



ДОСЛІДЖЕННЯ ДИНАМІКИ ВРОЖАЙНОСТІ ОЗИМОЇ ПШЕНИЦІ В РОЗРІЗІ ОБЛАСТЕЙ УКРАЇНИ

Грицюк П.М.
к.ф.-м.н., докторант

*Національний
університет водного
господарства та
природокористування*

м.Рівне





Мета дослідження

- *Гармонічний, статистичний, фрактальний та інтелектуальний аналіз часових рядів врожайності озимої пшениці*
- *Побудова та апробація прогнозних моделей врожайності*
- *Вироблення методики оцінки ризиків зерновиробництва*



Завдання дослідження

- ❑ *Дослідити характер циклічності врожайності озимої пшениці в різних областях України*
- ❑ *Оцінити ступінь детермінованості часових рядів врожайності для областей України*
- ❑ *Побудувати прогностні моделі врожайності (горизонт 1 рік і більше) та провести їх апробацію*
- ❑ *Виконати оцінки ризиків зерновиробництва для різних регіонів України*



1. Дослідження циклічності врожайності

*вихідні дані – ряди врожайності озимої пшениці в
областях України (1955–2008 роки)*



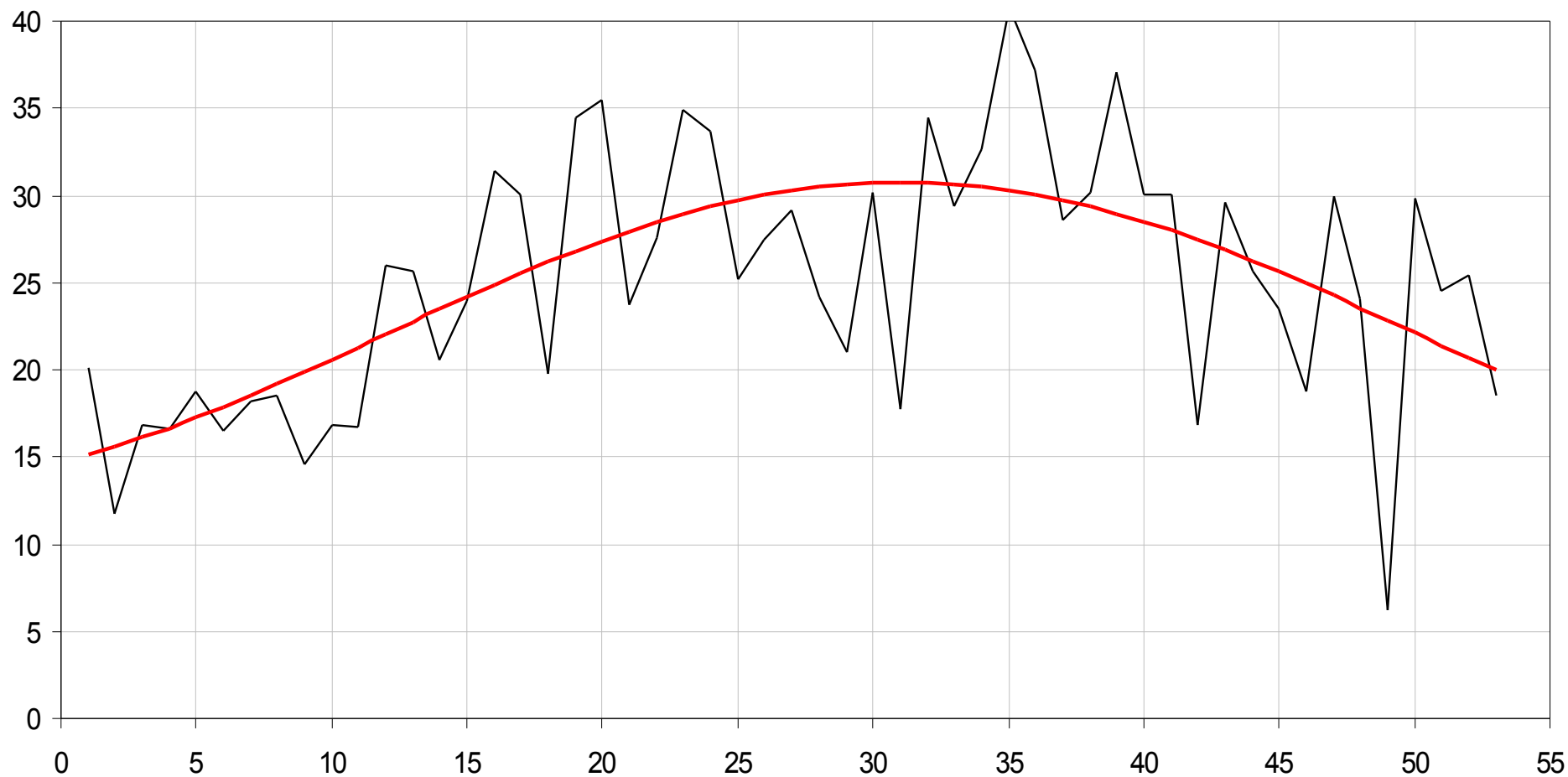
Дослідження циклічності врожайності

- ❑ В.Шоу (1906 р., Англія), Г.Мур (1919 р., США)
- ❑ В.Г.Михайловський, М.І.Семенов, М.Давидович (Росія, 1921-1922 роки)
- ❑ В.П.Тимошенко (1930-і роки, США)
- ❑ М.С.Четвериков (1930-і – 1950-і роки, Росія)
- ❑ А.І.Маннеля (1960-і – 1970-і роки, Росія)
- ❑ О.В.Олійник (2000-і роки, Україна)

Динаміка врожайності озимої пшениці для Херсонської області (1955–2008 роки).

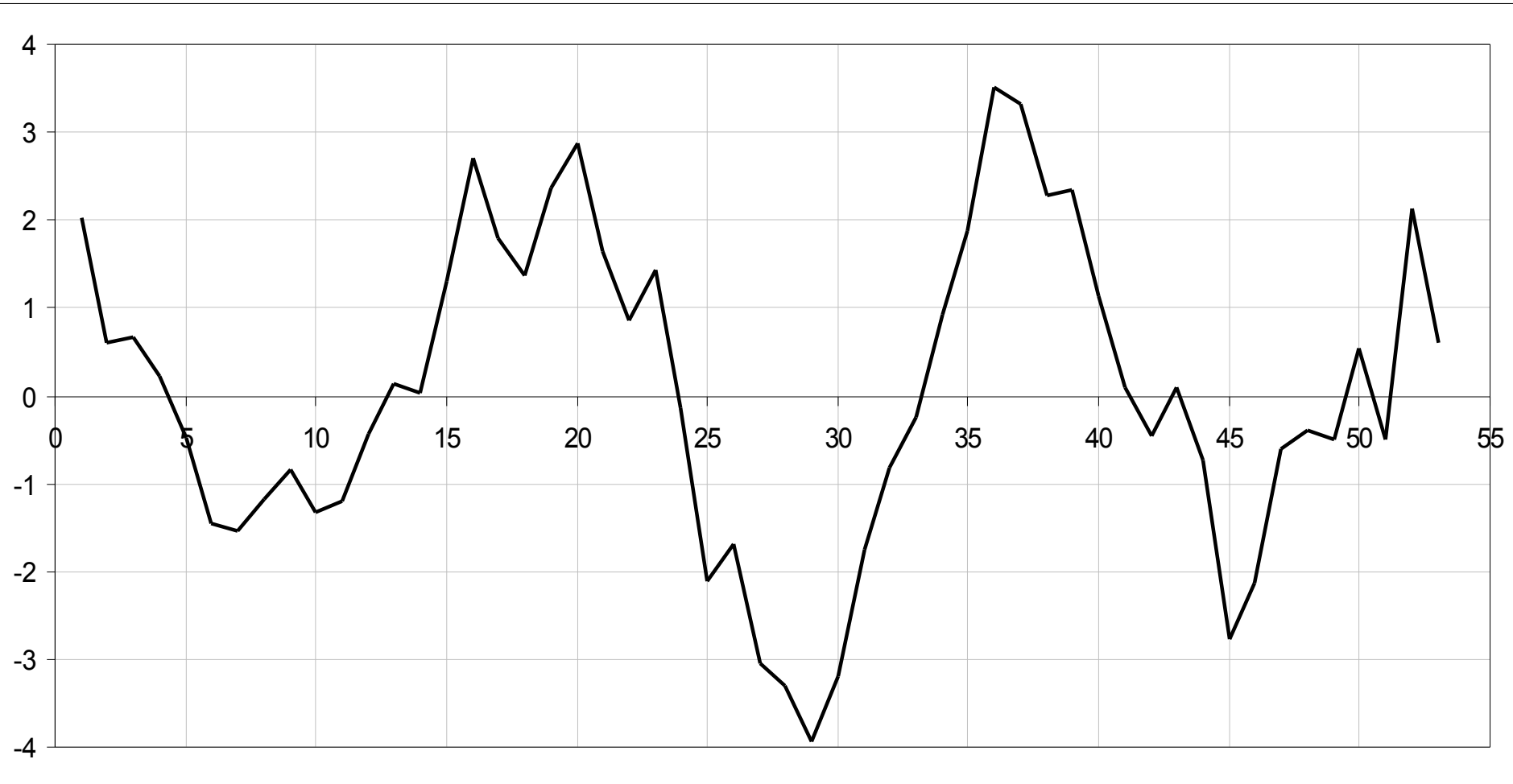


Гармонічний тренд
$$trend = a_0 + a_1 \cos\left(\frac{2\pi}{T_1}t\right) + b_1 \sin\left(\frac{2\pi}{T_1}t\right)$$



