Forecasting potato yield dynamics in the Tashkent region of the Republic of Uzbekistan

V. V. Vakhabov*, and M. A. Hidoyatova
Tashkent Institute of Irrigation and Agricultural Mechanization Engineers’ National Research University, Tashkent, Uzbekistan

Abstract. Observations of some phenomenon, the nature of which changes over time, generates an ordered sequence, which is called a time series. The statistical regularity of the series of dynamics $\bar{y}_t$—the average potato yield in the Tashkent region of the Republic of Uzbekistan (based on the materials of the CSO of the Republic of Uzbekistan for 2000-2018 years) was studied by the method of statistical analysis of time series. Point and interval estimates for the average potato yield were built with a 95% guarantee, explicit types of trends were identified and the yield in the Tashkent region was predicted for subsequent years. Using the Darbin-Watson statistical criteria, it was found that the average potato yield in the region has autocorrelation dependence.

1 Introduction

In each area there are phenomena that need to be studied in their development and change over time. For example, to try to predict the future based on knowledge of the past, to manage the process, to describe the characteristic features of a series based on a limited amount of information. When processing time series, the methods rely largely on the methods developed by mathematical statistics for distribution series. To date, statistics has a variety of methods for analyzing time series.

In general, the study of the yield of agricultural processes as a discrete dynamic series and the prediction of their yield based on experimental data play an important role in determining the economic efficiency of farms.

In general, the time series $\{y_t, t \in T\}$ consists of four components: trend; fluctuations relative to trend; seasonality effect; random component. The works of Anderson [1], Kendall [2], Lewis [3], Brillinger [4], Chetyrkin [5] and others are devoted to the study and analysis of dynamic series.

2 Results and Discussion

The geometric image of the observed data (Table 1) on a rectangular coordinate system gives grounds, in the first approximation, to assume the hypothesis that the trend part of the process under study has a linear dependence $y(t) = a_1 t + a_0$ (see Fig. 1).

*Corresponding author: v.v.vakhabov2019@gmail.com
Where the parameters are determined by the least squares method, i.e. based on the observed experimental data, solving the following (1) system of normal equations:

\[
\begin{align*}
    a_0 T + a_1 \sum t &= \sum y_t \\
    a_0 \sum t + a_1 \sum t^2 &= \sum y_t t
\end{align*}
\]

Using the calculation according to Table 1, we have

\[
\sum Y_t = 3900.1, \quad a_0 = \frac{1}{T} \sum Y_t = \frac{3900.1}{19} = 205.67, \quad a_1 = \frac{2943.8}{570} = 5.17
\]

From here, we find the equation of the linear trend (trend) harvest - potato news:

\[
y(t) = 5.17t + 205.67
\]
\[ y(t) = a_1 t + a_0 \]

- \( H_0: a_1 = 0 \) (null hypothesis)
- \( H_1: a_1 \neq 0 \) (alternative hypothesis)

Accepted significance level \( \alpha = 0.05 \).

Further research is necessary to calculate the following finite differences. Denote:

\[ \Delta \]

Using statistical criteria ([1]), it was found that in equation (2)

\[ Y(t + 1) - Y(t) = \Delta Y_t, \quad \Delta^2 Y_t = \Delta Y_{t+1} - \Delta Y_t, \quad \Delta^3 Y_t = \Delta^2 Y_{t+1} - \Delta^2 Y_t \]

<table>
<thead>
<tr>
<th>Year</th>
<th>( t )</th>
<th>( Y_t )</th>
<th>( t^2 )</th>
<th>( t \cdot Y_t )</th>
<th>( t^2 \cdot Y_t )</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>Year</th>
<th>( t )</th>
<th>( Y(t) )</th>
<th>( Y^2 )</th>
<th>( \Delta Y_t )</th>
<th>( \Delta^2 Y_t )</th>
<th>( \Delta^3 Y_t )</th>
</tr>
</thead>
</table>
\[
V = \frac{\sum_{t=k}^{T} (\Delta^k Y_t)^2}{(T-t)c_{2k}^k}
\]

Reason that there is a significant autocorrelation dependence between \( p \):

\[
d = \frac{\sum_{t=1}^{T-1} (Y_{t+1} - Y_t)^2}{\sum_{t=1}^{T} Y_t^2}
\]

\( d_{\text{observations}} = 0.001 \) and \( d_{\text{crit}} = 1.08 \).

