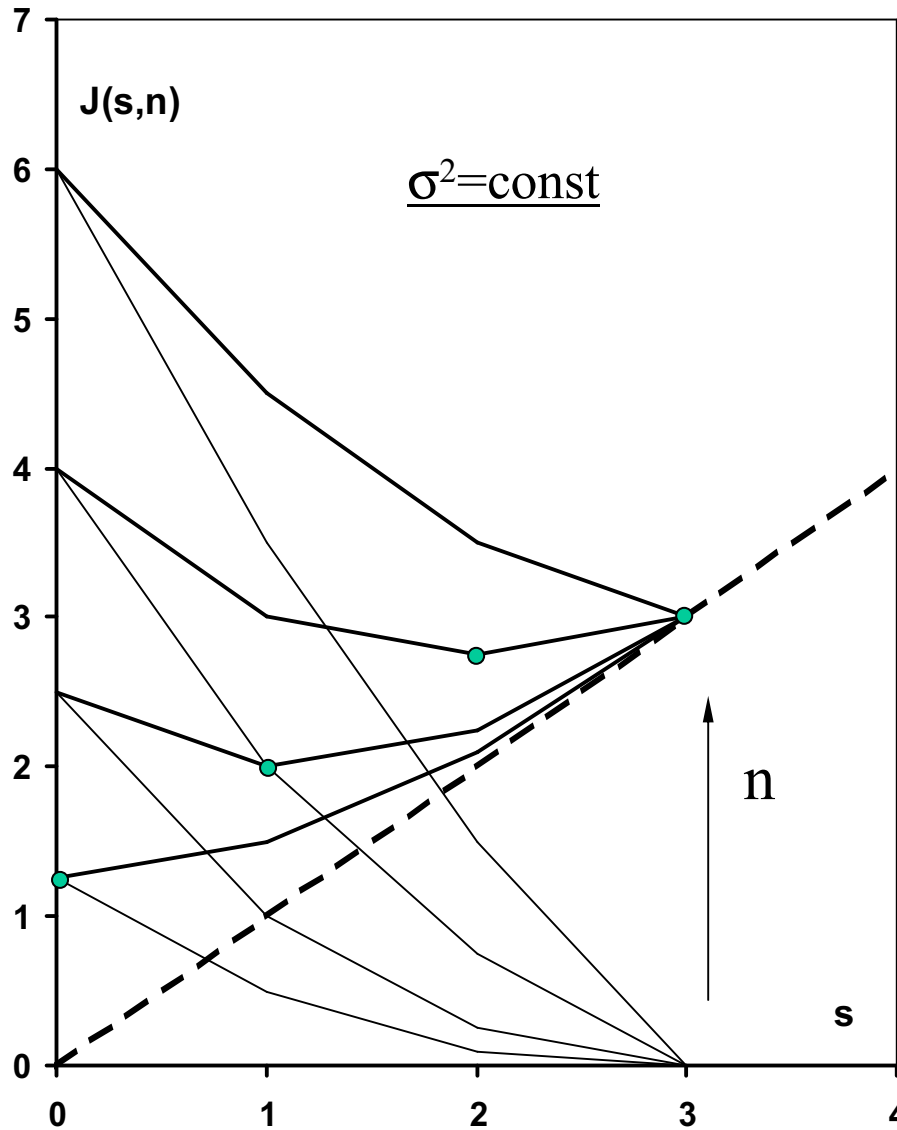


Рис.2. Дисперсія помилки моделі як функція дисперсії шуму σ^2 для структур різної складності (кружечками вказано критичні точки зміни структури моделі, яким відповідають критичні дисперсії $\sigma^2_{кр}(s,s+1)$; стрілками позначено *екстремаль* критерію $J(s)$)



Дисперсія помилки моделі як функція складності моделі S та числа точок Π

Метод критичних дисперсій

Критичні дисперсії (координати точок перемикання):

$$\sigma_{kp}^2(s, s+1): J(\sigma^2 | s) = J(\sigma^2 | s+1), \quad s = \overline{1, m-1} \quad (9)$$

Умова оптимальності:

$$\sigma_{kp}^2(s-1, s) > \sigma^2 \geq \sigma_{kp}^2(s, s+1) \quad (10)$$

Критичні дисперсії для $J(s)$, $J_B(s)$:

$$\sigma_{kp}^2(s, s+1) = \bar{\mathcal{G}}_{s+1}^2 \beta_{s+1}, \quad (11)$$

$$\sigma_{Fkp}^2(s, s+1) = 2\bar{\mathcal{G}}_{s+1} \beta_{s+1} \frac{a_F^T b_F}{b_F^T b_F} - \bar{\mathcal{G}}_{s+1}^2 \beta_{s+1} \quad (12)$$

Аналіз ефективності критеріїв

Клас критеріїв:

$$CR(s, n) = \rho_1(s, n)V(s, n) + \rho_2(s, n)\hat{\sigma}^2$$

Критерій Маллоуза:

$$C_p(s) = RSS(s) + 2\hat{\sigma}^2 s \quad (15)$$

Критерій Акаїке:

$$FPE(s) = \frac{n + s}{n - s} RSS(s); \quad (16)$$

Критерій регулярності $AR(s)$:

$$AR(s) = \|y_B - \hat{y}_{Bs}\|^2 = \|y_B - X_{Bs}\hat{\theta}_{As}\|^2 \quad (17)$$

$$W = (A^T B^T)^T, \quad A = (X_A, y_A), \quad B = (X_B, y_B), \quad n = n_A + n_B$$

Критичні дисперсії:

$$\sigma_{Cp}^2(s, s+1) = \sigma_{kp}^2(s, s+1) - 2(\bar{\sigma}^2 - \sigma^2) \quad (18)$$

$$\sigma_{FPE}^2(s, s+1) = \sigma_{kp}^2(s, s+1) \times$$
$$\times \left[1 + 2 \frac{(n-s)(s+1) - n\mu_s}{(n-s)(n-s-1)} \right] \quad (19)$$

$$\sigma_{AR}^2(s, s+1) = \frac{1}{1 + \rho_B^2} \sigma_{kp}^2(s, s+1) \quad (20)$$