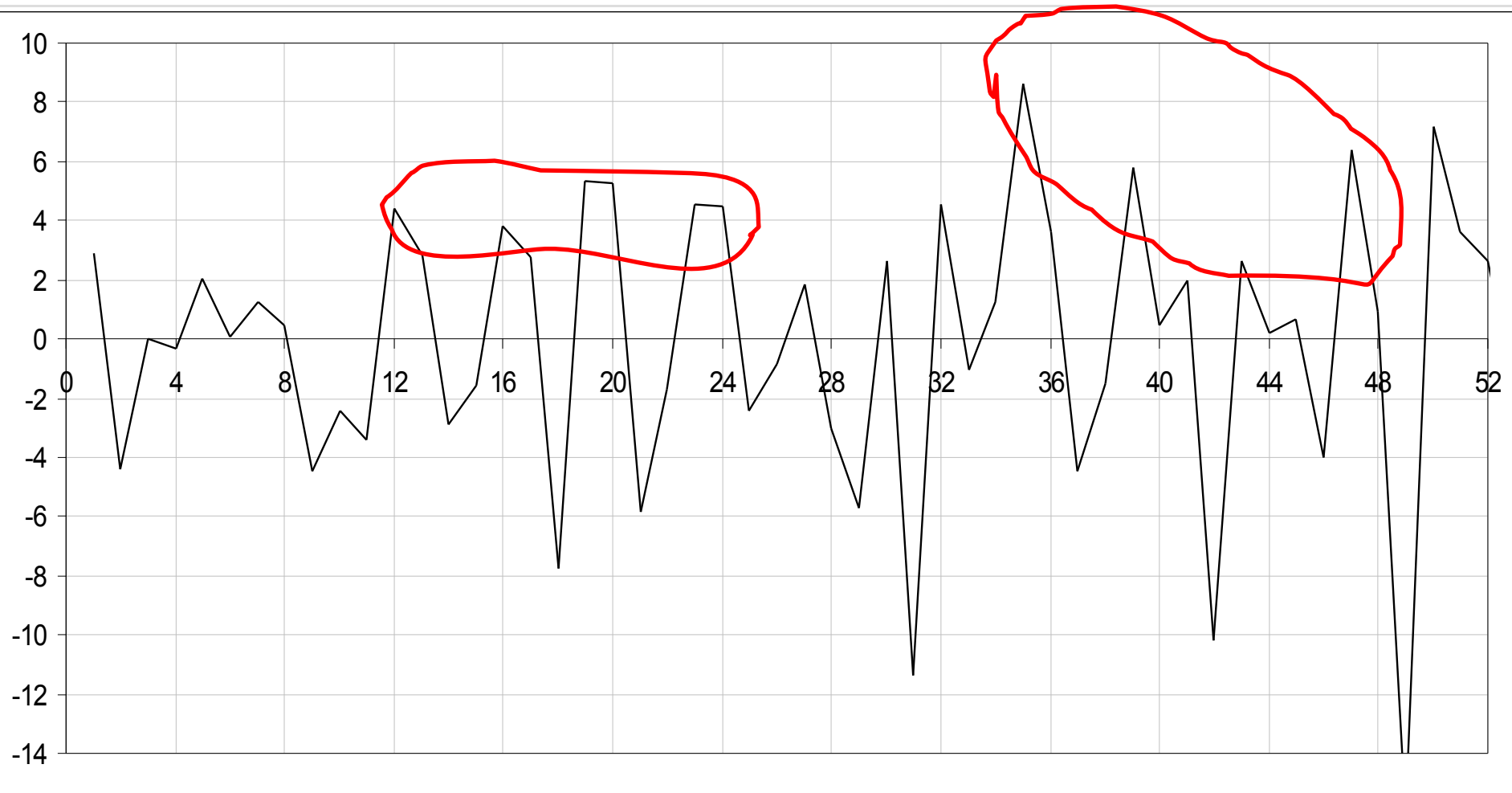


Херсонська область. Згладжений ряд залишків. Спостерігається цикл $T \approx 18$ років





Херсонська область. Ряд других залишків. Спостерігається цикл $T \approx 4$ роки





Полігармонічна модель динаміки врожайності

$$x_t = a_0 + \sum_{i=1}^m a_i \cos\left(\frac{2\pi}{T_i} t\right) + \sum_{i=1}^m b_i \sin\left(\frac{2\pi}{T_i} t\right) + E \quad (1)$$

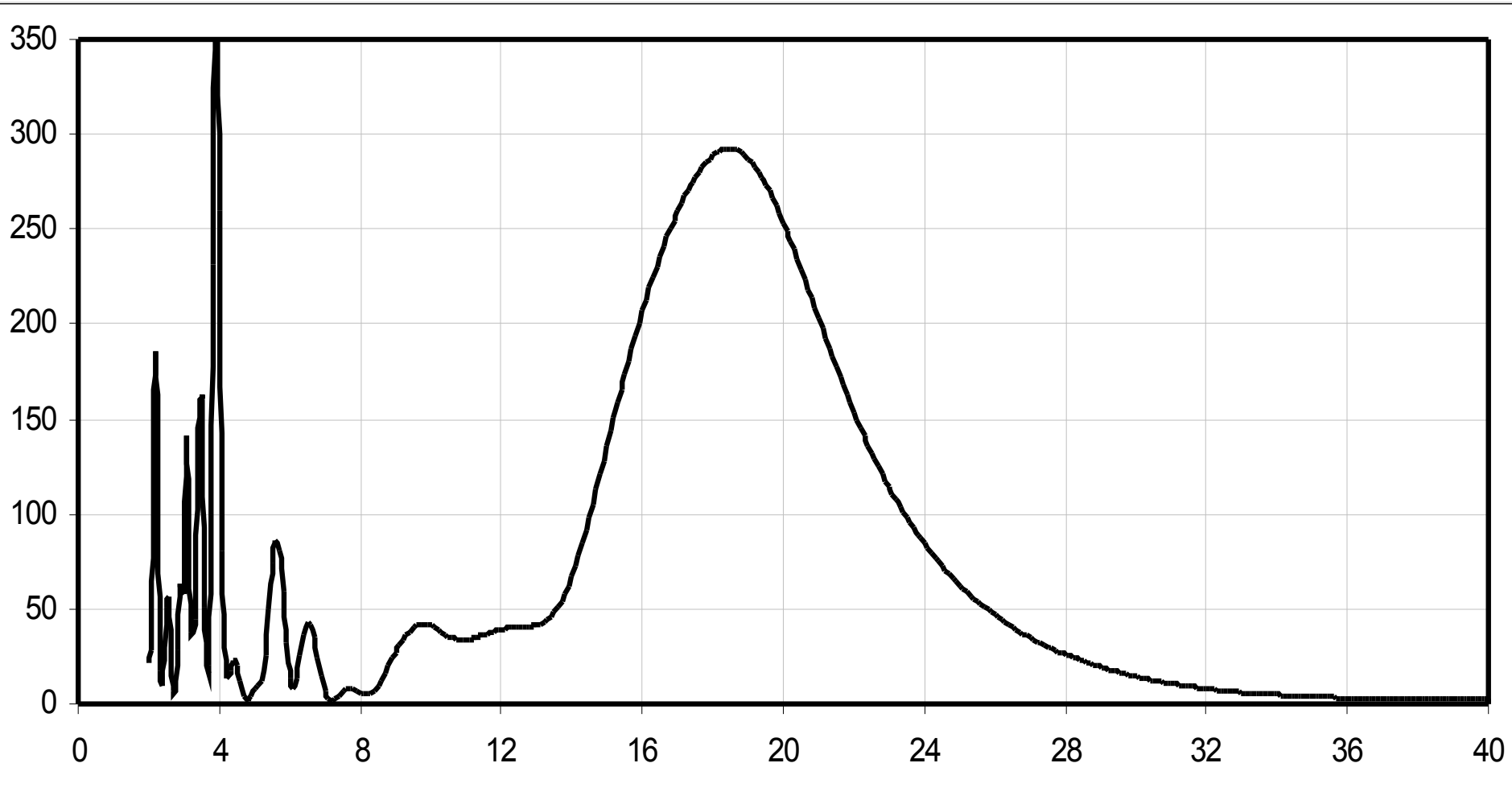
- Параметри моделі поетапно визначалися з мінімуму функціонала

$$\Psi = \sum_{i=1}^m \sum_{t=0}^{T \max} (x_t - a_0 - a_i \cdot \cos\left(\frac{2\pi}{T_i} t\right) - b_i \cdot \sin\left(\frac{2\pi}{T_i} t\right))^2 \rightarrow \min \quad (2)$$

- T – метод повного перебору; a_i, b_i – метод найменших квадратів
- *Перша гармоніка відіграє роль тренду*



