УДК 633.854.78:631 559:551 583(470.6)

Тебуев Х. Х.
TebuevKh.Kh.

МАТЕМАТИЧЕСКОЕ ОБОСНОВАНИЕ РЕПРЕЗЕНТАТИВНОСТИ ТРЕНДОВОЙ СОСТАВЛЯЮЩЕЙ ВРЕМЕННЫХ РЯДОВ

MATHEMATICAL JUSTIFICATION OF THE REPRESENTATIVENESS OF THE TREND COMPONENT OF TIME SERIES

На практике очень часто приходится исследовать временные ряды тех или иных природных и техногенных процессов. При этом, как правило, временные ряды имеют как закономерные (трендовые) составляющие, так и случайные отклонения от него. Для выявления статистической структуры рассматриваемых рядов и применения теории стохастических процессов проверяется их стационарность, однородность и нормальность вектора погрешности. Только в этом случае найденные методом наименьших квадратов оценки будут совместно репрезентативными.

Эти вопросы рассмотрены в работе на примере анализа пространственно-временной изменчивости урожайности подсолнечника на Северном Кавказе. Для описания временной последовательности и выявления статической структуры многолетних флуктуаций урожайности подсолнечника можно применить современный аппарат теории стохастических процессов, после проверки их на случайность и стационарность (процессы, для которых при любых выборках математические ожидания постоянны, а автокорреляционная функция зависит только от сдвига).

Для выделения закономерной составляющей (тренда) при исследовании климатических рядов использовался метод гармонических весов. Для оценки объективности выбранной линии тренда осуществлена проверка на случайность и стационарность ряда отклонения от тренда. Для проверки основной гипотезы (изменение случайной величины $\mathcal{E}_t$не связано с изменением времени) использовали критерии серии, основанной на медиане (для 5% уровня значимости), а условие стационарности случайного процесса на основании зависимости автокорреляционной функции только от величины сдвига разности аргументов $t_i - t_i = \tau$.

Одним из важных вопросов математической статистики является проверка статистических рядов на нормальность распределения и их оценки с помощью соответствующих параметров. Поэтому в данной работе для исследования исходных рядов были рассчитаны следующие статистические характеристики: математическое ожидание, среднеквадратическое отклонение, медиана, мода, коэффициенты - асимметрии, эксцесса и вариации.

Проверка основной гипотезы (о нормальности рассматриваемых распределений) основана в данной работе на методе проверки гипотезы нормальности распределения по x^2 - критерию. Применение критерия x^2 предполагает также использование свойств так называемого «стандартного распределения». Важность обязательности проведения этой процедуры (проверка нормальности распределения) диктуется тем, что среди различных несмещенных оценок, оценки найденные методом наименьших квадратов, могут быть совместно эффективными лишь, если вектор погрешности Δ нормален. Таким образом, наличие оптимальных свойств у метода наименьших квадратов тесно связано с нормальностью вектора погрешности.

Описанным методом на нормальность распределения проверялись ряды урожайности по всем рассматриваемым на Северном Кавказе районам, в том числе и ряд урожайности, усредненный в целом по экономическому району.

Very often we have to investigate period of time of different natural and techno genic processes in practice. Thus, as a rule, time rows have both as natural (trend) compounds, as casual deflections from it. For revealing of statistical frame of considered rows and application of the theory of stochastic processes uniformity and normality of a vector of an error is checked. Only in this case the assessments found at least with square method will be collaterally representative.

These questions of analysis of are considered in this article of the analysis of existential variability of productivity of sunflower in the North Caucasus. It is possible to apply the modern device of the theory of stochastic processes to the description of time sequence for revealing of static frame of perennial fluctuations of productivity of sunflower. After you must check them on accident and regularity (processes for which at any samples mathematical expectations are constant, and autocorrelation function depends only on alteration).

For allocation of a natural component (trend) at research of climatic rows the method of harmonious scales was used. For an assessment of objectivity of the chosen line of a trend check on accident and regularity of some a deflection from a trend is carried out. For check of the basic hypothesis (variety change $\varepsilon_t$is not bound to a delta time) criteria of the series based on a median (for 5 % of level of significance), were used and a condition of regularity of casual process on the basis of dependence of autocorrelation function only from size of alteration of a difference of arguments $t_i - t_i = \tau$. One of important questions of mathematical statistics is checking of statistical rows on normality of allocation and their assessment by means of the conforming parameters. Therefore in the yielded work for research of initial rows following statistical characteristics have been calculated: a mathematical expectation, a deflection, a median, a fashion, quotients - asymmetries, an excess and a variation.

Check of the basic hypothesis (about normality of considered allocations) is based in the yielded work on a method of check of a hypothesis of normality of allocation on x^2- to criterion. Criterion applicationx^2 assumes also use of properties of so-called standard allocation. Importance of compulsion of carrying out of this procedure (check of normality of allocation) is dictated to that among various unbiased estimators, assessments found at least square method, can be collaterally effective only if the error vector Δ is normal. Thus, presence of optimum properties at a least square method is intimately connected to normality of a vector of an error.

With the described method on normality of allocation checked productivity rows were checked on all areas of in the North Caucasus, including a productivity series as a whole on economic region.

Ключевые слова: временный ряд, автокорреляционная функция, стационарность, однородность, нормальность вектора погрешности, репрезентативность, тренд**.**

Key words: a temporary series, autocorrelation function, uniformity, normality of a vector of an error, representativeness, a trend**.**

Тебуев Хызыр Хасанович – кандидат географических наук, доцент кафедры «Природообустройство», ФГБОУ ВО «Кабардино-Балкарский ГАУ»

Тел.: 8-962-650-13-23

E-mail: senta48@mail.ru

Tebuev Khizir Khasanovich – Candidate of Geographical Sciences, Associate Professor, Department of Environmental Engineering, FSBEI HE "Kabardino-Balkarian State Agrarian University

На практике очень часто приходится исследовать временные ряды тех или иных природных и техногенных процессов [4,5,6,7 и др.]. При этом, как правило, временные ряды имеют как закономерные (трендовые) составляющие, так и случайные отклонения от него. Для выявления статистической структуры рассматриваемых рядов и применения теории стохастических процессов проверяется их эргодичность и нормальность вектора погрешности. Только в этом случае найденные методом наименьших квадратов оценки будут совместно репрезентативными (более подробная информация в работах [12,13]). Рассмотрим эти вопросы на примере анализа пространственно-временной изменчивости урожайности подсолнечника на Северном Кавказе.