Прикладні задачі теорії

МОДЕЛЮВАННЯ В УМОВАХ НЕВИЗНАЧЕНОСТІ

Збільшення кількості точок

$$\sigma_{kp}^2(n+1, m-1, m) = \sigma_{kp}^2(n, m-1, m) + \varphi^2(n, m) \mathcal{G}_{0m}^2 \quad (21)$$

Асимптотичний аналіз

Асимптотична незміщеність:

$$\bar{s}^o \rightarrow s_o : \lim_{n \rightarrow \infty} \bar{J}(\bar{s}^o) = \lim_{n \rightarrow \infty} \bar{J}(s_o) \quad (22)$$

Необхідна й достатня умова:

$$\sigma_{kp}^2(n, s, s+1) \xrightarrow{n} \infty. \quad (23)$$

Сильна регулярність: $\lim_{n \rightarrow \infty} \frac{1}{n} X_n^T X_n = \bar{H}$

Планування експерименту. Умова оптимальності:

$$\rho^2 X^T X = X_F^T X_F, \quad \rho^2 \neq 0 \quad (24)$$

Моделювання в неточному класі структур

$$X \neq (\overset{\circ}{X} : \overset{\circ}{\tilde{X}}), \quad \overset{\circ}{y} \neq X\theta_0 \quad (25)$$

Невизначеність, еквівалентна додатковому шумові

Оптимальність істинної структури

$$\sigma_{кр}^2(m-1, m) \equiv \mathcal{G}_{0m}^2 \beta_m \geq \sigma^2, \quad \beta_m \geq \frac{\sigma^2}{\mathcal{G}_{0m}^2} \quad (26)$$

$$\beta_m = x^T x - x^T X_{m-1} (X_{m-1}^T X_{m-1})^{-1} X_{m-1}^T x$$

Ілюстративний приклад

$$y = a + bt + \xi, \quad t = \overline{1, n}, \quad m = 2$$

Критичне число точок для b :

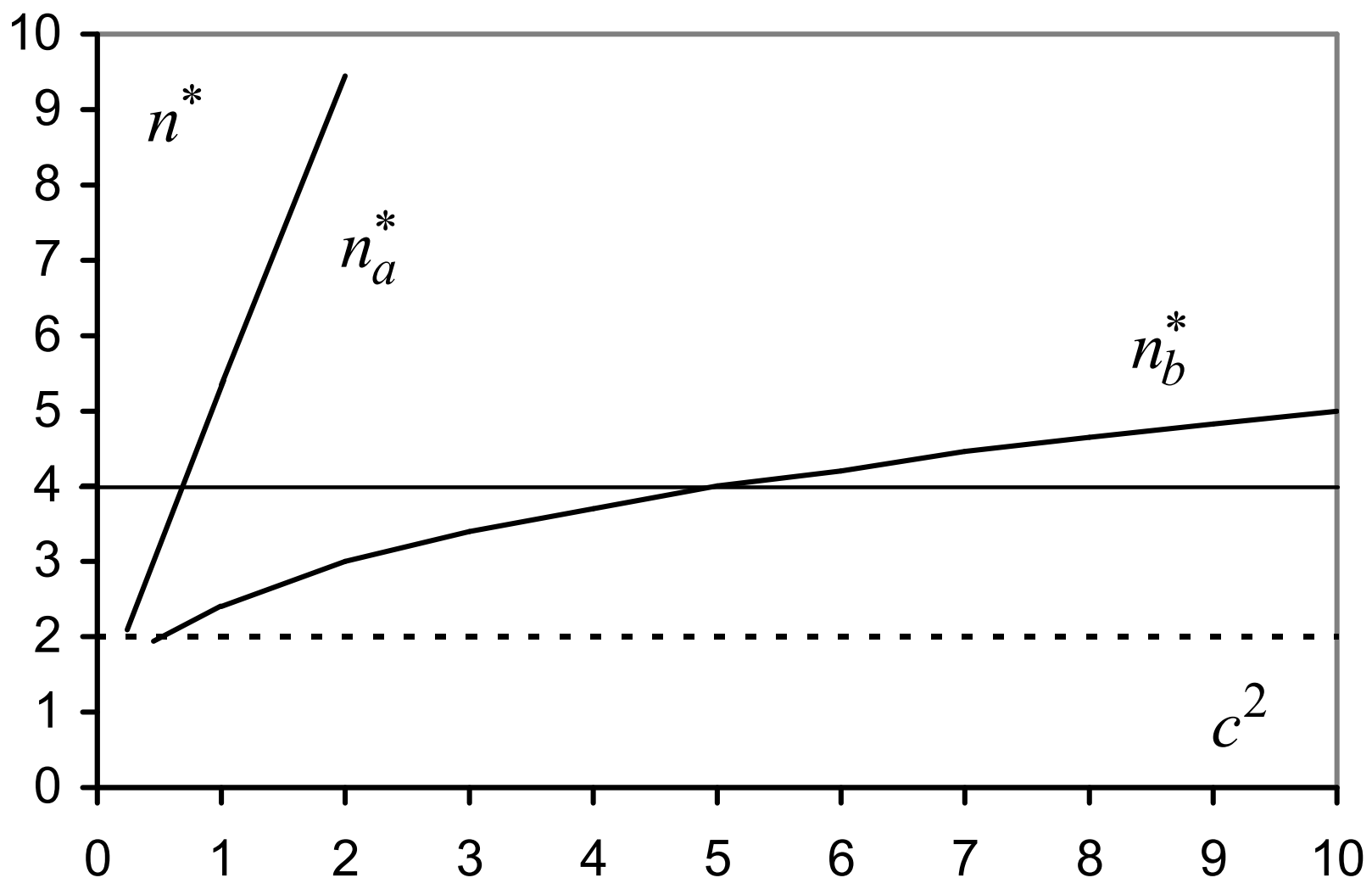
$$\frac{1}{12} n(n-1)(n+1) = \frac{\sigma^2 \Delta}{b^2} = c_b^2, \quad n^* \approx \sqrt[3]{c_b^2}$$