Херсонська область. Обернений функціонал похибки гармонічної моделі

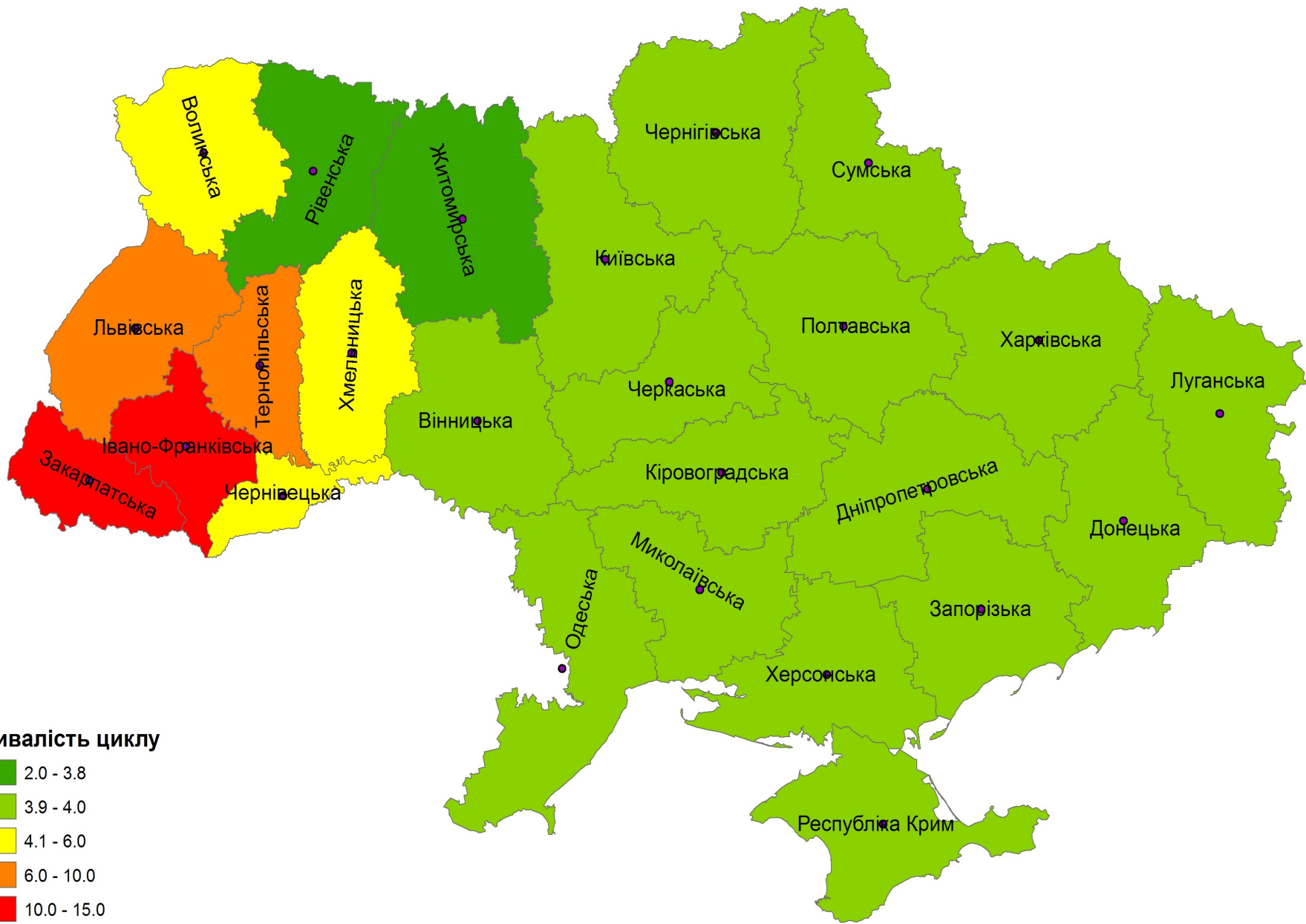


Подвійна циклічність урожайності озимої пшениці



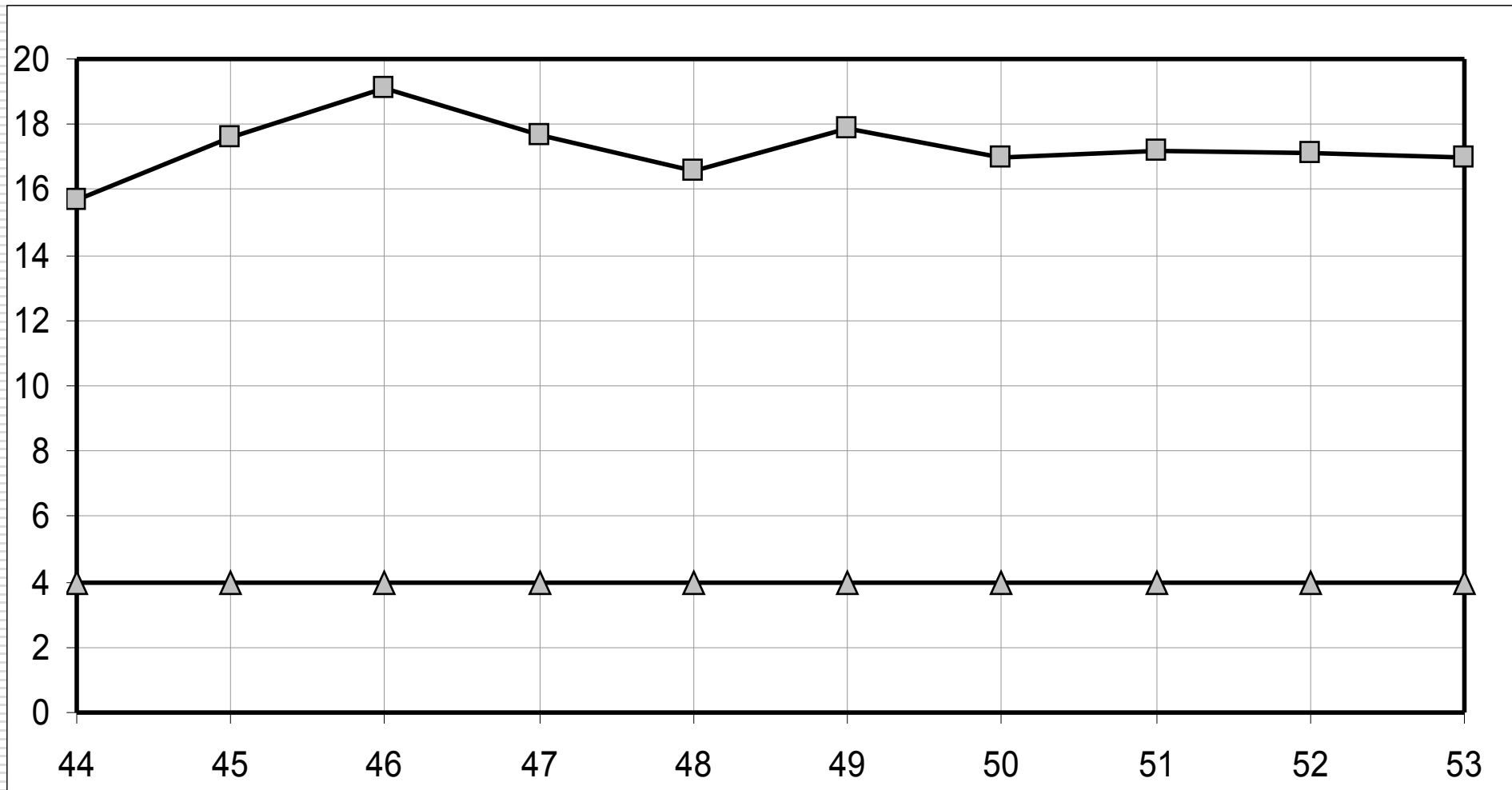
	T2	K1	T3	K2
Луганська	14.9	1.16	3.9	1.15
Донецька	15.8	1.11	4.0	1.12
Харківська	16.1	1.10	3.9	1.11
Запорізька	16.4	1.16	4.0	1.12
Дніпропетровська	16.3	1.10	3.9	1.11
АР Крим	18.4	1.15	3.9	1.15
Сумська	15.7	1.10	4.0	1.09
Полтавська	17.3	1.11	4.0	1.13
Херсонська	18.4	1.16	3.9	1.14
Кіровоградська	17.4	1.14	3.9	1.13
Чернігівська	15.6	1.13	4.0	1.08
Миколаївська	17.5	1.14	3.9	1.16
Черкаська	16.4	1.13	4.0	1.15
Київська	15.6	1.10	3.9	1.11
Одеська	18.6	1.10	3.9	1.10
Вінницька	17.3	1.11	4.0	1.09
Житомирська	15.9	1.09	3.0	1.07
Хмельницька	23.9	1.08	5.3	1.07
Рівненська	18.7	1.15	2.2	1.08
Чернівецька	15.4	1.09	5.7	1.09
Тернопільська	20.4	1.12	9.5	1.10
Волинська	19.3	1.17	5.3	1.08
Івано-Франківська	21.7	1.13	15.6	1.12
Львівська	19.0	1.11	8.7	1.06
Закарпатська	18.7	1.26	14.4	1.08

Тривалість короткого циклу врожайності





Дослідження стійкості циклів (нижня лінія - короткий цикл, верхня лінія - середній цикл)





ВИСНОВКИ 1

- ❑ Дослідження виконані із застосуванням методів ковзного середнього, кореляційного, спектрального та гармонічного аналізу виявили **подвійну циклічність** рядів врожайності озимої пшениці для більшості областей України.
- ❑ Короткий цикл має тривалість ≈ 4 роки, середній цикл має тривалість $\approx 16-18$ років.
- ❑ В західних областях ефект циклічності є слабковираженим



2.Ряди врожайності: детерміновані чи стохастичні?



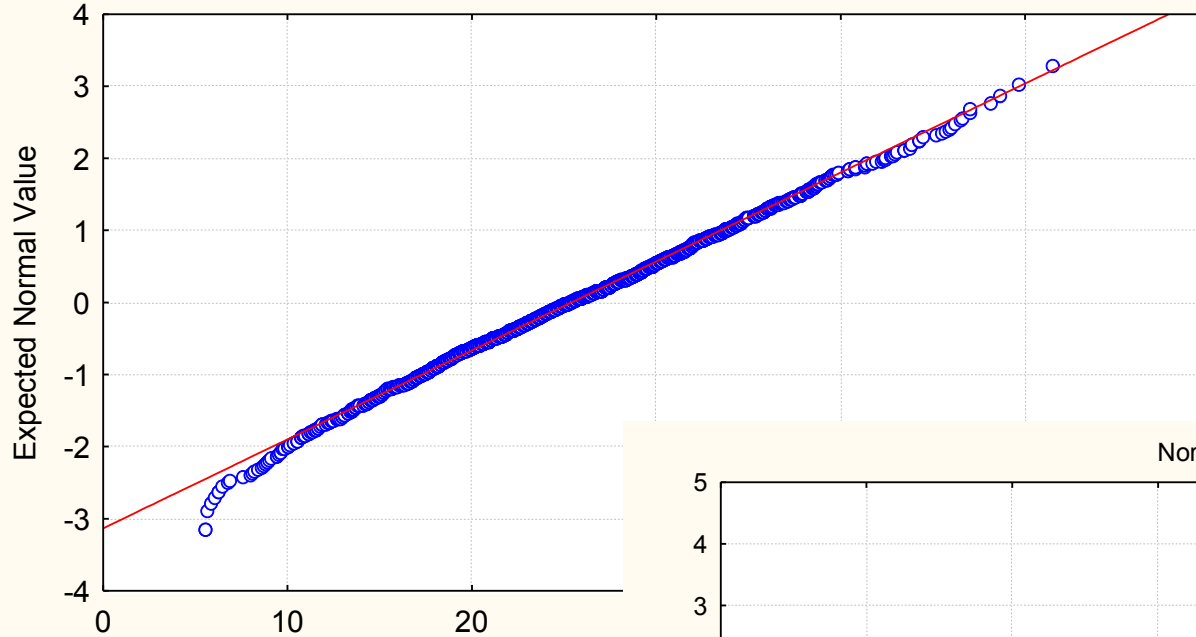
Оцінка ступеня детермінованості часових рядів врожайності

- ❑ Ефект циклічності є доказом присутності детермінованої поведінки динаміки врожайності.
- ❑ Але для різних областей співвідношення “детермінованість-стохастичність” може бути різним
- ❑ Для дослідження цього співвідношення застосуємо статистичні дослідження та методи фрактального аналізу



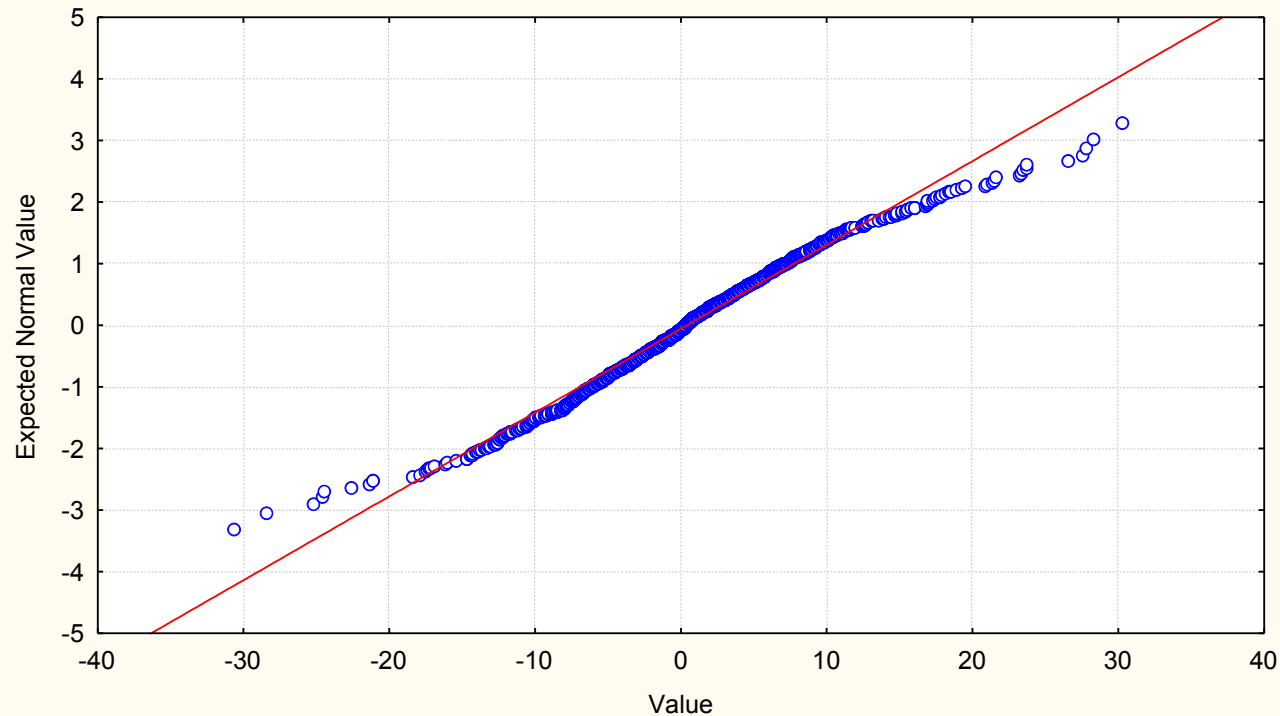
Перевірка гіпотези про нормальний розподіл

