UDC: 911.375:519.23(575.142)

MODELLING THE URBANIZATION LEVEL AND SHORT-TERM FORECAST SCENARIOS FOR KASHKADARYA REGION

Kurbonov Pakhlavon Rustamovich
Karshi State University, Doctor of Geographical Sciences (DSc), Associate Professor

Abstract. *The article models the dynamics of the urban population and the urbanization level of Kashkadarya region (Uzbekistan) for 2009-2025, using seven regression trend models (linear, quadratic, cubic, quartic, power, logarithmic, exponential) and the UN urban-rural growth difference (URGD) method. Models were evaluated by the coefficient of determination (R^2), Fisher's F-test, the Akaike information criterion (AIC) and the out-of-sample error for 2021-2025 (MAPE). The linear model proved the most reliable for urban population (R^2 = 0.9985; MAPE = 0.46%). The urbanization level series fits poorly over the whole period (R^2 = 0.36-0.43) because of statistical fluctuations in 2009-2010; for 2011-2025 the quadratic model gives a stable result (R^2 = 0.88). By 2030 the urban population is projected to reach 1.71-1.76 million, while the urbanization level in the inertial scenario stays at 42.3-42.4%. Reaching 45% by 2030 would require urban population to grow about 1.9 percentage points per year faster than rural population.*

Keywords: *urbanization, urban population, regression model, forecast, coefficient of determination, F-test, URGD, Kashkadarya region.*

МОДЕЛИРОВАНИЕ УРОВНЯ УРБАНИЗАЦИИ И КРАТКОСРОЧНЫХ ПРОГНОЗНЫХ СЦЕНАРИЕВ ДЛЯ КАШКАДАРЬИНСКОЙ ОБЛАСТИ

Курбанов Пахлавон Рустамович

Каршинский государственный университет, доктор географических наук (DSc), доцент

Аннотация. *В статье моделируется динамика численности городского населения и уровня урбанизации Кашкадарьинской области за 2009-2025 гг. с помощью семи типов регрессионных моделей и метода разницы темпов роста городского и сельского населения (URGD) ООН. Модели оценены по коэффициенту детерминации (R^2), критерию Фишера (F), критерию Акаике (AIC) и ошибке ретропрогноза на 2021-2025 гг. (MAPE). Для городского населения наиболее надёжной оказалась линейная модель (R^2 = 0,9985; MAPE = 0,46%). Уровень урбанизации за весь период описывается слабо (R^2 = 0,36-0,43) из-за статистических колебаний 2009–2010 гг.; для 2011–2025 гг. устойчивый результат дала квадратичная модель (R^2 = 0,88). По прогнозу, к*

2030 г. городское население области составит 1,71-1,76 млн человек, а уровень урбанизации в инерционном сценарии останется на уровне 42,3–42,4%. Для достижения 45% к 2030 г. городское население должно расти в среднем на 1,9 п.п. в год быстрее сельского.

Ключевые слова: *урбанизация, городское население, регрессионная модель, прогноз, коэффициент детерминации, критерий Фишера, URGD, Кашкадарьинская область.*

Introduction. Urbanization is a multifactor process reflecting the concentration of population and economic activity in cities, the spread of the urban way of life and the socio-economic modernization of a territory [10; 11; 13]. The "Uzbekistan – 2030" Strategy sets the task of raising the national urbanization level to 60% [7]. However, current trends show that this target cannot be reached evenly across regions: as of 1 January 2026 the national urbanization level was 50.9%, whereas in Kashkadarya region it was only 42.6% [5; 17].

Kashkadarya region has a population of over 3.72 million, a high natural increase (crude birth rate of 27.1‰ in 2025) and a negative net migration balance [17]. Under these conditions, although the urban population is growing in absolute terms, its share is declining. Quantitative modelling of urbanization and scenario-based assessment of its future state are therefore a pressing task for urban development policy.