Критичне число точок для a :

$$\frac{n(n-1)}{2(2n+1)} = \frac{\sigma^2 \Delta}{a^2} = c_a^2, \quad n^* \approx c_a^2$$

Балансове співвідношення
(індикатор інформативності):

$$r^*(n) = \frac{a^2}{b^2} = \frac{1}{6} (n+1)(2n+1)$$



Залежність критичної довжини вибірки від рівня шуму

Комбінаторно-селекційний
алгоритм
поліноміальної складності
для послідовного пошуку
оптимальної моделі
(результату повного перебору)

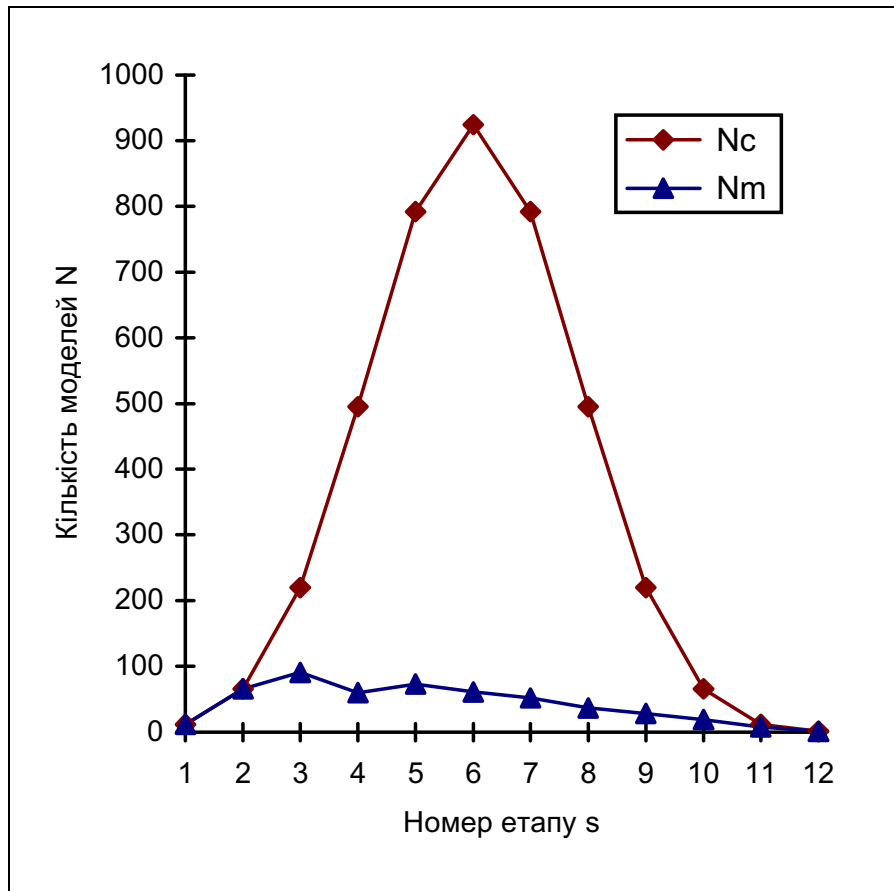


Рис. 2. Обсяг перебору моделей на кожному етапі

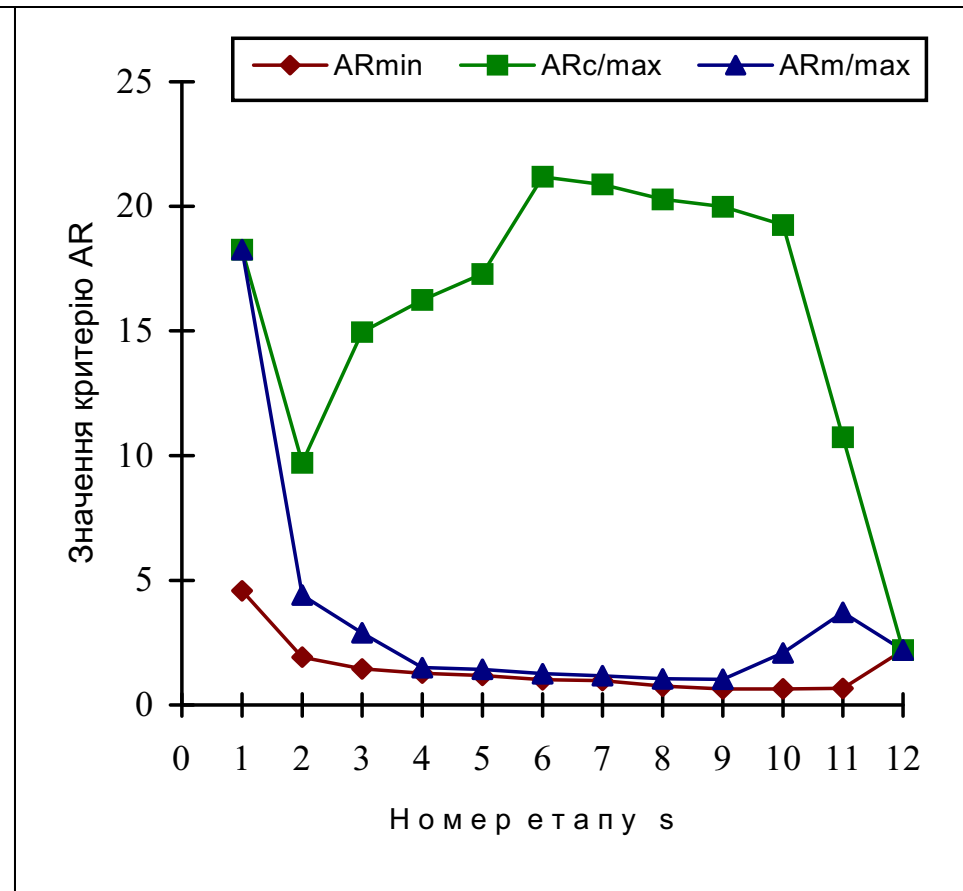


Рис. 3. Діапазон зміни значень критерію регулярності на кожному етапі при $m=11$

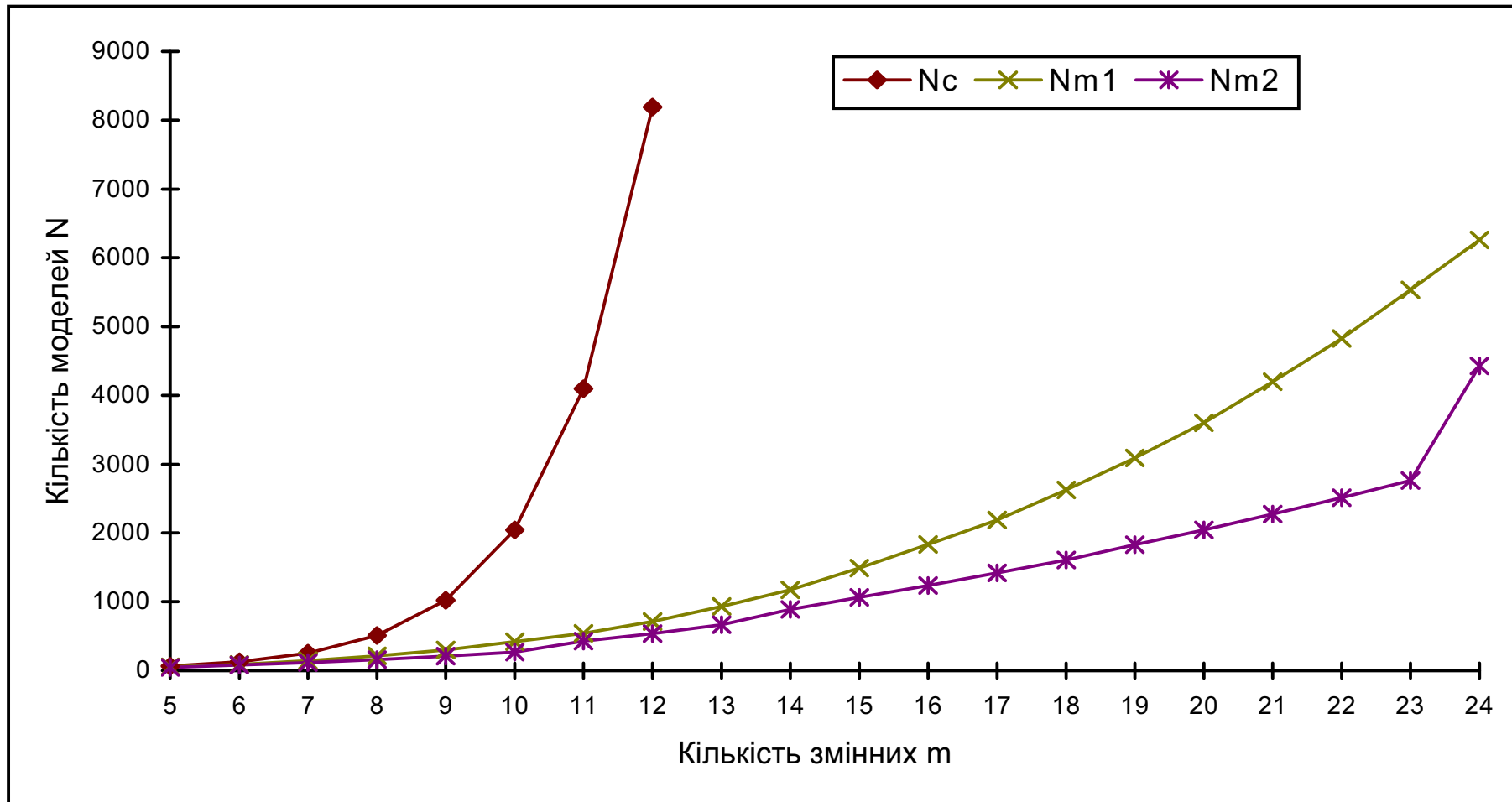


Рис. 1. Залежність кількості порівнюваних моделей від числа змінних