Продуктивность подсолнечника зависит от соответствия климатических ресурсов биологическим потребностям и от агротехники возделывания культуры, т.е. урожайность является интегральным показателем, отражающим влияние всего комплекса условий сельскохозяйственного производства. Различают изменчивость урожаев, связанную с ростом культуры земледелия и варьированием агрометеорологических условий периода вегетации конкретных лет. При этом, изменчивость урожайности за счет роста культуры земледелия значительно (на порядок и более) ниже, чем вызванные флуктуациями метеорологических условий. Поэтому рост урожайности за счет повышения культуры земледелия рассматривают как тенденцию урожайности (тренд). Проблема выделения трендов – одна из центральных при исследовании случайных процессов, так как она является необходимой предпосылкой для разделения процесса на две основные составляющие – детерминированную и собственно случайную часть ряда.

В агрометеорологии для прогнозирования урожайности сельскохозяйственных культур в динамико-статистических моделях используют выражение

$Y_n = Y_{t+1}.C$ [7], а в статистических моделях $Y_t = f(t) + \mathcal{E}_t$ [8],

где Y_t – ряд урожайности (ц/га); $\mathrm{f}(t)$ – некоторая не случайная функция времени (тренд); $\mathcal{E}_t$ – случайная составляющая временного ряда; Y_{t+1} – тенденция урожайности культуры на прогнозируемый год (ц/га); С – оценка степени отличия складывающихся на дату составления прогноза агрометеорологических условий формирования урожая, от многолетних на фоне, которых формируется тенденция урожайности.

Для описания временной последовательности и выявления статической структуры многолетних флуктуаций урожайности подсолнечника можно применить современный аппарат теории стохастических процессов, после проверки их на случайность и стационарность (процессы, для которых при любых выборках математические ожидания постоянны, а автокорреляционная функция зависит только от сдвига, т.е. эти процессы эргодичны). Существуют различные методы экстраполяции тенденции урожайности (полиномы 1-ой и более высоких степеней, кубическая сплайн-функция [8], модель Бокса-Дженкинса [11] и др.). Для выделения закономерной составляющей (тренда) при исследовании климатических рядов часто используют метод гармонических весов [3]. В данной работе мы и воспользуемся этим методом.

Сущность метода гармонических весов состоит в том, что значение временного ряда (Y_t) взвешивают так, чтобы более поздние наблюдения имели больший вес. В качестве линии тренда берется некоторая ломаная линия, сглаживающая данное число точек временного ряда $\left(f(t)\right)$. Для определения движения скользящего тренда принимается линейный закон изменения за отдельные фазы. Длина фазы $\boldsymbol{K}$ определяется согласно основной гипотезе – изменение случайной величины не зависит от изменения времени.

На основе фактических данных временного ряда предварительно образуют скользящие серии одинаковой длины **K** и рассчитывают уравнения линейных отрезков вида:

$Y_{i(t)} = a_i + b_i\, t \qquad (i = 1,2, \dots, N - K + 1),$ (1),

где N- длина ряда (общее число точек); K - число сглаживаемых точек $K < N$). Общее число уравнений равно $N - K + 1$ причем:

i=1, если $t = 1, \dots, K$

i=2, если $t = 2, \dots, K+1$

…..   …..

$i = N - K + 1$,если $t = N - K + 2, \dots, N$

Параметры a и b в системе уравнений (1) определяются методом наименьших квадратов. Значение каждой функции $Y_i(t)$ находится в точках

$t = i - h - 1 \quad (h = \overline{1, K})$.

Рассчитанные значения $Y_i(t)$ в каждой точке усредняют по полученным уравнениям следующим образом:

$\bar{Y}_i(t) = \frac{1}{g_i}\sum_{j=1}^{g_i} Y_i\,(t)$ (2),

где g_i - количество определения $Y_i(t)$ в каждой точке t_i. Предсказываемое значение $\bar{Y}(t+1)$ временного ряда определяют как:

$\bar{Y}(t+1) = \bar{Y}(t) + \bar{\omega}(t+1)$,   (3)

где $\bar{\omega}(t+1)$ -среднее значение прироста функции f(t) , которую вычисляют по выражению:

$\bar{\omega}(t+1) = \sum_{t=1}^{N-1} C_N\,(t+1)\omega(t+1)$ (4),

Прирост $\omega(t+1)$ функции $f(t)$ определяется по формуле:

$\omega(t+1) = f(t+1) - f(t) = \bar{Y}(t+1) - \bar{Y}(t)$ (5).

Гармонический вес $C_N(t+1)$ находится из соотношения:

$C_N(t+1) = m(t+1)/(N-1)$ (6).

Гармонические коэффициенты $m(t+1)$ рассчитывают согласно формуле:

$m(t+1) = m(t) + \frac{1}{N-t}, (t = 1, \overline{N-1})$ (7),

Отметим, что сумма найденных весов равна единице:

$\sum_{t=1}^{N-1} C_{t+1}^{N-1} = 1$ (8),

Таким образом, при различной длине линейных отрезков **K** получают разные приближения функции $f(t)$.

Нами анализировалась динамика урожайности семян подсолнечника по всем областям (края, республикам) на Северном Кавказе за 29 лет (1987 – 2016 гг.). В качестве примера на рис.1.(а), 2.(а), 3.(а) приводятся данные о средней областной урожайности семян подсолнечника (ц/га) и линии тренда, построенной на основе метода гармонических весов. На них видно, в какие годы наблюдается максимальная или минимальная урожайности, и в какие средние. Представление о колебании урожайности по годам дает рис.1.(б) – 3.(б), где приведены значения отклонений от тренда. Анализ динамики урожайности семян подсолнечника показывает, что на фоне повышения урожайности, в отдельные годы наблюдается значительное ее снижение.