Normal P-Plot: Var1



Об'єднаний ряд
врожайностей.
 $25 \cdot 54 = 1350$
елементів. Гіпотеза
не відхиляється

Normal P-Plot: Var1



Об'єднаний ряд різниць
врожайностей. Гіпотеза
відхиляється

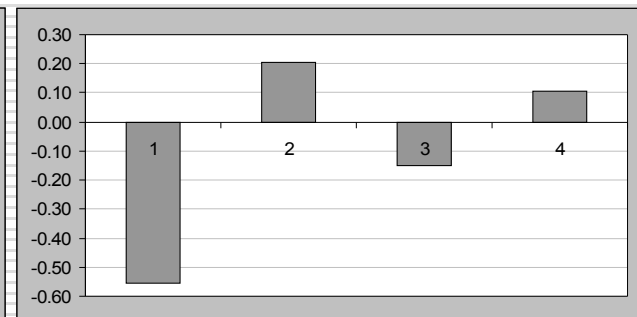
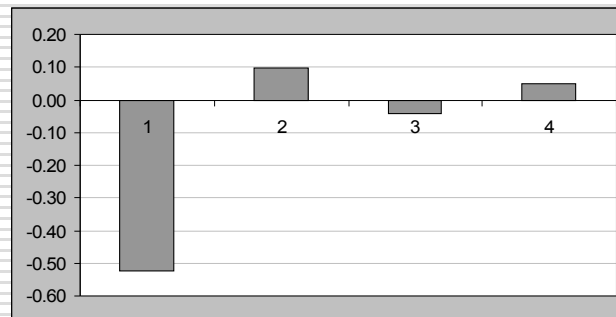
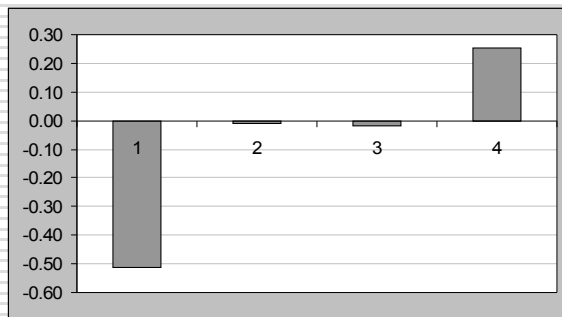
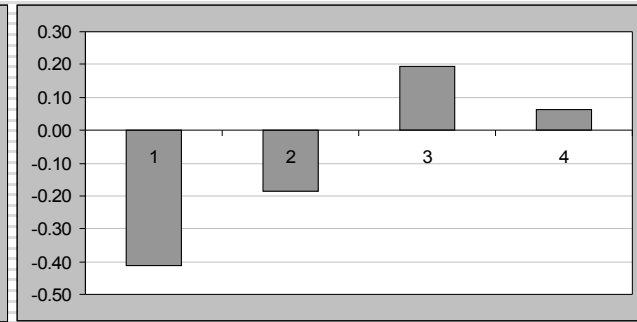
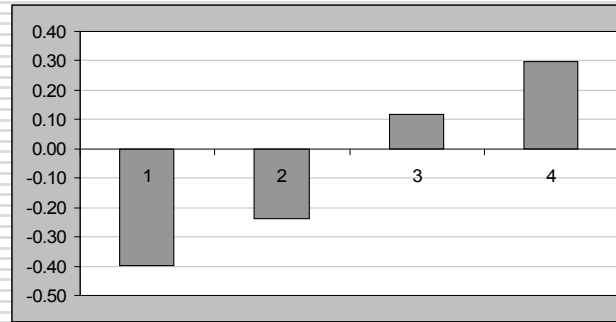
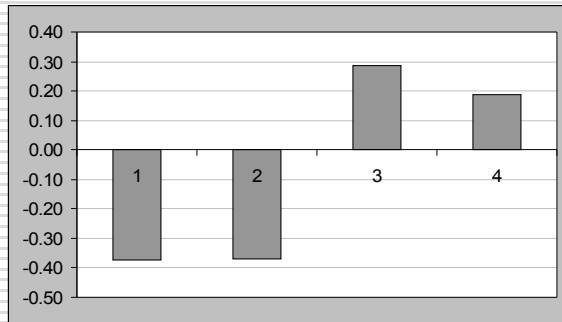


Статистичний аналіз різницевих серій

- Перейдемо від ряду врожайності до **ряду перших різниць**
Проведемо **бінарне кодування** різниць: додатна різниця – символ “1”, від’ємна різниця – символ “0”. В результаті отримаємо ряд виду (10010111000101101001...)
Він складається з різницевих серій різної довжини: 1+00+1+0+111+000+1+0+11... Для **випадкового ряду з незалежними приростами** частота серій повинна залежати від довжини серії за законом $p=1/2^n$ (0.5, 0.25, 0.125, ...)
- *Фактична частота серій одиничної довжини $p_1=0.61$.* Це означає, що після приросту імовірність наступного спаду 61%, імовірність ще одного приросту – 39%
- Це є свідченням **антиперсистентної** поведінки ряду (частота коливань є вищою, ніж у випадкового ряду – **ознака детермінованості**)



Автокореляційна функція ряду перших різниць врожайності для різних областей



Від'ємне значення $AR(1)$ підтверджує властивість антиперсистентності рядів врожайності



R/S – аналіз

- $(R/S)_N = c \cdot N^H$ (3)
 R – розмах коливань акумульованого ряду y_k , S – стандартне відхилення ряду, R/S – нормований розмах, N – часовий інтервал, H – показник Херста

$$y_k = \sum_{i=1}^k (x_i - \bar{x}); \quad k = 1, \dots, N \quad (4)$$

- **Степенева залежність нормованого розмаху R/S від часового інтервалу N є характерною для фракталів. Часові ряди, для яких виконується співвідношення (3) називають фрактальними**

$$\log(R/S)_N = \log(c) + H \cdot \log(N) \quad (5)$$

Hurst H. E. Long-term Storage of Reservoirs // Transactions of the American Society of Civil Engineers. - 1951. - Vol. 116. - pp. 776.

Мандельброт Б. Фрактальная геометрия природы. – М.: «Институт компьютерных исследований», 2002. - 656 с.



Вимоги до R/S - аналізу

- Ряд повинен бути **стаціонарним і незалежним** (відсутні автокореляції)
- Довжина ряду не менше 10^3

Досягнення стаціонарності

- Фінансові ряди:** нестаціонарні, геометричний ріст, сильна автокореляція. **Ряд прибутковостей** - $y_i = \log(x_{i+1} / x_i)$ - стаціонарний, незалежний
- Природні ряди:** нестаціонарні, арифметичний ріст, помітна автокореляція
Ряд різниць $y_i = x_{i+1} - x_i$ - стаціонарний, незалежний

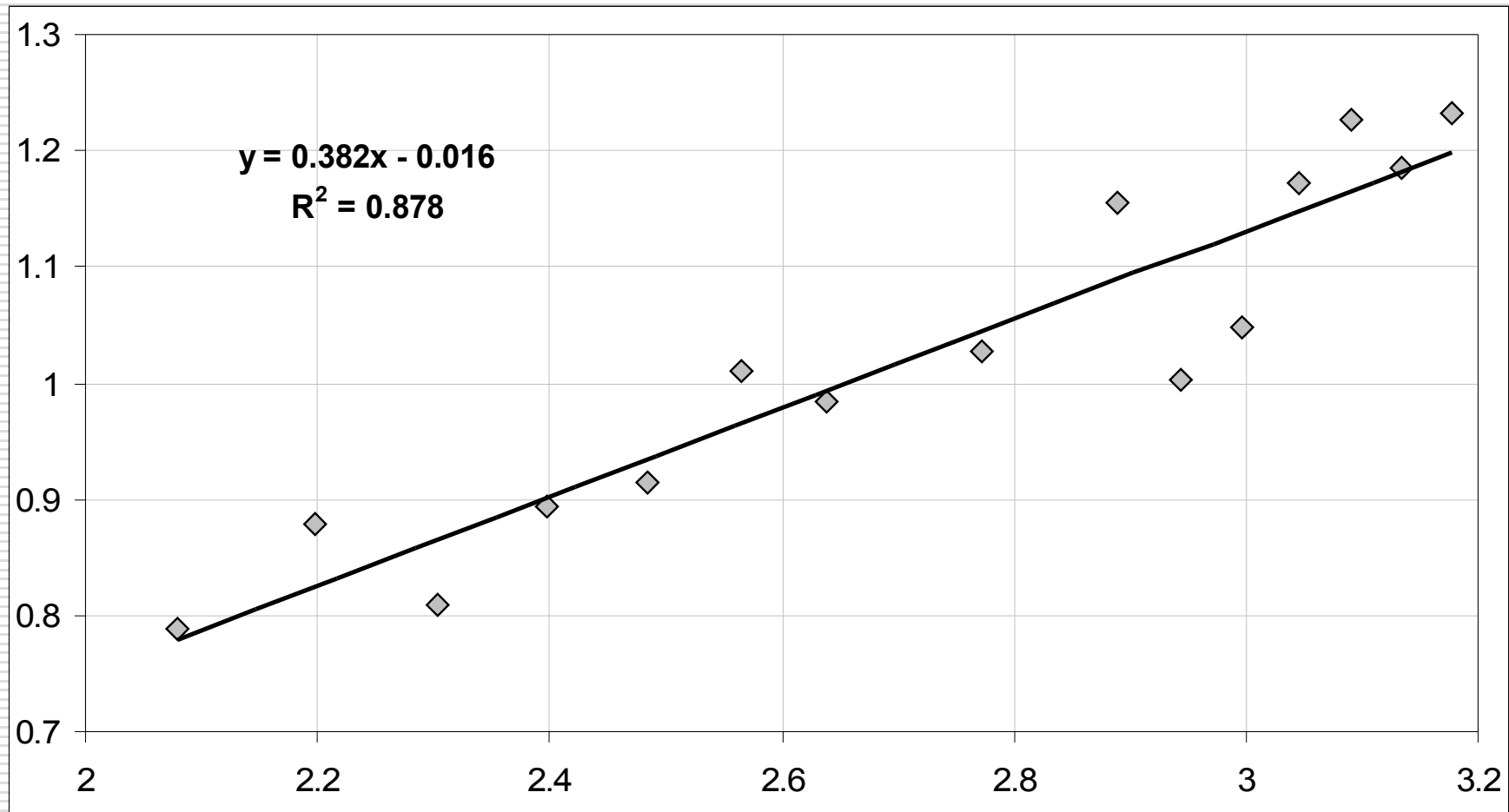


Метод ковзаючого вікна

- ❑ Підхід Петерса: вибирати фрагмент ряду довжини n від $n = 10$ до $n = N \text{ div } 2$, так щоб n було дільником довжини ряду N
- ❑ Підхід автора: **метод ковзаючого n -вікна** – досліджуваний фрагмент ковзає вздовж часового ряду: 1 крок: x_1, x_2, \dots, x_n ; 2 крок: x_2, x_3, \dots, x_{n+1} ; ... $k+1$ крок: $x_{nk}, x_{2k}, \dots, x_{nk}$. Отримані кумулятивні суми усереднюються
- ❑ **Метод ковзаючого вікна** розширює статистичну базу досліджень і дозволяє розраховувати показник Херста H для рядів, які містять $\sim 10^2$ елементів