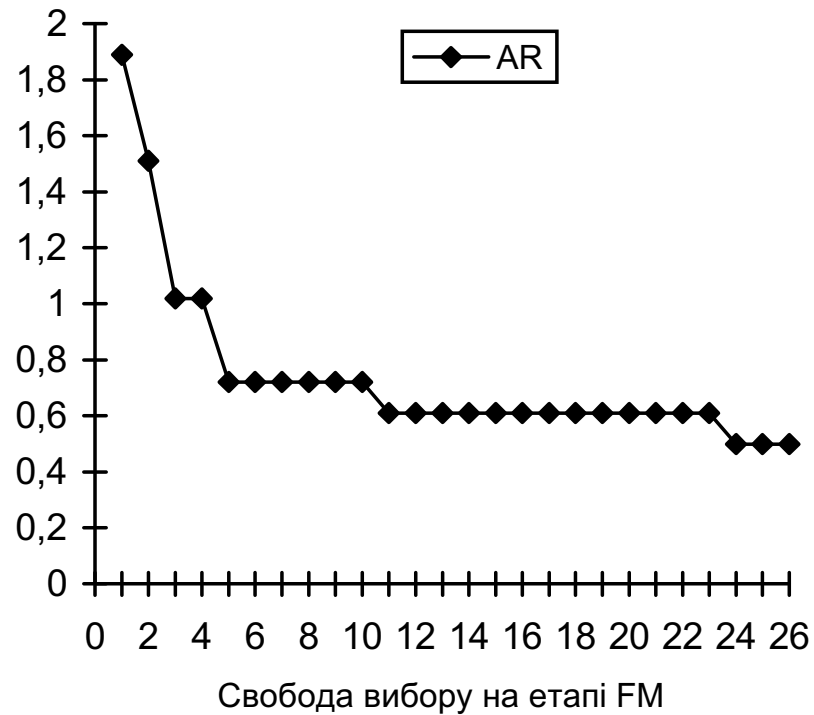


Рис.5 Залежність значень критерію якості моделей AR від свободи вибору FM ($m=46$)

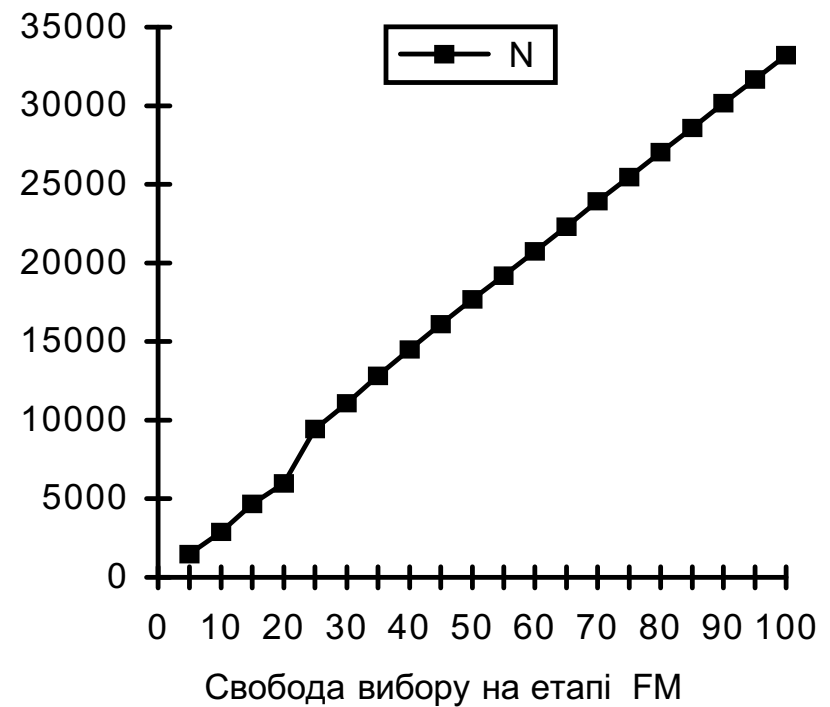


Рис. 6. Залежність числа моделей N від свободи вибору на етапі FM ($m=46$)

Порівняння ефективності
застосування регресійного аналізу
та МГУА для прогнозування
економічних процесів