Рис. 1

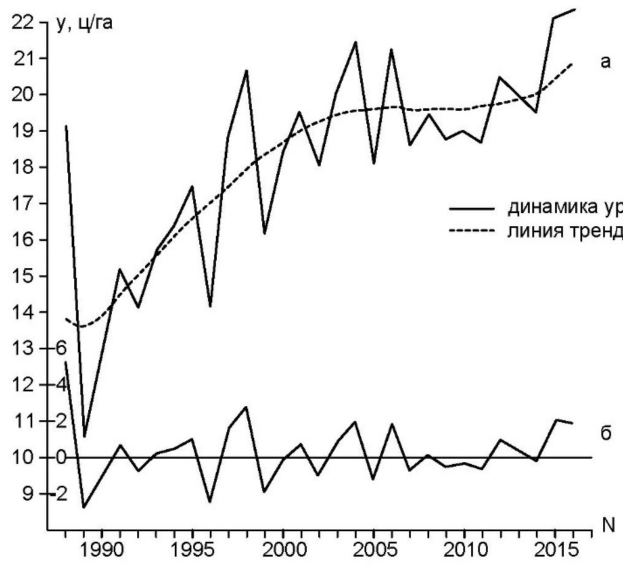

Рис. 2

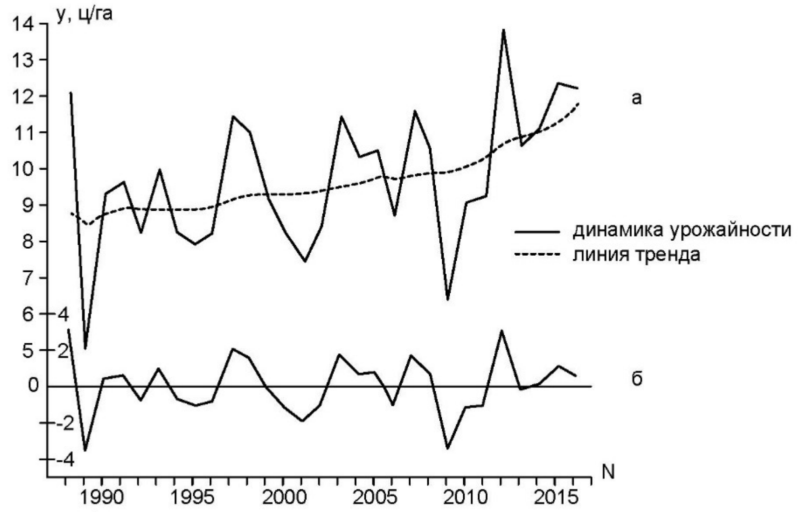

Рис. 3

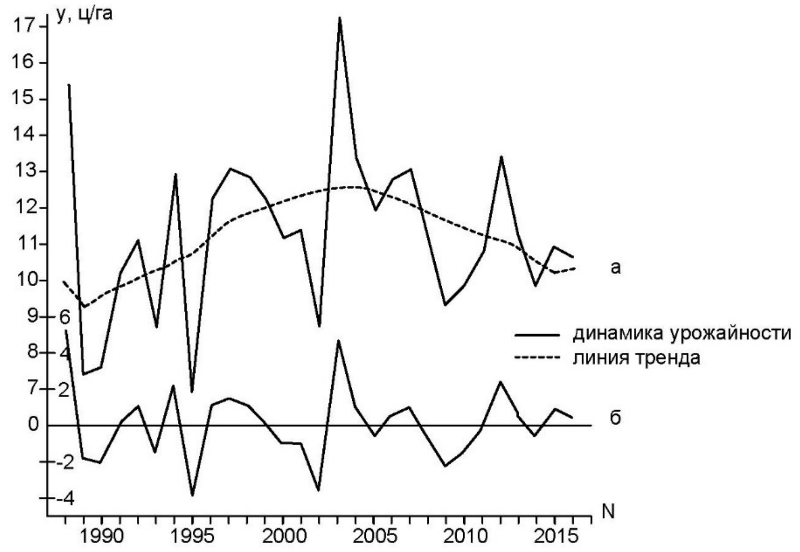

Рис. 1 а, б - динамика урожайности подсолнечника в Краснодарском крае;
Рис. 2 а, б - динамика урожайности подсолнечника в Ставропольском крае;
Рис. 3 а, б - динамика урожайности подсолнечника в Ростовской области;

Для оценки объективности выбранной линии тренда требуется проверка на случайность и стационарность ряда отклонения от тренда. Для проверки основной гипотезы (изменение случайной величины $\mathcal{E}_t$не связано с изменением времени) воспользуемся критерием серий, основанным на медиане $\mathcal{E}_{med}$ [9]. Для того, чтобы исходный ряд представлял случайную выборку, протяженность $[K_m(n)]$ самой длинной серии (последовательность плюсов и минусов, полученных путем сопоставления каждого члена с медианой) не должна быть слишком большой, а общее число серий $U(n)$ - слишком маленькой. Выборка признается случайной, если выполняется следующее условие (для 5% уровня значимости):

$$\begin{cases} K_m(n) < [3{,}3(lgn + 1)] \\ U(n) > \left[\frac{1}{2}\left(n + 1 - 1{,}96\sqrt{n-1}\right)\right] \end{cases} \qquad (9)\ .$$

Чтобы получить левые части системы неравенства (9) из отклонения от тренда $\mathcal{E}_t, i = \overline{1,n},$ образуем для каждой из рассматриваемых областей (табл.1) вариационный ряд $\mathcal{E}^i,\ i = \overline{1,n}$ так, чтобы $\mathcal{E}^{(1)} > \mathcal{E}^{(2)} > \cdots > \mathcal{E}^{(m)}$, $\mathcal{E}_{med}$ - медиана этого вариационного ряда. По величине отклонения $\mathcal{E}_i$ от $\mathcal{E}_{med}$ ставится соответствующий знак, по следующему правилу:

$$\mathcal{E}_i - \mathcal{E}_{med} = \begin{cases} > 0\ + \\ < 0\ - \\ = 0\ \ \textit{пропуск} \end{cases} \qquad (10).$$