To calculate the autocorrelation coefficient, we use the formula (4), we determine the value of the autocorrelation coefficient \( \rho \) by the formula (3), we compare them with \( \nu \).

\[
R_L = \frac{\sum_{t=1}^{N-L} Y_{t+L} - \frac{\sum_{t=1}^{N-L} Y_t}{N-L} \sum_{t=1}^{N-L} Y_{t+L} \frac{\sum_{t=1}^{N-L} Y_t}{N-L}}{\sqrt{\sum_{t=1}^{N-L} Y_t^2 - \left( \frac{\sum_{t=1}^{N-L} Y_t}{N-L} \right)^2} \sqrt{\sum_{t=1}^{N-L} Y_{t+L}^2 - \left( \frac{\sum_{t=1}^{N-L} Y_{t+L}}{N-L} \right)^2}}
\]

\( R_L: R_1, R_2, R_3, R_4, R_5 \)

**Table 3.**

<table>
<thead>
<tr>
<th>( Y_t )</th>
<th>( Y_t \cdot Y_{t+1} )</th>
<th>( Y_t \cdot Y_{t+2} )</th>
<th>( Y_t \cdot Y_{t+3} )</th>
<th>( Y_t \cdot Y_{t+4} )</th>
<th>( Y_t \cdot Y_{t+5} )</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Based on sample data, using the x7.2019 software package and Excel computers, numerical characteristics $y_t$ for the average potato yield of the region are calculated (Table 4):

**Table 4. Numerical characteristics $y_t$ for theaverage potato yield of the region**

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>Estimates sample characteristics</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Average potato yield</strong> $\bar{y}_T$</td>
<td>205.67</td>
</tr>
<tr>
<td><strong>Variance</strong> $\sigma^2$</td>
<td>923.24</td>
</tr>
<tr>
<td><strong>Mean square deviation</strong> $\sigma_T$</td>
<td>30.38</td>
</tr>
<tr>
<td><strong>Coefficient of variation</strong> $v$ ($%$)</td>
<td>14.77</td>
</tr>
<tr>
<td><strong>Asymmetry</strong> $A$</td>
<td>-0.54</td>
</tr>
<tr>
<td><strong>Excess</strong> $E$</td>
<td>1.19</td>
</tr>
<tr>
<td><strong>Error average value</strong> $\bar{y}_T, m_y$</td>
<td>6.97</td>
</tr>
<tr>
<td><strong>Limit error</strong> $m_y'$</td>
<td>2.11</td>
</tr>
<tr>
<td><strong>Error mean square deviation</strong> $\sigma_T$</td>
<td>6.16</td>
</tr>
<tr>
<td><strong>Interval estimates (95%)</strong> $\bar{y}_T \pm t m_y$</td>
<td>205.67 ± 14.71 (190.96; 220.38) c/ha</td>
</tr>
</tbody>
</table>

**3 Conclusions**

Tashkent region of the Republic of Uzbekistan with a reliability of $\gamma = 0.95$, the following conclusions can be drawn:

1) Point and interval statistical estimates for their sample characteristics are constructed;
2) The explicit type of trend is determined and its linearity is established;
3) The Darbin-Watson criterion established that autocorrelations in the considered series of dynamics have linear trends;
4) Using statistical criteria, it was found that average potato crop of the region forms a non-stationary time series.

**References**


8. Abdullayev, A., Kholturayev, K., Safarbayeva, N. Exact method to solve of linear heat transfer problems. E3S Web of Conferences (2021), 264, 02059. https://doi.org/10.1051/e3sconf/202126402059