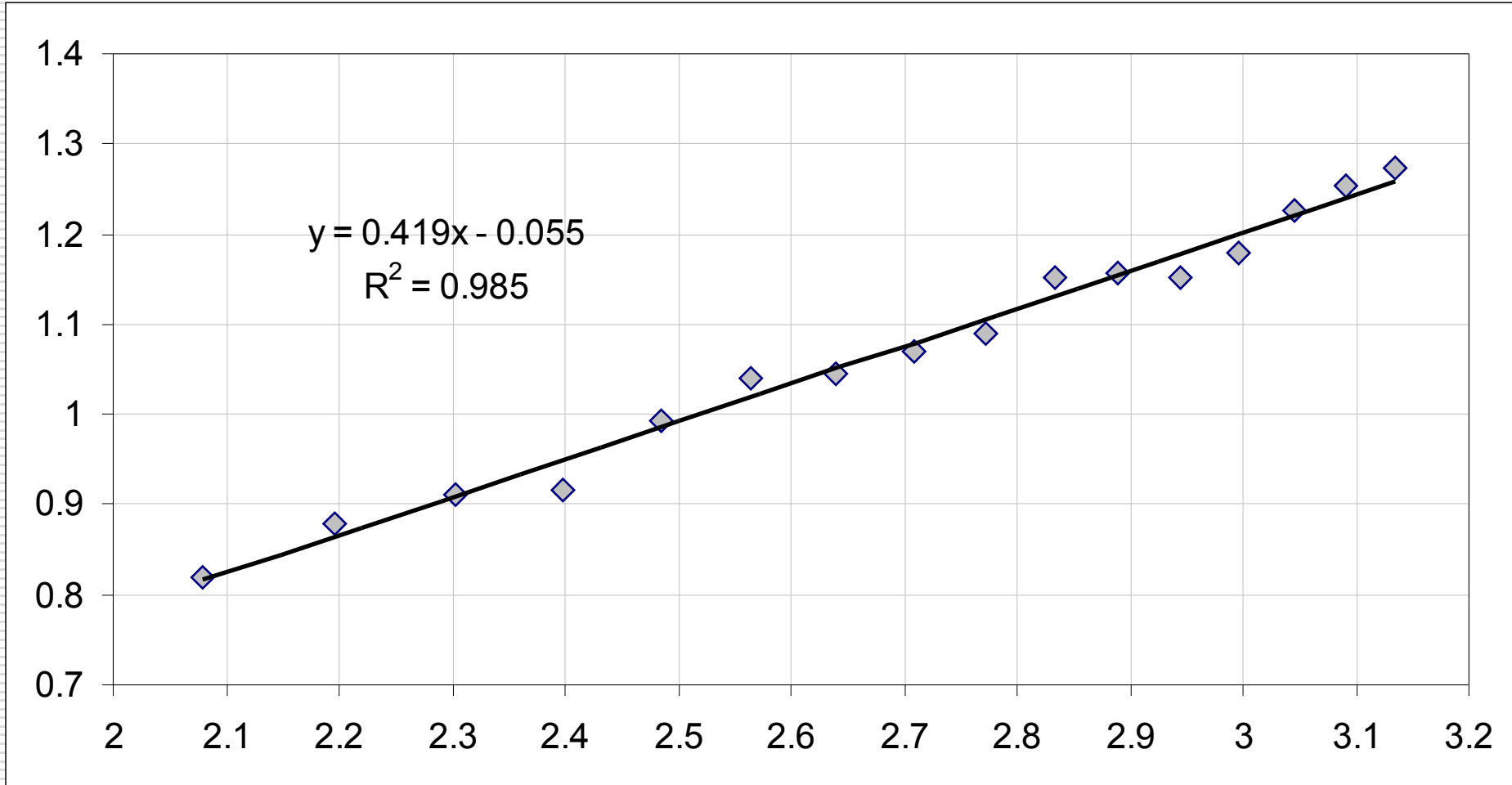
Метод ковзаючого вікна. АР Крим: 1-ий крок



Вибираючи фрагмент часового ряду довжиною $n=48$ і послідовно пересуваючи його вздовж часового ряду врожайності ми отримуємо наступні значення показника Херста: 0.38, 0.34, 0.40, 0.40, 0.44, 0.45, 0.43



АР Крим. Метод ковзаючого вікна. Усереднення кумулятивних сум



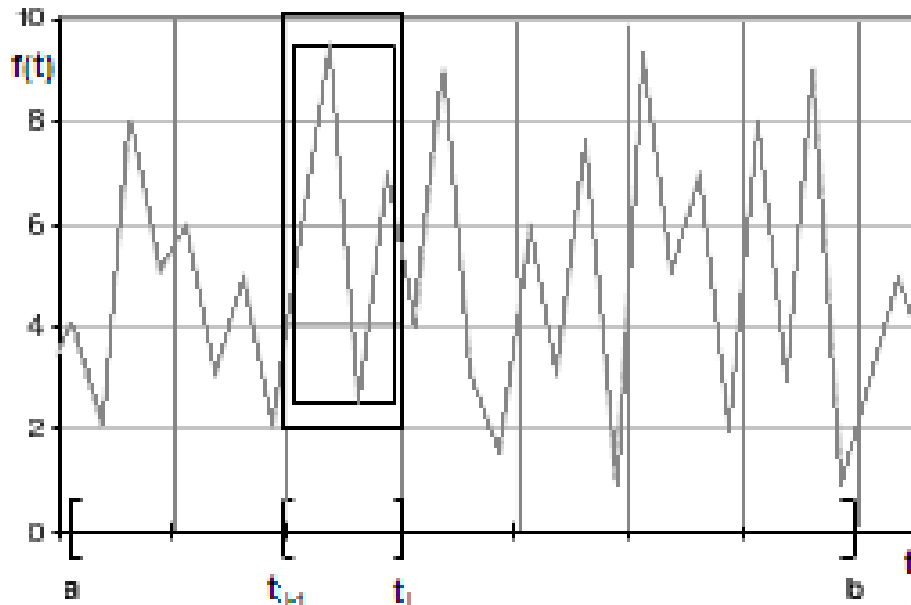
$H = 0.419 \pm 0.040$; $\alpha = 0,95$; $R^2 = .985$



Класифікація процесів

- ❑ $0.45 \leq H \leq 0.55$ – випадковий процес
- ❑ $H > 0.55$ – персистентний (трендостійкий) процес
- ❑ $H < 0.45$ – антиперсистентний процес (процес з підвищеною коливністю)
- ❑ Розрахунки коефіцієнта Херста для більшості областей (21 з 25) дали значення $H < 0.45$ – **антиперсистентний** процес

Фрактальний аналіз часових рядів



*M.M. Dubovikov, N.V. Starchenko,
M.S. Dubovikov. Dimension of the
minimal cover and fractal analysis of
time series // Physica A. - v339.
-2004. - pp591-608*

- ❑ Введемо поняття індекса фрактальності μ $\mu = D_H - D_T$ (6)
- ❑ Для обчислення індекса μ використаємо n послідовно вкладених відрізків

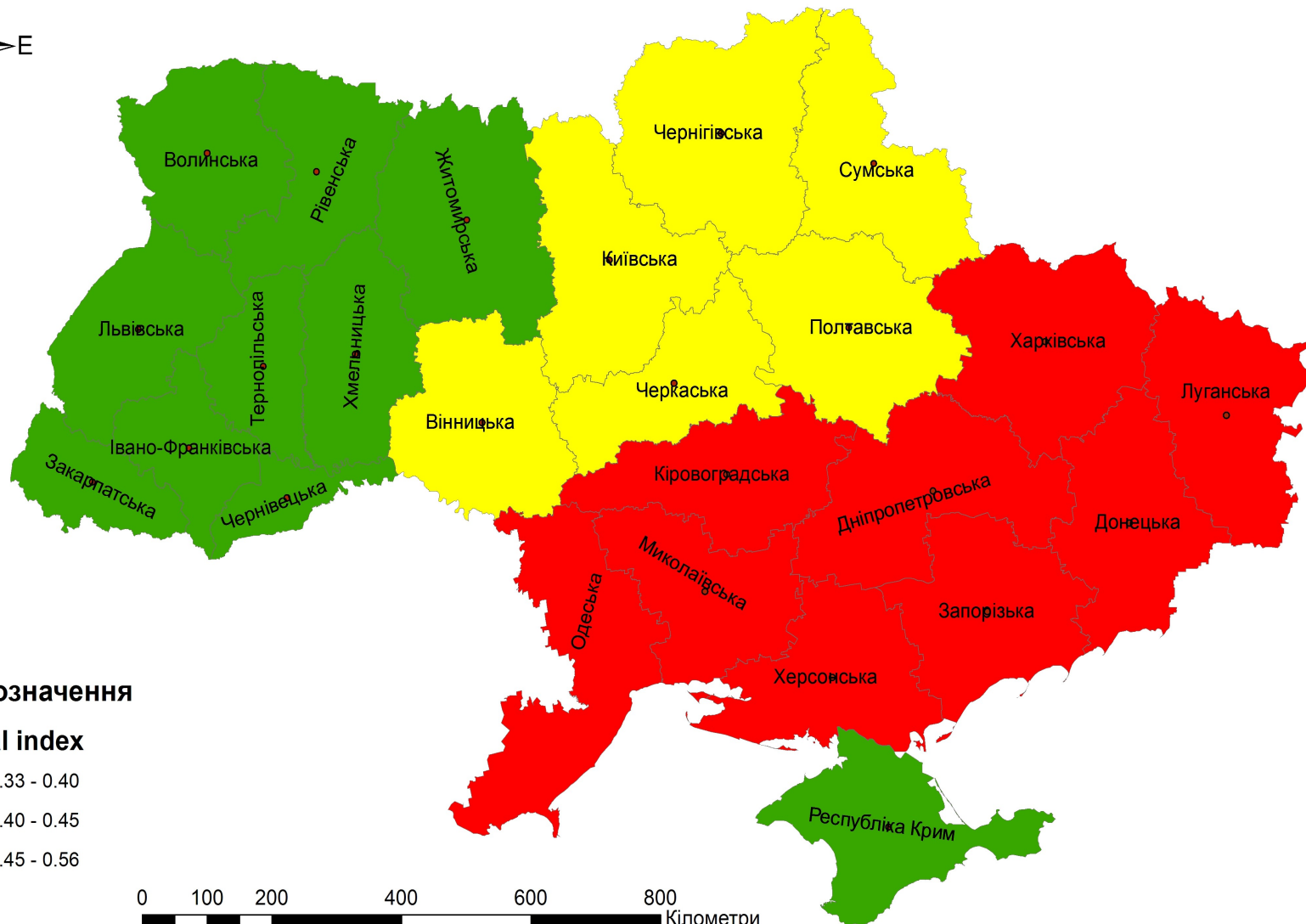
$$\omega_m = [a = t_0 < t_1 < \dots < t_m = b], \quad m = 2^n; n = 0, 1, 2, \dots$$
 (7)
- ❑ D_H - фрактальна розмірність Хаусдорфа $D_H = \lim_{\delta \rightarrow 0} [\ln N(\delta) / \ln(1/\delta)]$ (8)
- ❑ D_T - топологічна розмірність
- ❑ Для гауссових рядів $H_f = 1 - \mu$ співпадає з коефіцієнтом Херста H
- ❑ Вдало побудована передфрактальна модель (мінімальні покриваючі прямокутники) дає можливість розрахунку H_f вже при $N \geq 32$