Таблица 1 – Оценка случайности отклонения от тренда

номер года	Северо-Кавказский экономический район			Краснодарский край			Ставропольский край			Ростовская область			Дагестанская республика			Кабардино-Балкарская республика			Северо-Осетинская республика			Чеченская и Ингушская республики*		
	ε	В порядке убывания	Серии	ε	В порядке убывания	Серии	ε	В порядке убывания	Серии	ε	В порядке убывания	Серии	ε	В порядке убывания	Серии	ε	В порядке убывания	Серии	ε	В порядке убывания	Серии	ε	В порядке убывания	ери и
1	4,821	4,821	+	5,27	5,27	+	3,25	3,25	+	5,19	5,19	+	0,97	2,57	+	3,47	3,47	+	2,22	2,73	+	2,47	2,47	+
2	-2,424	2,387	-	2,86	2,71	-	-3,51	3,14	-	-1,93	4,76	-	-1,08	2,38	-	-0,27	2,47	-	-1,40	2,47	-	-2,61	2,13	-
3	-0,859	1,748	-	-1,31	1,94	-	0,47	2,20	+	-2,04	2,38	-	2,07	2,07	+	0,04	2,03	-	1,04	2,22	+	1,74	1,74	+
4	0,577	1,686	+	0,69	1,83	+	0,72	1,86	+	0,23	2,37	-	-0,70	2,00	-	0,34	1,85	+	-0,16	1,64	-	0,53	1,31	+
5	-0,139	1,613	+	-0,89	1,74	-	-0,66	1,79	-	1,11	1,54	+	-1,97	1,69	-	-1,24	1,54	-	-1,76	1,50	-	-1,82	1,14	-
6	-0,213	1,193	-	0,14	1.58	+	1,11	1,72	+	-1,61	1,11	-	-0,65	1,67	-	-0,13	1,53	-	-0,91	1,45	-	-0,45	0,98	-
7	1,99	1,121	+	0,33	1.35	+	-0,71	1,11	-	2,37	1,08	+	1,46	1,46	+	-0,05	0,64	-	1,37	1,37	+	0,38	0,93	+
8	-1,380	1,093	-	1,02	1.02	+	-1,05	1,05	-	-3,94	1,06	-	-1,16	0,97	-	0,14	0,51	+	0,02	1,30	+	0,75	0,91	+
9	-1,591	0,922	-	-2,72	0.74	-	-0,83	0,82	-	1,08	0,98	+	0,46	0,72	+	-0,98	0,50	-	-0,50	1,09	-	-0,95	0,85	-
10	1,613	0,838	+	1,35	0.73	+	2,20	0,72	+	1,54	0,91	+	2,00	0,68	+	-0,98	0,34	-	1,64	0,95	+	0,93	0,75	+
11	1,686	0,577	+	2,71	0.69	+	1,72	0,70	+	1,06	0,75	+	0,11	0,46	+	0,50	0,29	+	-0,20	0,65	-	-0,06	0,66	-
12	-0,468	0,311	-	-2,23	0.49	-	-0,11	0,66	-	0,15	0,49	-	-0,42	0,42	+	2,03	2,25	+	0,65	0,17	+	1,14	0,60	+
13	-0,782	0,137	-	-0,23	0.33	-	-1,10	0,52	-	-0,98	0,38	-	-1,23	0,25	-	0,25	0,20	+	-0,84	0,06	-	-1,30	0,53	-
14	-0,523	-0,139	-	0,49	0.14	+	-1,92	0,47	-	-0,94	0,30	-	-1,59	0,11	-	-1,30	0,14	-	-0,77	0,02	-	-0,39	0,38	-
15	-2,087	-0,213	-	-1,18	0.12	-	-0,99	0,14	-	-3,78	0,23	-	-1,89	-0,46	-	0,10	0,10	-	-0,76	-0,02	-	-0,72	0,32	-
16	2,387	-0,282	+	0,74	-0.14	+	1,86	-0,11	+	4,76	0,15	+	2,57	-0,50	+	-0,42	0,04	-	0,95	-0,16	+	0,32	-0,06	-
17	1,121	-0,348	+	1,94	-0.26	+	0,70	-0,20	+	0,91	-0,44	+	1,69	-0,53	+	1,53	-0,05	+	0,17	-0,20	+	2,13	-0,39	+
18	-0,285	-0,383	-	-1,49	-0.54	-	0,82	-0,66	+	-0,57	-0,57	-	-1,25	-0,56	-	0,29	-0,13	+	-1,12	-0,50	-	-0,47	-0,45	-
19	0,838	-0,468	+	1,74	-0.64	+	-1,14	-0,71	-	0,49	-0,65	+	0,68	-0,65	+	-0,41	-0,17	-	1,45	-0,76	+	-1,08	-0,47	-
20	0,311	-0,523	+	-0,04	-0.85	-	1,79	-0,83	+	0,98	-0,71		0,25	-0,70	+	0,20	-0,27	+	0,06	-0,77	+	0,98	-0,48	+
21	-0,348	-0,782	-	-0,14	-0.89	-	0,66	-0,99	+	-0,71	-0,94	-	-0,52	-1,08	-	-1,47	-0,41	-	2,47	-0,84	+	-0,62	-0,61	-
22	-1,704	-0,859	-	-0,80	-1.02	-	-3,56	-1,05	-	-2,26	-0,98	-	-0,50	-1,16	-	1,54	-0,42	+	-0,88	-0,88	-	0,50	-0,72	+
23	-1,253	-0,942	-	-0,64	-1.04	-	-1,07	-1,07	-	-1,56	-1,58	-	-0,46	-1,23	-	-2,01	-0,98	-	-1,76	-0,91	-	-0,48	-0,95	-
24	-0,942	-1,253	-	-1,02	-1.18	-	-1,13	-1,10	-	-0,44	-1,61	-	-0,56	-1,25	-	0,51	-0,98	+	-1,78	-1,12	-	-1,44	-1,08	
25	1,748	-1,380	+	0,73	-1.31	+	3,14	-1,13	+	2,38	-1,93	+	1,67	-1,59	+	-1,58	-1,24	-	1,50	-1,40	+	0,91	-1,30	
26	0,137	-1,591	+	0,12	-1.49	-	-0,20	-1,14	-	0,38	-2,04	+	0,72	-1,89	+	-0,17	-1,30	-	-2,10	-1,76	-	-1,43	-1,43	
27	-0,383	-1,704	-	-0,54	-2,23	-	0,14	-1,92	-	-0,65	-2,26	-	2,38	-1,97	+	0,64	-1,47	+	-0,02	-1,76	-	0,85	-1,44	
28	1,193	-2,087		1,83	-2,72	+	1,05	-3,51	+	0,75	-3,72	+	-4,28	-3,39	-	1,85	-1,58	+	2,73	-1,78	+	1,31	-1,82	
29	0,922	-2,424	+	1,58	-2,86	+	0,52	-3,56	+	0,30	-3,94	+	-3,39	-4,28	-	2,47	-2,01	+	1,30	-2,10	+	0,66	-2,61	
	$\varepsilon_{med} = 0.21$			$\varepsilon_{med} = 0.12$			$\varepsilon_{med} = 0.14$			$\varepsilon_{med} = 0.23$			ε_{med}=- 0.46			ε_{med}=0.10			$\varepsilon_{med} = -0.02$			$\varepsilon_{med} = 0.32$		

*- объединенные данные

Затем подсчитываем протяженность самой длинной серии $K_m(n)$ и общее число серий каждой области (табл.2). Сравнение (табл.2) левых и правых частей неравенств (9) показывает, что оба неравенства справедливы.