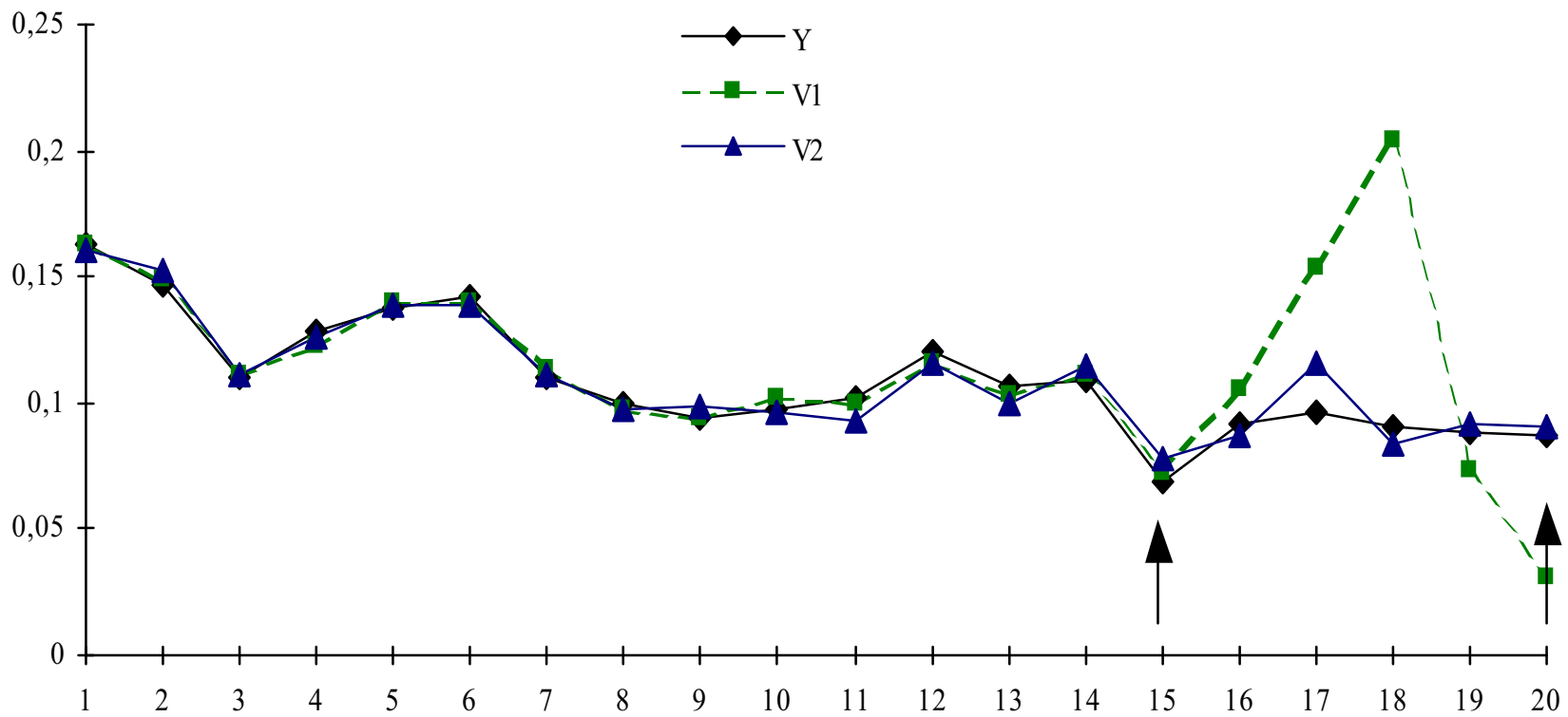


Рис. 1.. Порівняння точності моделей МНК (V1) та МГУА (V2) обсягу виробництва легкої промисловості (стрілками відзначено початок і кінець екзаменаційної послідовності)

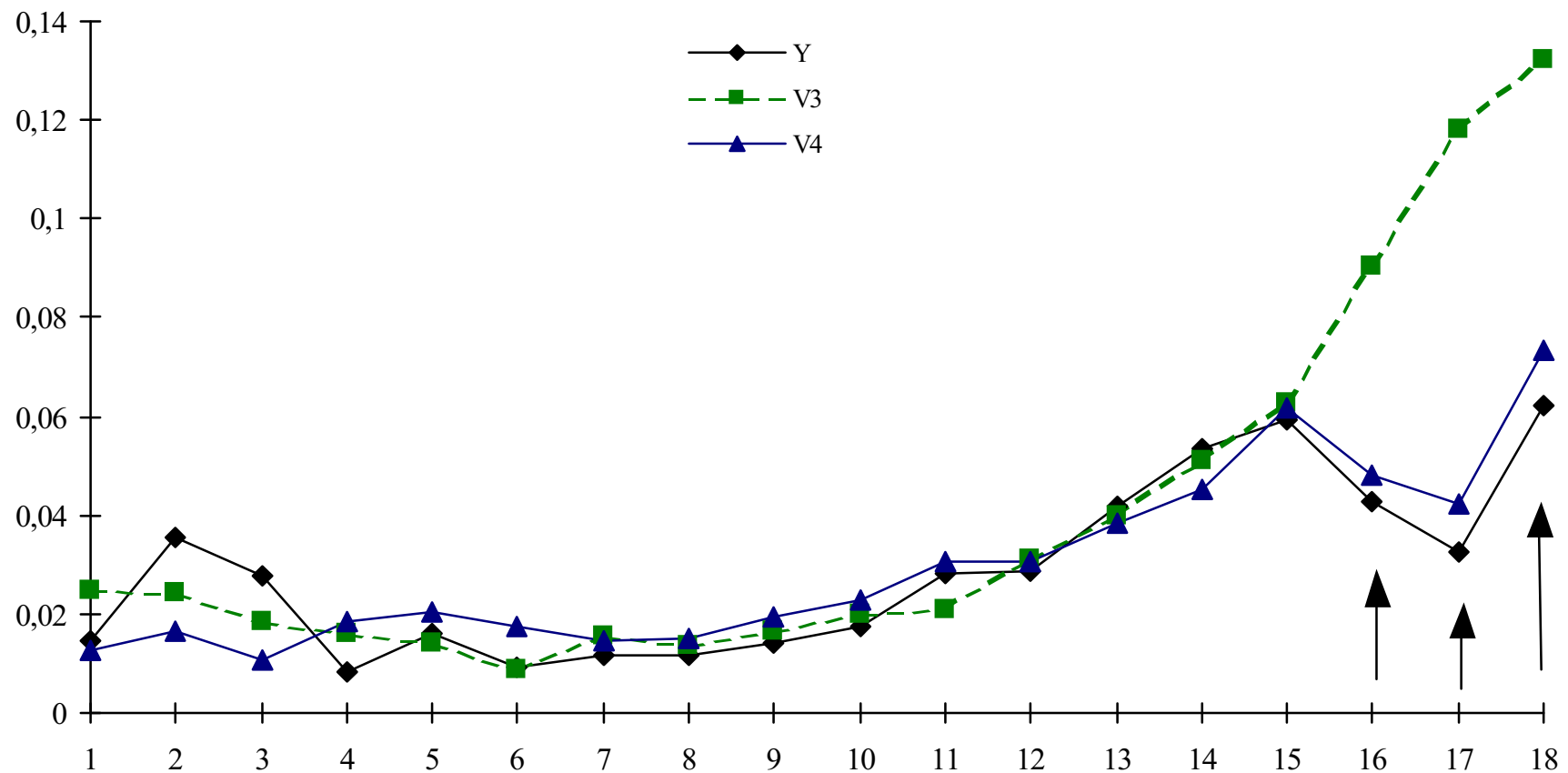


Рис. 2. Порівняння реального перебігу інфляції Y та її оцінок, одержаних за МНК ($V3$) та МГУА ($V4$) (стрілками відзначено екзаменаційні точки)