Коефіцієнт Херста і коефіцієнт фрактальності рядів врожайності озимої пшениці для областей України



<i>Область</i>	<i>H</i>	<i>H_μ</i>	<i>Область</i>	<i>H</i>	<i>H_μ</i>
Луганська	0.46	0.34	Київська	0.41	0.44
Донецька	0.43	0.36	Одеська	0.38	0.36
Харківська	0.39	0.33	Вінницька	0.41	0.45
Запорізька	0.43	0.35	Житомирська	0.41	0.47
Дніпропетровська	0.36	0.35	Хмельницька	0.43	0.50
АР Крим	0.42	0.46	Рівненська	0.38	0.49
Сумська	0.42	0.41	Чернівецька	0.41	0.46
Полтавська	0.43	0.41	Тернопільська	0.43	0.54
Херсонська	0.38	0.37	Волинська	0.38	0.49
Кіровоградська	0.52	0.39	Івано-Франківська	0.44	0.49
Чернігівська	0.43	0.43	Львівська	0.48	0.49
Миколаївська	0.48	0.40	Закарпатська	0.52	0.55
Черкаська	0.41	0.45			

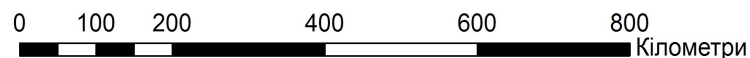
Значення фрактального коефіцієнта рядів врожайності озимої пшениці



Умовні позначення

Fractal index

- 0.33 - 0.40
- 0.40 - 0.45
- 0.45 - 0.56



1:6 000 000



ВИСНОВКИ 2

- ❑ *Статистичні дослідження та методи фрактального аналізу показали, що ряди врожайності для більшості областей є антиперсистентними*
- ❑ *Ступінь детермінованості має яскраво виражений географічний тренд, зменшуючись за напрямком “схід-захід”*



Антиперсистентність інших природних рядів

- ❑ Ряд урожайності зернових в Росії за період 1900 – 2006 роки: **$H = 0.42$**
- ❑ Ряд урожайності озимої пшениці в США за період 1866 – 2007 роки : **$H = 0.44$**
- ❑ Річні дані про рівень річки Ніл за 622 – 1470 роки : **$H = 0.38$**
- ❑ Місячні дані про кількість сонячних плям за 1749 – 1937 роки : **$H = 0.43$**
- ❑ **Антиперсистентність природних рядів є відображенням загального принципу гомеостазу (принцип рівноваги природних систем)**



3. Методи прогнозування рядів врожайності



1. Покращена полігармонічна модель

- ❑ Враховуючи стійку циклічну природу динаміки врожайності, застосуємо для прогнозування удосконалену полігармонічну модель ($m = 3$)

$$x_t = a_0 + \sum_{i=1}^3 a_i \cos\left(\frac{2\pi}{T_i} t\right) + \sum_{i=1}^3 b_i \sin\left(\frac{2\pi}{T_i} t\right) \quad (9)$$

- ❑ Суть удосконалення:
 - а) модель будується для ряду перших різниць;
 - б) застосовується попереднє згладжування ряду різниць
- ❑ Модель розроблена автором

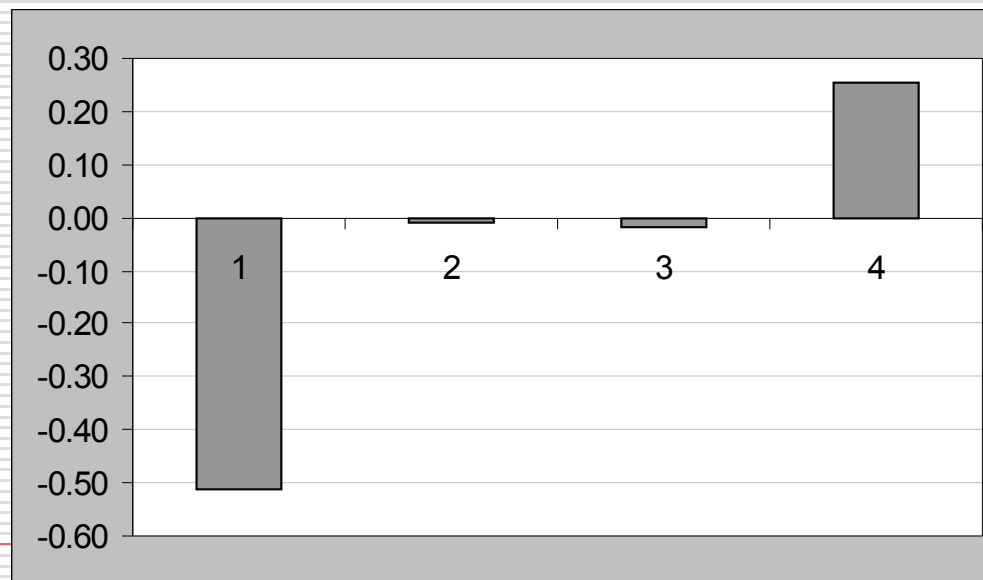


2. Авторегресійна модель для різниць

- За результатами автокореляційного аналізу ряду перших різниць будується модель AR4

$$\Delta y_{t+1} = a_0 + a_1 \Delta y_t + a_2 \Delta y_{t-1} + a_3 \Delta y_{t-2} + a_4 \Delta y_{t-3} \quad (10)$$

- Деякі коефіцієнти моделі можуть дорівнювати нулю





3. Статистичний аналіз різницевих серій

- ❑ В результаті аналізу статистики ряду різниць визначається **фактична частота серій** (одинарної, подвійної, потрійної)
 - ❑ На підставі аналізу кінцевого фрагмента ряду будується **імовірнісний прогноз**
 - ❑ **Приклад:** $p_1 = 0.60$. Кінцевий фрагмент ряду "...01". Ймовірність приросту – 40%, ймовірність спаду – 60%. Середнє значення приросту 5.8, середнє значення спаду – 5.2.
- Прогнозований приріст:** $5.8 \cdot 0.4 + (-5.2) \cdot 0.6 = -0.8$
- ❑ Модель розроблена автором



4. Метод найближчих сусідів

- Використовуючи методику Паккарда, відновлюємо **фазовий портрет** динамічної системи у лаговому просторі за даними часового ряду. Він складається із векторів (x_1, x_2, x_3) , (x_2, x_3, x_4) , (x_3, x_4, x_5) , ... (для розмірності простору $D=3$)
- Метод найближчих сусідів ґрунтується на ідеї про те, що **близькі фазові вектори** на короткому відрізку часу **еволюціонують однаково**. Для того, щоб оцінити зміну фазового вектора y_k , необхідно знайти m найближчих до нього векторів (найближчих сусідів). Найпростіший вектор-прогноз має наступний вигляд

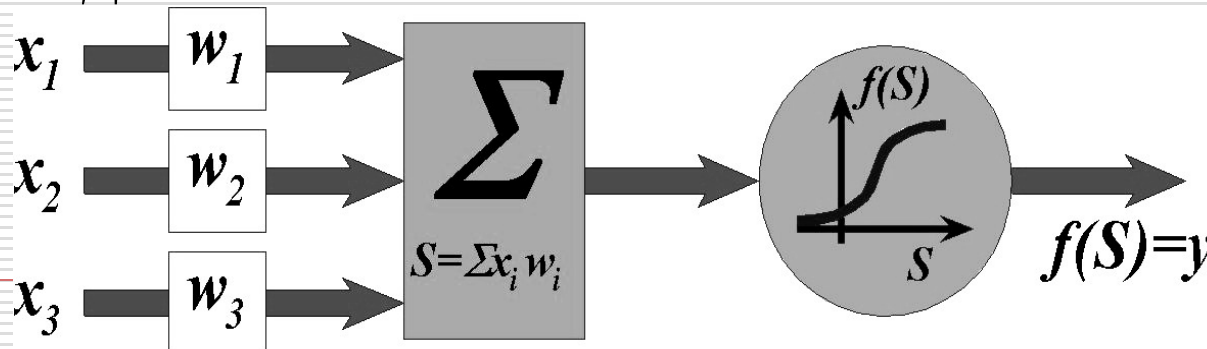
$$y_{k+1} = \frac{1}{m} (y_{n_1+1} + y_{n_2+1} + \dots + y_{n_m+1}) \quad (11)$$

5. Метод нейронних мереж

У прогностичних розрахунках був використаний метод зворотного поширення помилки, сигмоїдна активаційна функція виду

$$y = \frac{e^{\lambda V} - 1}{e^{\lambda V} + 1} \quad (12)$$

та вхідне вікно розміром три роки. Значення врожайності за четвертий рік служило виходом мережі. Таким чином, була використана двохшарова мережа з трьома входами і одним виходом





6. Метод Бокса-Дженкінса

- Метод авторегресії – ковзаючого середнього $ARIMA(p,q,d)$, запропонований Боксом і Дженкінсом, широко використовується для прогнозування часових рядів

$$Y'_t = \alpha_0 + \alpha_1 Y_{t-1} + \alpha_2 Y_{t-2} + \dots + B_1 \varepsilon_{t-1} + B_2 \varepsilon_{t-2} + \dots + u_t \quad (13)$$

$$\varepsilon_i = Y_i - Y'_i$$

Бокс Дж., Дженкинс Г. Анализ временных рядов. Прогноз и управление. – М.: Мир, 1974. – 608с.