Таблица 2 – Оценка правильности выбора тренда.

Территория	$K_m(n)$	$U(n)$	A	B
Северокавказский экономический регион	4	13	8,13	9,82
Краснодарский край	5	15	8,13	9,82
Ставропольский край	3	15	8,13	9,82
Ростовская область	4	15	8,13	9,82
Дагестанская республика	4	14	8,13	9,82
Кабардино-Балкарская республика	3	17	8,13	9,82
Северо-Осетинская республика	3	17	8,13	9,82
Чеченская и Ингушская республики*	3	19	8,13	9,82

*- используются объединенные данные по этим республикам

Гипотеза о случайном характере отклонений уравнений временного ряда урожайности от тренда, определенного методом гармонических весов, принимается. Таким образом, тенденция возрастания урожайности, характеризующая уровень культуры земледелия при средних почвенно-климатических условиях, исключена из временных рядов верно.

Проверим гипотезу стационарности случайной компоненты временных рядов урожайности [9]. Основным условием стационарности случайного процесса является условие зависимости автокорреляционной функции только от величины сдвига разности аргументов $t_i - t_i = \tau$. На основе данных об отклонениях урожаев от тренда $\mathcal{E}_t(t - \overline{1,n})$ рассчитываем автокорреляционные коэффициенты для $N, N-1, \dots, N-K$ наблюдений, причем из ряда последовательно будем исключать 0.1 K наблюдений. Для каждого полученного ряда $\mathcal{E}_t(K)$ автокорреляционные коэффициенты определяются со сдвигом фаз на $\tau = 1,2,\dots,l$. В итоге имеем матрицу автокорреляционных коэффициентов размером $K \times l$. В табл.3. представлены значения автокорреляционных функций для каждой области при разном количестве наблюдений.

Таблица 3 – Проверка стационарности отклонении урожайности подсолнечника от тренда

Количество наблюдений	Критерии	Сдвиги												
		1	2	3	4	5	6	7	8	9	1 0	1 1	12	3
Краснодарский край														
29	r	-0.385	-0.290	0.350	-0.186	-0.236	0.198	0.063	-0.449	0.226	0.291	-0.194	0.030	0.080
	z	-0,406	-0.299	0.366	-0.188	-0.240	0.200	0.063	-0.483	0.230	0.300	-0.197	0.030	0.080
28	r	-0,212	-0.215	0.374	-0.123	-0.303	0.225	-0.032	-0.255	0.131	0.013	0.131	0.113	0.036
	z	-0,215	-0.219	0.394	-0.123	-0.313	0.229	-0.032	-0.260	0.132	0.013	0.132	0.114	0.036
27	r	-0,321	-0.197	0.348	-0.132	-0.315	0.311	-0.253	-0.192	0.396	-0.217	0.104	0.167	-0.116
	z	-0,333	-0.199	0.364	-0.133	-0.326	0.321	-0.259	-0.195	0.419	-0.221	0.104	0.169	-0.116
26	r	-0,310	-0.232	0.356	-0.128	-0.296	0.252	-0.220	-0.107	0.342	-0.246	0.125	0.100	-0.089
	z	-0,321	-0.236	0.372	-0.128	-0.305	0.258	-0.224	-0.107	0.356	-0.251	0.125	0.100	-0.089
25	r	-0,330	-0.234	0.355	-0.147	-0.271	0.235	-0.308	-0.067	0.354	-0.255	0.168	0.088	-0.165
	z	-0,310	-0.238	0.371	-0.148	-0.278	0.240	-0.318	-0.067	0.370	-0.261	0.170	0.089	-0.166
24	r	-0,333	-0.232	0.382	-0.232	-0.252	0.342	-0.423	-0.082	0.373	-0.327	0.193	0.192	-0.262
	z	-0,312	-0.236	0.403	-0.236	-0.257	0.356	-0.451	-0.082	0.392	-0.340	0.195	0.194	-0.268
23	r	-0,303	-0.238	0.437	-0.242	-0.300	0.394	-0.423	-0.083	0.392	-0.329	0.207	0.205	-0.308
	z	-0,313	-0.243	0.468	-0.247	-0.310	0.416	-0.451	0.083	0.415	-0.341	0.210	0.208	-0.318
22	r	-0,313	-0.235	0.435	-0.304	-0.303	0.401	-0.429	-0.074	0.391	-0.376	0.250	0.219	-0.305
	z	-0,324	-0.240	0.466	-0.314	-0.313	0.425	-0.459	-0.074	0.413	-0.395	0.255	0.222	-0.315
21	r	),275	).282	.404	).254	).293	.395	).401	).097	.359	).347	.187	.301	).293
	z	),232	).290	.428	).260	).302	).417	).425	).097	.375	).362	).190	).310	).302
	$\bar{Z}$	),313	).244	).403	).197	).294	).318	).284	).161	.345	).207	).132	).160	).162
	χ^2	,494	).184	).291	).810	).158	.396	).927	.705	.868	).840	.414	.241	.198

Для проверки однородности коэффициентов автокорреляции входящие в τ группу для каждого $r_\tau(n-k)$ были рассчитаны величины z - критерия по формуле:

$$Z_{\tau_k} = \frac{1}{2}\ln\frac{1+\tau_\tau(n-k)}{1-\tau_\tau(n-k)} \qquad (11)$$

Далее для каждого сдвига была получена величина:

$$\overline{Z_\tau} = \sum_{k=0}^{k} Z_{\tau_k}/(K+1) \qquad (12)$$

Доказано [9], что величина:

$$\Omega = \sum_{k=0}^{k} \frac{\left(Z_{\tau_k} - \bar{Z}_\tau\right)}{1/(n-k-3)} \qquad (13)$$

Распределена как x^2 с K степенями свободы. Верхняя граница для x^2 с 13 степенями свободы при 5% уровне значимости равна 22,36. Данные табл. 3. показывают, что все вычисленные значения x^2 меньше этой величины. Исходя из этого, гипотеза об однородности коэффициентов временных рядов урожайности представляет собой стационарный случайный процесс. Он протекает приблизительно однородно, и с течением времени ни средняя амплитуда, ни характер случайных колебаний урожайности вокруг линии тренда не будут обнаруживать существенных изменений. Отметим, что расчеты по выше описанному алгоритму произведены по программе $TREN$, составленной в отделе агрометпрогнозов ВНИИСХМ.