ЗАЛЕЖНІСТЬ ЕФЕКТИВНОСТІ
ПРОГНОЗУВАННЯ
ПОДАТКОВИХ НАДХОДЖЕНЬ
ВІД ДОВЖИНИ СТАТИСТИЧНОЇ
ВИБІРКИ

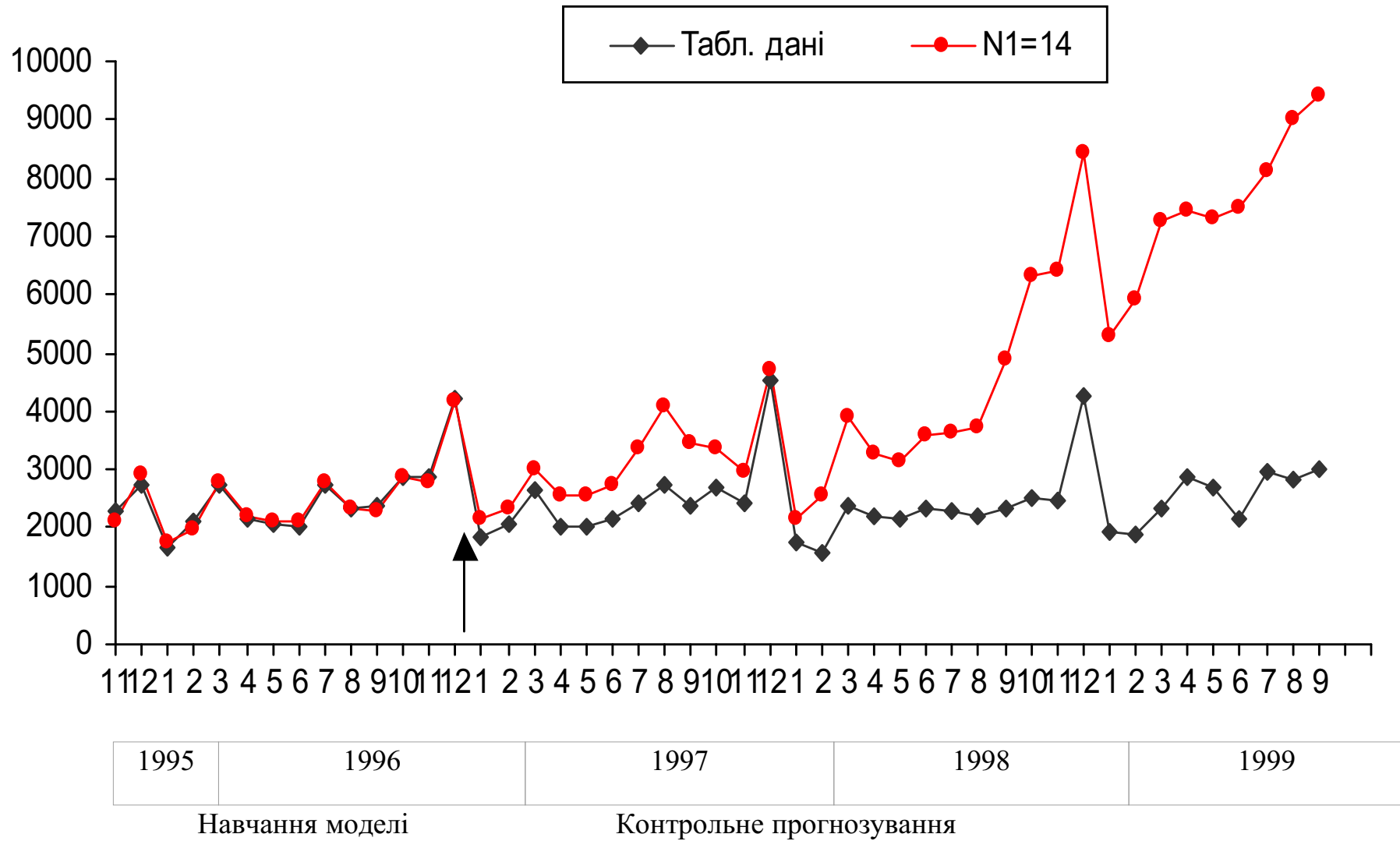


Рис.1. Якість прогнозування за моделлю (1) (14 точок навчання)