4. Методи оцінки прогностичних моделей

- а) Середня похибка*
- б) Дисперсія похибки*
- в) Коефіцієнт справджуваності знаку приростів*



Оцінка прогностичних моделей

(результат апробації на 25 рядах врожайності озимої пшениці для областей України)

	<i>ПГМ</i>	<i>АРМ</i>	<i>МАРС</i>	<i>МНС</i>	<i>НМ</i>	<i>ARIMA</i>
<i>Середня похибка</i>	0.196	0.195	0.220	0.230	0.219	0.247
<i>Дисперсія похибки</i>	0.41	0.49	0.57	0.62	0.55	0.66
<i>Коефіцієнт справджуваності</i>	80%	78%	71%	68%	71%	68%

ПГМ – полігармонічна модель; **АРМ** – авторегресійна модель;

МАРС – метод аналізу різницевих серій;

МНС – метод найближчих сусідів; **НМ** – метод нейронних мереж;

ARIMA – метод Бокса - Дженкінса

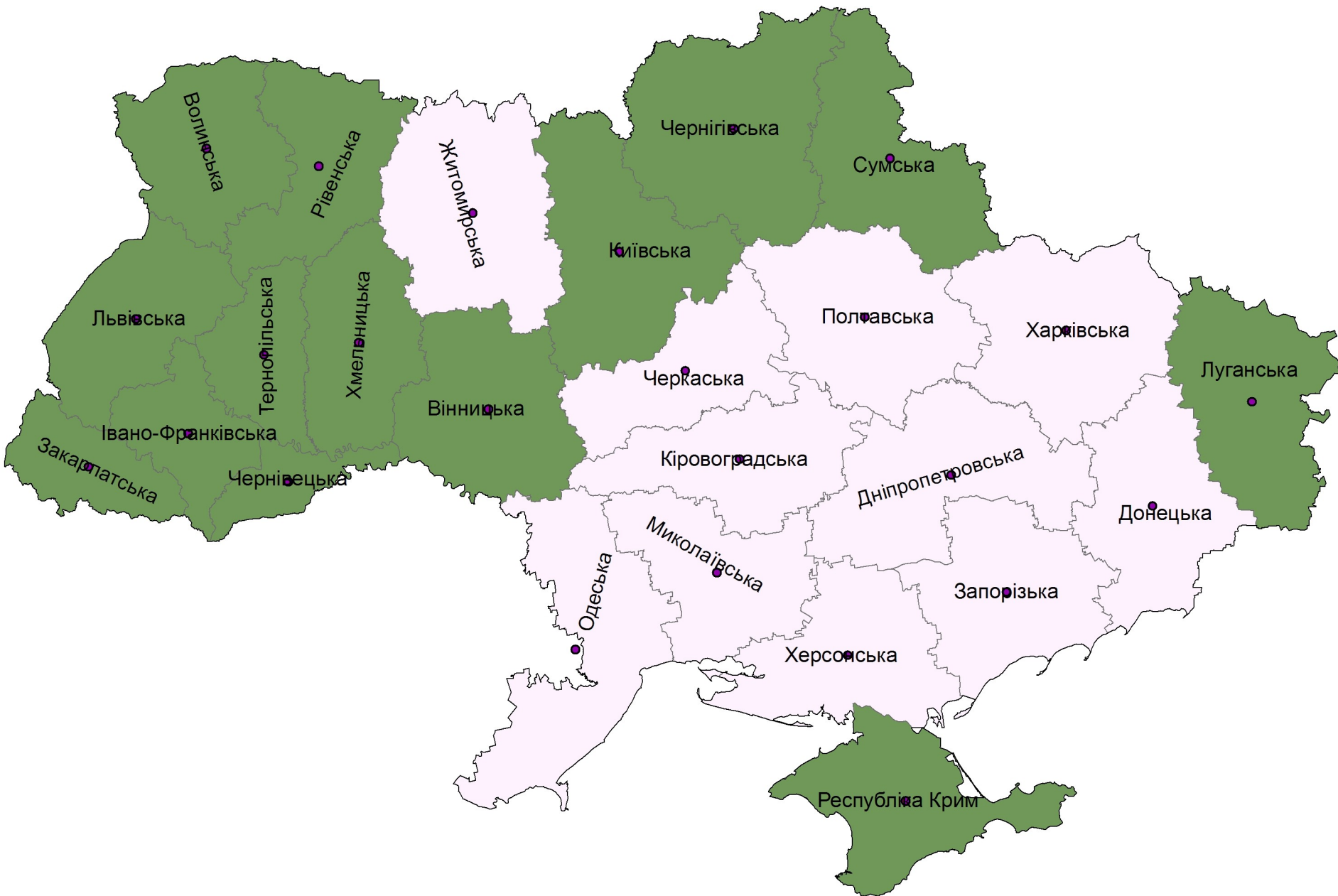
Висновок: Найбільш ефективними прогностичними моделями врожайності є **полігармонічна модель** та **авторегресійна модель**

Похибка прогнозування для окремих областей



		ПГМ	АРМ
1	Луганська	30.6%	30.4%
2	Донецька	27.8%	22.9%
3	Харківська	29.5%	26.9%
4	Запорізька	24.1%	19.9%
5	Дніпропетровська	32.4%	26.7%
6	АР Крим	8.0%	10.6%
7	Сумська	21.3%	23.6%
8	Полтавська	27.7%	27.4%
9	Херсонська	22.5%	19.1%
10	Кіровоградська	29.8%	27.0%
11	Чернігівська	14.6%	18.9%
12	Миколаївська	31.0%	28.3%
13	Черкаська	23.0%	21.0%
14	Київська	16.5%	16.4%
15	Одеська	25.5%	20.4%
16	Вінницька	16.1%	18.5%
17	Житомирська	13.1%	12.1%
18	Хмельницька	16.7%	19.3%
19	Рівненська	13.3%	14.7%
20	Чернівецька	11.9%	21.2%
21	Тернопільська	14.1%	17.8%
22	Волинська	10.5%	11.0%
23	Івано-Франківська	11.3%	11.4%
24	Львівська	7.9%	9.2%
25	Закарпатська	11.2%	13.7%

Порівняння точності гармонічної та авторегресійної моделей





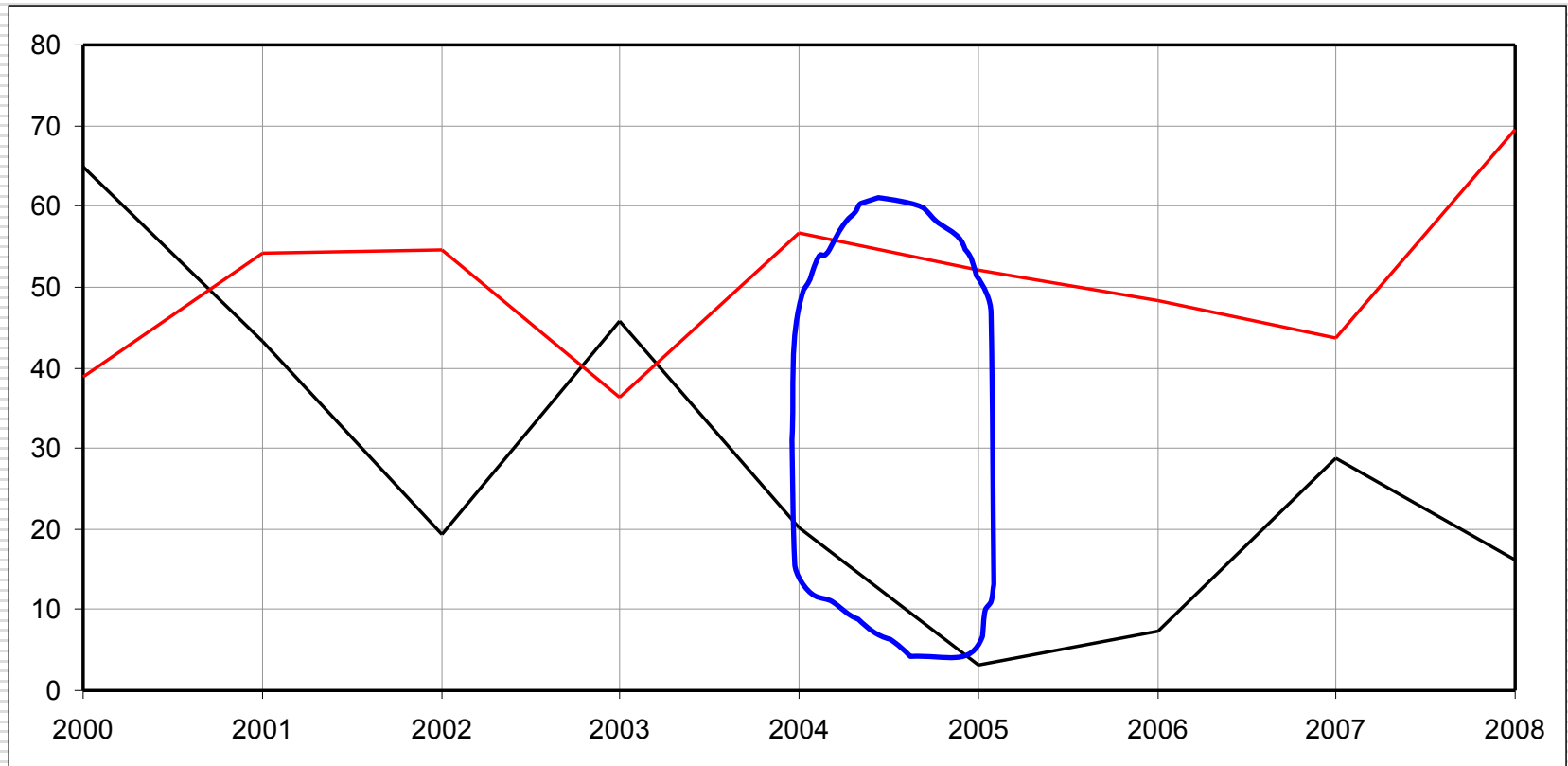
5.Методика оцінювання ризиків зерновиробництва

- а) Модель “врожайність - рентабельність”*
- б) Прогноз врожайності*
- в) Прогноз рентабельності*