Одним из важных вопросов математической статистики является проверка статистических рядов на нормальность распределения и их оценки с помощью соответствующих параметров. Поэтому в данной работе для исследования исходных рядов были рассчитаны следующие статистические характеристики: математическое ожидание, среднеквадратическое отклонение, медиана, мода, коэффициенты - асимметрии, эксцесса и вариации.

Использование статистических параметров предполагает наличие однородности фактических рядов. Для проверки однородности ряда его крайние значения сравниваются с математическим ожиданием и среднеквадратическим отклонением.

При доверительной вероятности, равной $1-p$ (где p – уровень значимости), для наблюдения x_i справедлива доверительная оценка:

$$|x_i - м| < U_{1-p}\sigma, \qquad (14)$$

где $м, \sigma$ - соответственно математическое ожидание и среднеквадратическое отклонение ряда; U_{1-p}- квантиль стандартного нормального распределения.

Критерий проверки крайнего элемента состоит в том, что при уровне значимости P элемент x_k не принадлежит данной генеральной совокупности, если выполняется условие:

$$|x_k - м| > U_{1-p,N}\ \sigma \text{ или } \frac{|x_k - м|}{\sigma} > U_{1-p,N} \qquad (15)$$

В случае достаточно большого объема выборки $(N > 25)$ в качестве оценок $м$ и σ берутся среднеарифметическое и среднеквадратическое отклонения фактического ряда. При $(N < 25)$ в силу смещения рассчитанных оценок ряда от характеристик генеральной совокупности используют τ - критерий, который имеет специальное распределение [4]. В этом случае крайний элемент выборки не относится к генеральной совокупности, если:

$$\left|\frac{x_k - \bar{x}}{\sigma_x} > \tau_{1-p}\right| \qquad (16)$$

Существуют и другие методы отсева грубых погрешностей [1,2].

Проверка основной гипотезы (о нормальности рассматриваемых распределений) основана в данной работе на методе проверки гипотезы нормальности распределения по x^2 - критерию. Применение критерия x^2 предполагает также использование свойств так называемого стандартного распределения [2,4]. Уравнение кривой стандартного нормального распределения имеет вид:

$$Y = f(z) = \frac{1}{\sqrt{2П}} l^{-\frac{z^2}{2}} = 0{,}4 l^{\frac{z^2}{2}} \qquad (17),$$

где $z = (x - м)/\sigma$

Опишем основные этапы реализации алгоритма этого метода.

1. Методом гармонических весов, алгоритм которого изложен выше, определяется ряд тенденции урожайности $\widehat{Y}_i(t+1), i=\overline{1,N}$

2. Находится относительная урожайность по формуле:

$$P=\frac{Y(y)}{\hat{Y}\ (t)}\cdot 100\% \quad , \qquad (18)$$

где $Y(t)$ - заданный временной ряд урожайности, определенная методом гармонических весов.

3. Анализируется ряд относительной урожайности в процентах, согласно критерию x^2. В результате определяется - существенно ли распределение повторяемостей отличается от нормального распределения?

Для этого значения временного ряда делятся на градации:

$$\Omega=\sum_{j=1}^{G_n}\frac{\left(Q_j-H_j\right)^2}{H_j} \qquad (19)$$

где G_n - число градации; Q_j - число попадания в j - ю градацию согласно временному ряду; H_j - число попадания в j - ю градацию согласно нормальному закону, с математическим ожиданием, равным среднему по выборке и дисперсией, соответствующей выборочной оценке дисперсии.

При этом, если Ω меньше соответствующего значения в таблице x^2 гипотеза принимается. Число попадания Q_j находится по формуле:

$$Q_j=PS_j\left(x_{2j}-x_{1j}\right) \qquad (20),$$

где $PS_j=\frac{\sum nWK_j}{n}$ - частота повторяемости, которая определяется как отношение суммы событий ряда попадающих в j градацию (WK_j) на число событии в j градаций (n_1).

Накопление повторяемости (rn_{1j}) получается путем последовательного суммирования приведенных частот градации:

$$rn_{1j}=\frac{\sum_j nWK_j}{n}+rn_{1j-1} \qquad (21)$$

Для нахождения накопленных повторяемостей (rn_{1j}) используется выражение:

$$rn_{2j} = rn_{1j} \cdot 100\% \qquad (22)$$

За медиану (med) принимается то значение абсциссы функции накопления повторяемости в процентах, которое соответствует ординате 50%. Мода (mod) определяется, как наиболее вероятное значение случайной переменной приблизительно, с точностью до градации. Математическое ожидание (м) находится из соотношения:

$$\text{м} = \frac{\sum_{i=1}^{n} WK_i}{n} \qquad (23)$$

Среднеквадратическое отклонение

$$\sigma = \sqrt{\frac{\sum_{i}^{n} (Wk_i)^2}{i}} \quad . \qquad (24)$$

Коэффициенты асимметрии (as), эксцесса (es) и вариации (cov) соответственно равны:

$$as = \frac{3(\text{м} - med\,)}{\sigma} \qquad (25)$$

$$es = \sum_{i}^{n} (WK_i - \text{м})^4 / (\sigma^4 n) \ . \qquad (25)$$

$$cov = \frac{\sigma}{\text{м}} \cdot 100\% \ (27)$$

Важность обязательности проведения этой процедуры (проверка нормальности распределения) диктуется тем, что среди различных несмещенных оценок, оценки, найденные методом наименьших квадратов, могут быть совместно эффективными лишь, если вектор погрешности Δ нормален. Таким образом, наличие оптимальных свойств у метода наименьших квадратов тесно связано с нормальностью вектора погрешности [2].

Описанным методом на нормальность распределения проверялись ряды урожайности по всем рассматриваемым на Северном Кавказе районам, в том числе и ряд урожайности, усредненный в целом по экономическому району. Выяснили, что по критерию x^2 можно принять гипотезу о нормальности распределения относительных рядов урожаев в (табл.4):

1. В целом по Северокавказскому экономическому региону
2. Краснодарском крае

3. Ростовской области

4. Кабардино-Балкарской республике

5. Северо – Осетинской республике

6. (Чеченской и Ингушской) республикам

Некоторое смещение от нормального распределения обнаружили относительные ряды урожаев в Ставропольском крае и Дагестанской республике. Вычисленные статистические характеристики для соответствующих рядов приводятся в табл. 4.