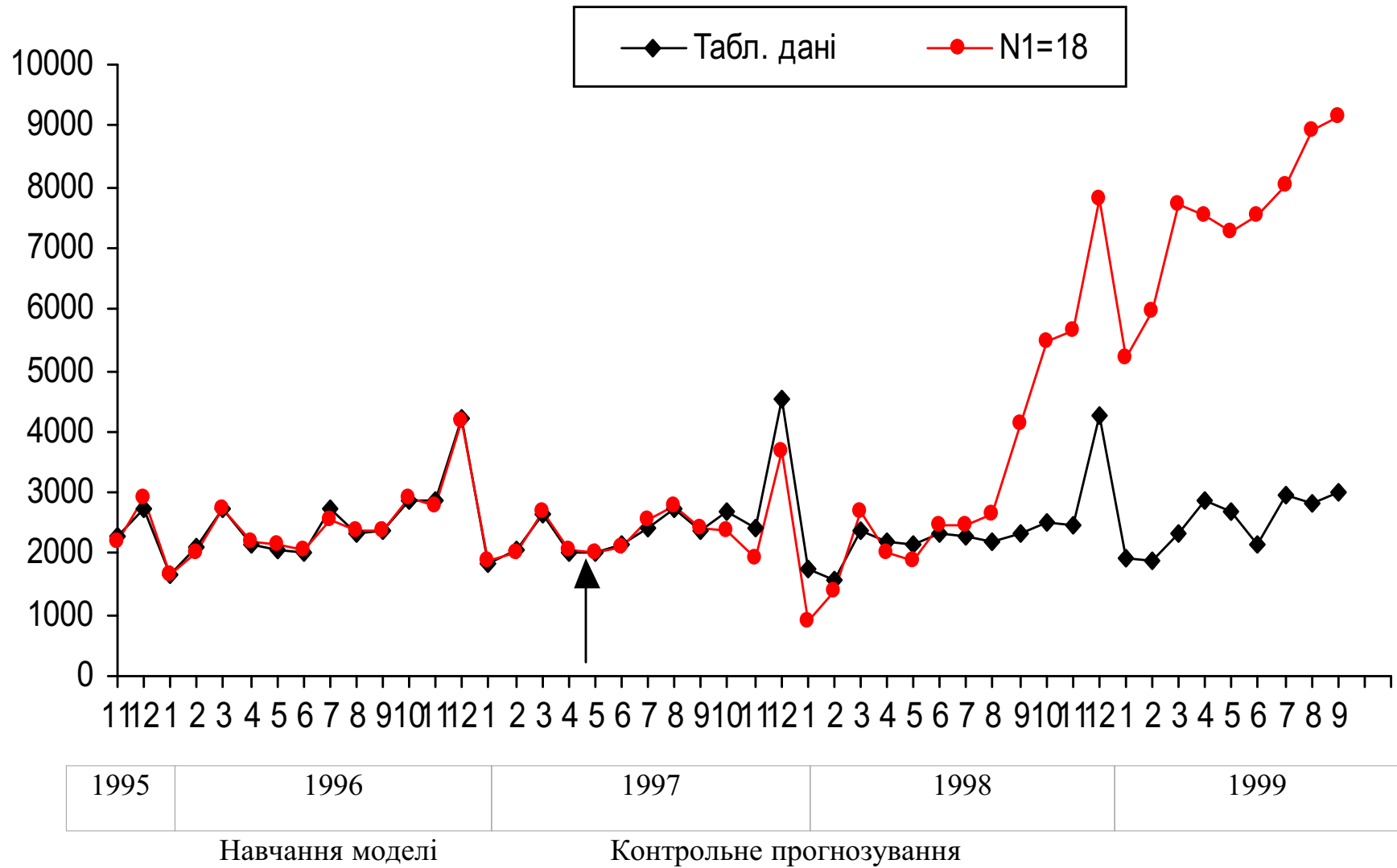


Рис. 2. Якість прогнозування за моделлю (2) (18 точок навчання)

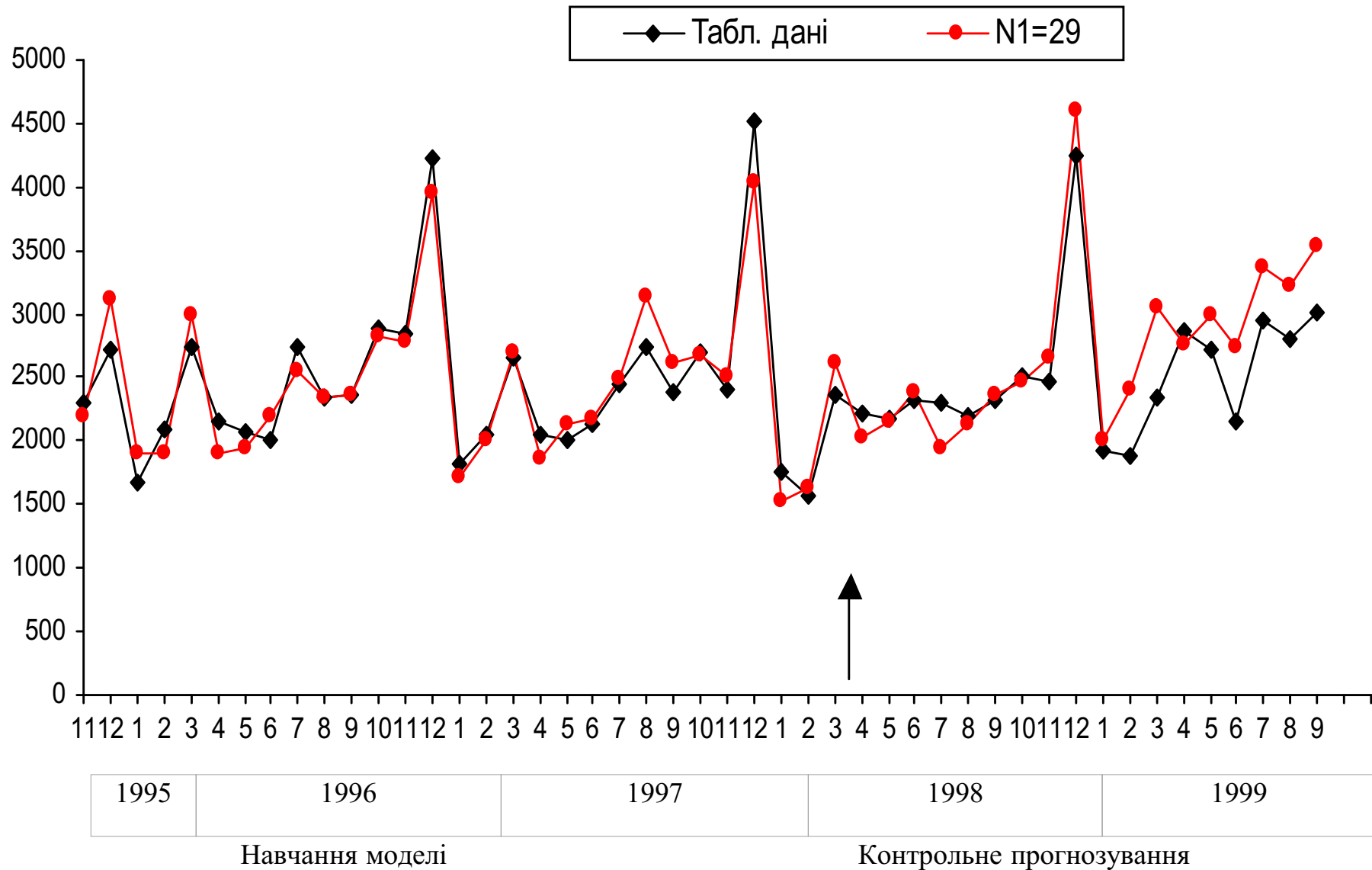


Рис. 3. Якість прогнозування за моделлю (3) (29 точок навчання)