Дослідження зв'язку між врожайністю та рентабельністю зернових в Україні



— валовий збір x 2 — рентабельність

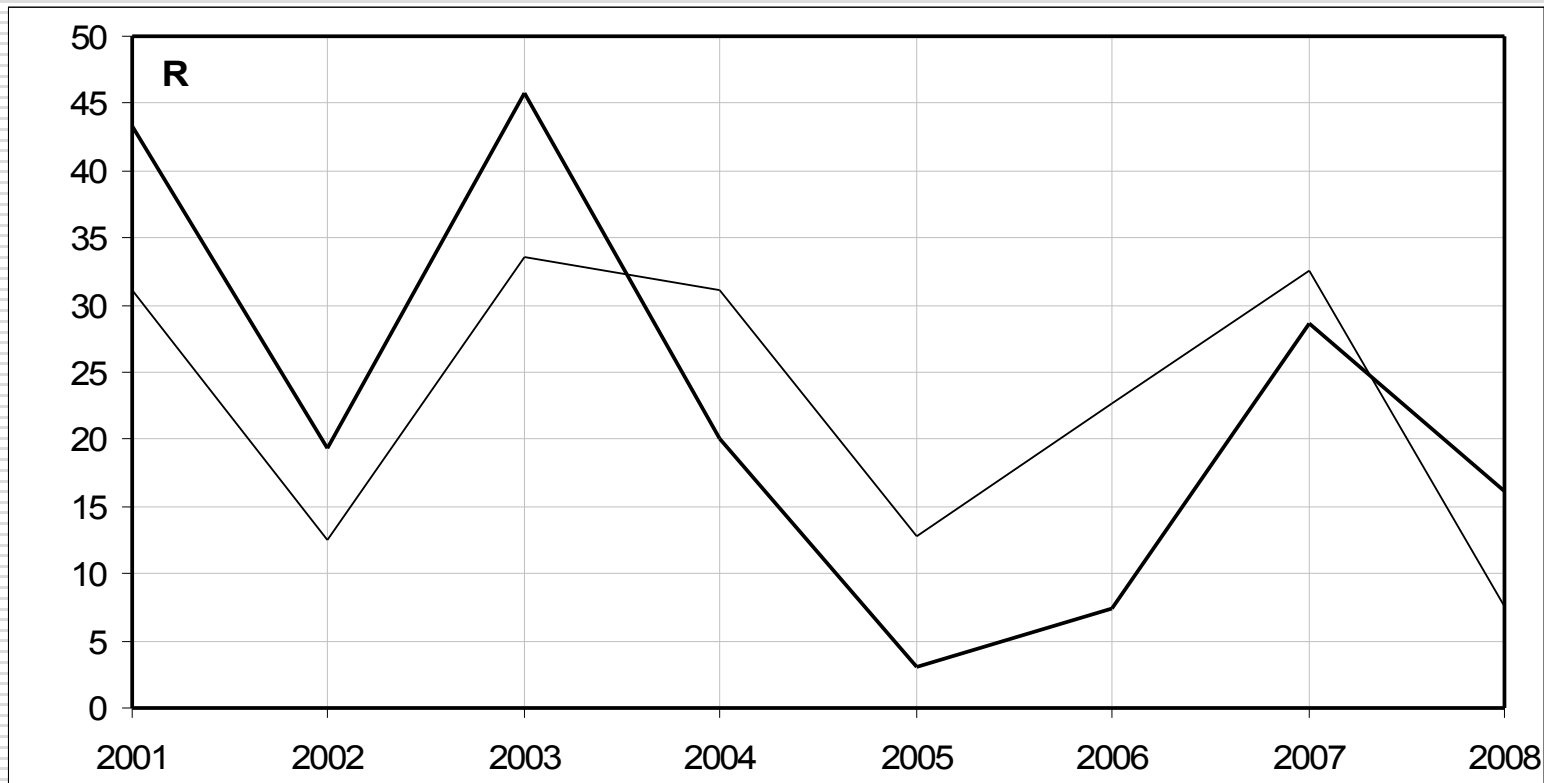




Моделювання рентабельності зернових

R – рентабельність, **W** – валовий збір зернових

$$R_i = a_0 + a_1 \cdot (W_i + W_{i-1}) \quad (14)$$



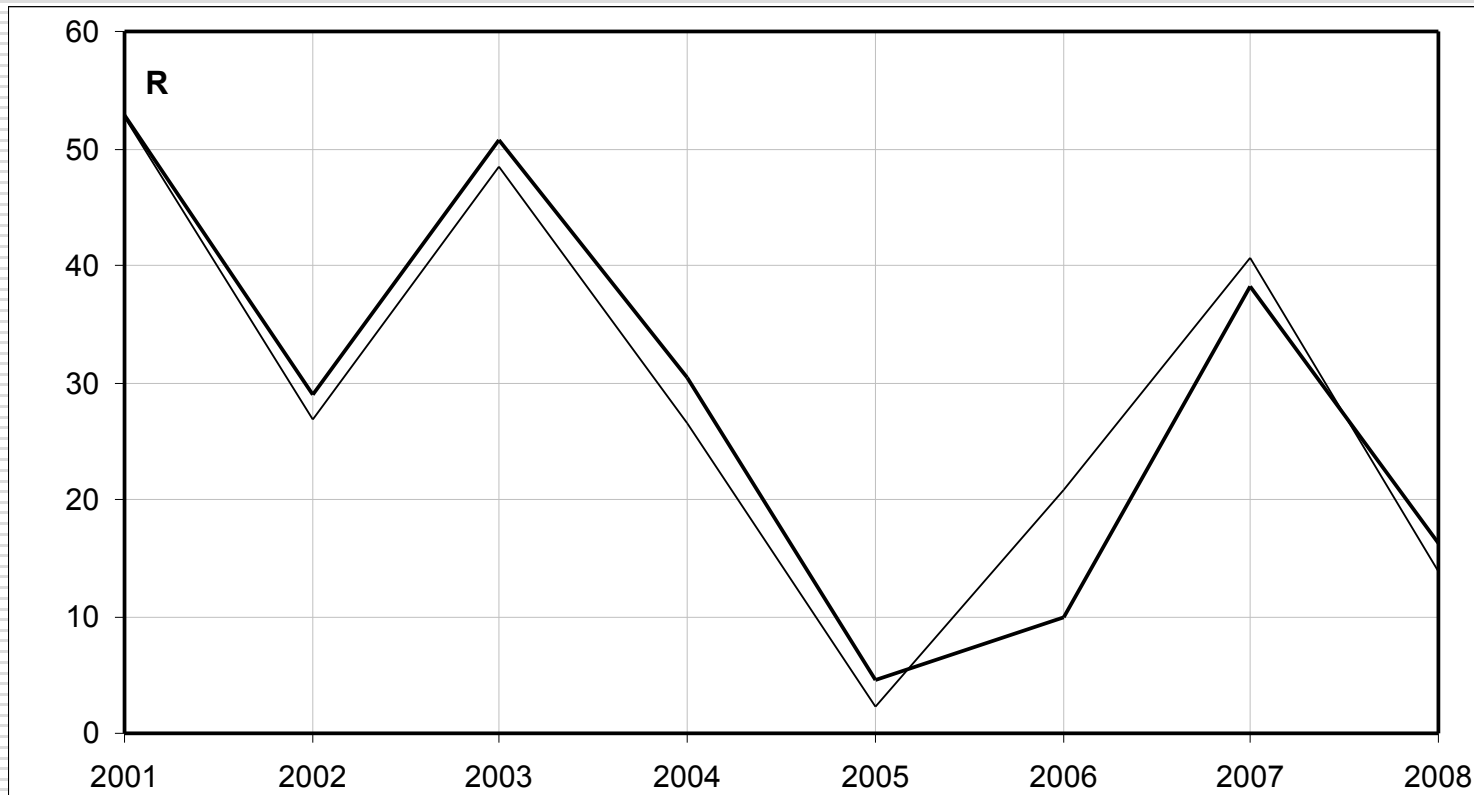
Висновок: Рентабельність залежить від **двохрічного запасу зерна**



Моделювання рентабельності озимої пшениці

R – рентабельність, **W** – валовий збір зернових,
V – валовий збір озимої пшениці

$$R_i = a_0 + a_1 \cdot (W_i + W_{i-1}) + a_2 \cdot (V_i + V_{i-1}) \quad (15)$$





Оцінка рентабельності для регіонів

- ❑ Регіональний коефіцієнт затрат k_i є відношенням затрат на 1 га зернових для даної області до середньоукраїнського показника
- ❑ Рентабельність $R = P/Z - 1 = Y \cdot C/Z - 1$ (16)
- ❑ Якщо прийняти гіпотезу про те, що ціна на зерно однакова в усіх регіонах України, то для рентабельності окремого регіону отримаємо вираз

$$\frac{R_i + 1}{R_0 + 1} = \frac{1}{k_i} \cdot \frac{Y_i}{Y_0} \quad (17)$$

Прогнозні показники для озимої пшениці на 2009р.



	Ціна грн/т – 900	Урожайність	Рентабельність	Площа	Вал	Дохід	Прибуток
		цн / га		тис га	тис тонн	тис грн	тис грн
1АР Крим		21.13	-16.0%	366	774	696403	-132693
2Вінницька		27.05	-3.6%	386	1044	939230	-34752
3Волинська		25.33	-7.8%	137	347	312319	-26461
4Дніпропетровська		27.08	4.7%	510	1381	1242485	56243
5Донецька		24.57	-9.0%	393	967	869925	-86517
6Житомирська		24.75	4.4%	112	277	249035	10594
7Закарпатська		30.01	-11.0%	33	100	90210.1	-11104
8Запорізька		24.82	6.0%	517	1282	1153981	65688
9Ів.-Франківська		24.01	-19.5%	47	112	100914	-24436
10Київська		30.13	2.7%	272	819	737311	19588
11Кіровоградська		28.95	16.5%	354	1025	922086	130924
12Луганська		22.7	6.1%	292	663	596760	34506
13Львівська		25.41	-1.5%	151	383	344636	-5370
14Миколаївська		22.79	1.8%	452	1031	927918	16456
15Одеська		24.98	5.4%	535	1337	1203237	61417
16Полтавська		26.86	-5.4%	361	969	871714	-49427
17Рівненська		24.89	-13.4%	112	279	251115	-38730
18Сумська		23.94	-0.9%	235	563	506331	-4614
19Тернопільська		23	-6.7%	231	531	477756	-34540
20Харківська		28.74	10.1%	478	1375	1237429	113472
21Херсонська		23.14	-5.0%	404	935	841787	-44136
22Хмельницька		24.58	1.8%	216	530	476950	8484
23Черкаська		29.94	3.2%	254	760	683620	20900
24Чернівецька		22.42	-10.1%	31	70	62955.4	-7108
25Чернігівська		23.04	-6.1%	176	405	364539	-23759
Україна		25.67	8.3%	7054	17956		14626



НОВІ РЕЗУЛЬТАТИ ОТРИМАНІ АВТОРОМ

- ❑ *Урожайність озимої пшениці на території України володіє подвійною циклічністю: цикл тривалістю 4 роки і цикл тривалістю 16-18 років*
- ❑ *Динаміка врожайності озимої пшениці є антиперсистентною. Це ж стосується інших природних рядів*
- ❑ *Ступінь детермінованості рядів врожайності зменшується за напрямком "схід - захід"*
- ❑ *Побудовані нові та апробовані відомі прогностні моделі врожайності. Найкращими з них є полігармонічна модель та авторегресійна модель*
- ❑ *Побудована модель "врожайність - рентабельність", яка дозволяє за рахунок маневру посівних площ*
 - а) зменшити втрати; б) отримати прибуток*

ДЯКУЮ ЗА УВАГУ!