Иногда для проверки близости эмпирического распределения к нормальному используют критерий Колмогорова [2]. Для проверки гипотезы нормальности распределения вычислим:

$$\left.\begin{array}{l} S_{G_1}=\sqrt{\dfrac{6n(n-1)}{(n-2)(n+1)(n+3)}} \\ \hline S_{G_2}=\sqrt{\dfrac{24n(n-1)^2}{(n-3)(n-2)(n+3)(n+5)}} \end{array}\right\} \quad (26)$$

Таблица 4 – Характеристики распределения

Район, область, край, республика	Характеристики							Наблюдаемые частоты распределены нормально, принимается на 10% уровне
	$м$	σ	med	mod	as	es	cov	
Северо-Кавказкий	98.86	13.03	97.64	97.23	0.28	3.28	13.18	$\hat{x}^2 = 4.70 << x^2_{(17.10)} = 24.77$
Краснодарский	98.81	11.86	98.52	97.33	0.075	3.45	12.07	$\hat{x}^2 = 4.66 << -,,-$
Ставропольский	101.37	18.39	104.82	89.87	-0.56	2.84	18.14	$\hat{x}^2 = 32.49 > -,,-?$
Ростовская	97.70	19.60	96.48	95.25	0.19	3.16	20.06	$\hat{x}^2 = 23.92 < -,,-$
Дагестанская	92.90	28.66	96.49	96.49	-0.37	4.54	30.85	$\hat{x}^2 = 36.52 > -,,-?$
Кабардино-Балкарская	99.48	13.34	98.65	91.15	0.188	2.09	13.41	$\hat{x}^2 = 2.81 << -,,-$
Северо-Осетинская	99.37	14.90	97.51	93.78	0.37	2.36	14.99	$\hat{x}^2 = 12.29 << -,,-$
(Чеченской и Ингушской)	99.19	14.04	100.51	100.95	-0.28	2.89	14.16	$\hat{x}^2 = 7.73 << -,,-$

Если $|as| \leq 3S_{G_1}$ и $|es| \leq 5S_{G_2}$, то гипотеза о нормальности распределения не отвергается. По данному критерию нормальность распределения по Ставропольскому краю ряда отклонений урожайности не отвергается:

|as| = 0,56 << $3S_{G_1}$ = 1,3005

|es| = 2,84 << $5S_{G_1}$ = 4,008

Нормальность распределения не принимаются для ряда отклонений урожайности по Дагестанской республике.

Таким образом, можем принять, что в целом в рассматриваемом районе вектор погрешности Δ нормален и, следовательно, найденные нами методом наименьших квадратов статистические оценки будут более эффективными для рассматриваемых рядов, чем различные несмещенные оценки.

После такой проверки объективности выбранной линии тренда и установления нормальности распределения случайной составляющей можем провести анализ изменчивости урожайности. В связи с тем, что изменчивость урожайности сельскохозяйственных культур зависит, в основном, от культуры земледелия и агрометеорологических условий среды обитания, в работе [5] предлагается рассматривать отдельно изменчивость урожайности, связанную с повышением культуры земледелия:

$$C_a = \frac{\sigma}{\bar{y}} = \frac{1}{\bar{y}}\sqrt{\frac{\sum(\hat{Y}_i - \bar{Y})^2}{n-2}} \qquad (27)$$

И изменчивость урожайности, связанную с вариабельностью метеорологических факторов:

Таблица 5 – Изменчивость урожайности семян подсолнечника в Северокавказском экономическом регионе

Область, край, республика	Характеристика				
	$\sigma_{общ}$	$\sigma_{зем}$	$C_{v\ общ}$	$C_{V\ зем}$	σ_m
Краснодарский	2.81	2.17	0.16	0.12	1.79
Ставропольский	1.97	0.80	0.20	0.18	1.80
Ростовский	2.34	1.66	0.21	0.15	1.65
Дагестанская	1.75	1.20	0.23	0.17	1.27
Кабардино-Балкарская	1.69	1.04	0.15	0.10	1.33
Северо-Осетинская	2.07	1.42	0.20	0.14	1.51
Чеченская и Ингушская	1.30	0.37	0.14	0.04	1.25

$$C_m = \frac{1}{\bar{y}}\sqrt{\frac{\sum_i^n (Y_i)^2 - \sum_i^n (\hat{Y}_i - \bar{Y})^2}{n-1}} \qquad (28),$$

где $\bar{Y}$- средняя многолетняя урожайность; Y_i- урожайность конкретного года; $\hat{Y}_i$- динамическая средняя величина (урожайность по тренду в конкретном году); n – количество исследованных лет. Заметим, что величина C_m

не может считаться устойчивой, так как обнаруживает значительную изменчивость при разной длине рассматриваемого периода [5,10 и др.], поэтому и объективная информативность ее ограничена. Данная методика была применена нами для оценки климатической составляющей изменчивости урожаев подсолнечника на территории Северного Кавказа. Вычисленные на основании формулы (30) значения C_m нанесены на рис. 4, остальные характеристики приводятся в табл. 5.

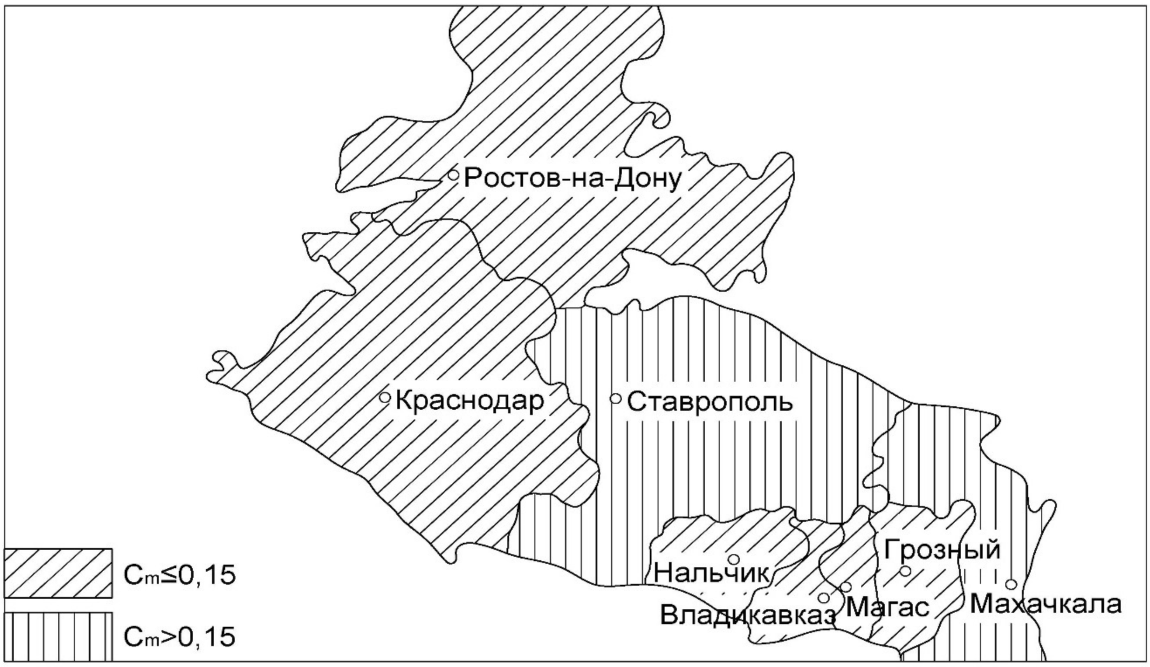

Рис.4. Климатическая изменчивость (Cm) урожая подсолнечника в Северо-Кавказском экономическом районе

Таким образом, исследование временных рядов и последующий анализ пространственно-временной изменчивости урожайности подсолнечника на Северном Кавказе показали, что при выборе линии тренда, определенного методом гармонических весов, изменения случайной величины $\mathcal{E}_t$ не связаны с изменением времени (вектор погрешности Δ в целом нормален), т.е. тренд вычленен верно, и найденные нами методом наименьших квадратов статистические оценки будут более эффективными для рассматриваемых рядов, чем различные несмещенные оценки [11,15]. Убедившись, что вектор погрешности Δ нормален, можем приступить к исследованию случайной составляющей временного ряда с целью нахождения связей $\mathcal{E}_t$ с погодными характеристиками.

Данную методику можно применить для исследования любых временных рядов, где требуется определить детерминированную составляющую и случайные отклонения от тренда.

Литература

1. Давитая Ф.Ф., Дроздов О.А. Проблемы горной климатологии. - Тбилиси, 1969.-19 с.
2. Львовский Е.Н. Статистические методы построения эмпирических формул. - М: Высшая школа, 1982.-224 с.
3. Методические указания по составлению прогноза урожайности и валового сбора всех зерновых и зернобобовых культур, а также основных сельскохозяйственных культур в Прибалтике, Белоруссии и европейской части РСФСР./Под ред. А. Н. Полевого/.- Обнинск, ВНИИГМИ-МЦД,1987.-108 с.
4. Пустыльник Е.И. Статистические методы анализа и обработки наблюдений. - М: Наука,1968.-288 с.
5. Статистический анализ экономических временных рядов и прогнозирование.-М.: Наука,1973.-295 с.
6. Тебуев Х.Х. Методы оценки агрометеорологических условий и прогнозы урожайности подсолнечника с использованием динамической модели продукционного процесса. Дип. МЦД, 1987, №716-ГМ 570. -27с.
7. Тебуев Х.Х. Моделирование влияния агрометеорологических условий на формирование продуктивности подсолнечника. - Метеорология и гидрология, 1988.№ 10
8. Тебуев Х.Х. Метод долгосрочного прогноза урожайности семян подсолнечника. Метеорология и гидрология, 1989, № 6.
9. Френкель А.А. Математические методы анализа динамики и прогнозирования производительности труда. - М.:Экономика,1972.-189 с.
10. Чирков Ю.И. Агрометеорологические условия и продуктивность кукурузы. - Л: Гидрометеоиздат, 1969. – 251 с.
11. Box G. E. P. and Jenkins G. M. (1976). Time Series analysis: Forecasting and Control. — San Francisco: Holden-Day.

12. Ewdards A. W. F. (1972). Likelihood. — Cambridge: Cambridge University Press.
13. Hoel P. G. (1962). Introduction to Mathematical Statistics, third edition.- New York: Wiley
14. Kendall M. G. (1973). Time — series. — London: Griffin.
15. Kennedy J. O. S. (1972). A model for determining optimal marketing and Feeding policies for beef cattle. — Journal of Agricultural Economics, 23, 147—159.

References

1. Davitaya F.F., Drozdov O.A. Problemy gornoj Klimatologii. - Tbilisi, 1969.-19 s.
2. L'vovskij E.N. Statisticheskie metody postroeniya ehmpiricheskih formul. - M: Vysshaya shkola, 1982.-224 s.
3. Metodicheskie ukazaniya po sostavleniyu prognoza urozhajnosti i valovogo sbora vsekh zernovyh i zernobobovyh kul'tur, a takzhe osnovnyh sel'skohozyajstvennyh kul'tur v Pribaltike, Belorussii i evropejskoj chasti RSFSR./Pod red. A. N. Polevogo/.- Obninsk VNIIGMI-MCD,1987.-108 s.
4. Pustyl'nik E.I. Statisticheskie metody analiza i obrabotki na-blyudenij. - M: Nauka,1968.-288 s.
5. Statisticheskij analiz ehkonomicheskih vremennyh ryadov i progno-zirovanie.- M.: Nauka,1973.-295 s.
6. Tebuev H.H. Metody ocenki agrometeorologicheskih uslovij i prognozy urozhajnosti podsolnechnika s ispol'zovaniem dinamiche-skoj modeli produkcionnogo processa. Dip. MCD, 1987, №716-GM 570. -27s.
7. Tebuev H.H. Modelirovanie vliyaniya agrometeorologicheskih uslovij na formirovanie produktivnosti podsolnechnika. - Meteorologiya i gidrologiya, 1988.№ 10
8. Tebuev H.H. Metod dolgosrochnogo prognoza urozhajnosti semyan podsolnechnika. - Meteorologiya i gidrologiya, 1989, № 6.
9. Frenkel' A.A. Matematicheskie metody analiza dinamiki i prognozirovaniya proizvoditel'nosti truda. - M.:EHkonomika,1972.-189 s.
10. CHirkov YU.I. Agrometeorologicheskie usloviya i produktivnost' kukuruzy. - L: Gidrometeoizdat, 1969. – 251 s.
11. Box G. E. P. and Jenkins G. M. (1976). Time Series analysis: Fore-casting and Control. — San Francisco: Holden-Day.
12. Ewdards A. W. F. (1972). Likelihood. — Cambridge: Cambridge University Press.
13. Hoel P. G. (1962). Introduction to Mathematical Statistics, third edi-tion.- New York: Wiley
14. Kendall M. G. (1973). Time — series. — London: Griffin.
15. Kennedy J. O. S. (1972). A model for determining optimal marketing and Feeding policies for beef cattle. — Journal of Agricultural Economics, 23, 147—159.