There are two main approaches to modelling urbanization. The first is the logistic approach based on R. Northam's S-shaped curve and the UN urban–rural growth difference (URGD) method [9; 13; 14]. The second is fitting trend functions to statistical series and selecting them by goodness-of-fit criteria [3; 6]. P. Bocquier showed that the UN model overestimates urbanization in developing countries [2]. In the Russian school of urban studies, G. M. Lappo and Yu. L. Pivovarov emphasised the need to distinguish between the quantitative and qualitative aspects of urbanization [4; 12]. In his doctoral dissertation (Section 3.3), the author built seven types of models for Southern Uzbekistan using data for 2009–2024 [15].

The aim of the study is to update the dissertation models with official data for 2025, to rank them by out-of-sample validation and to develop scenario forecasts of urbanization in Kashkadarya region up to 2030. The objectives are: to re-estimate seven types of trend models; to analyse urban and rural population growth separately; and to compare the inertial and target scenarios.

Literature review and methodology. ***Data.*** The study uses a 17-year series (2009–2025) of the urban population (S, thousand people) and the urbanization level (Y, %) of Kashkadarya region. Data for 2009–2023 were taken from Table 3.9 of the dissertation [15]. For 2024 and 2025, new official data of the Regional Statistics Department were used: as of 1 January 2025 the resident population was 3,639.3 thousand and the urban population 1,553.0 thousand (42.67%); as of 1 January 2026, 3,717.0 and 1,583.8 thousand respectively (42.61%) [16; 17]. The 2024 value given in the dissertation (1.56 million; 42.75%) was replaced with the revised official figure. Each year's value refers to the end-of-year state (1 January of the following year). Total population was reconstructed as $P = S/Y \cdot 100$ and rural population as $R = P - S$.

Trend models. The time variable was defined as t = 1 (2009) … 17 (2025); in the forecast period t = 18 (2026) … 22 (2030). In Appendices 23–24 of the dissertation the calendar year itself (x = 2009 …) was used as the argument; high powers of large numbers make polynomial coefficients unstable, so t is used here instead. Seven types of functions were estimated by ordinary least squares:

$$1)\ y = a_0 + a_1 t; \qquad 2)\ y = a_0 + a_1 t + a_2 t^2;$$

$$3) y = a_0 + a_1 t + a_2 t^2 + a_3 t^3;$$

$$4)\ y = \sum_{i=0}^{4} a_i t^i\ ;$$

$$5) y = a t^b; \qquad 6)\ y = a + b\ln t; \qquad 7)\ y = a e^{bt}$$

Goodness-of-fit criteria. The models were compared using four criteria (n — number of observations, m — number of regressors, RSS — residual sum of squares):

$$R^2 = 1 - \frac{\sum(y_i - \hat{y}_i)^2}{\sum(y_i - \bar{y})^2}, \qquad F = \frac{R^2/m}{(1 - R^2)/(n - m - 1)}, \qquad AIC = n\ln\frac{RSS}{n} + 2(m+1)$$

$$MAPE = \frac{100\%}{h}\sum_{j=1}^{h}\left|\frac{y_j - \hat{y}_j}{y_j}\right|$$

MAPE was calculated for out-of-sample validation: each model was re-estimated on 2009–2020 data, and its forecast for 2021–2025 (h = 5) was compared with the actual values. R^2 and F show the fit to past data, AIC penalises superfluous parameters, and MAPE reflects practical forecast accuracy [1; 3; 6]. For higher-order polynomials R^2 always increases, but extrapolation becomes unstable (overfitting); therefore, the final choice relied on AIC and MAPE.

URGD method. According to the UN methodology, the difference between urban and rural growth rates equals the annual change in the logit of the urbanization level [13; 14]:

$$URGD_t = \ln\frac{S_t}{S_{t-1}} - \ln\frac{R_t}{R_{t-1}} = \ln\frac{Y_t}{100 - Y_t} - \ln\frac{Y_{t-1}}{100 - Y_{t-1}}$$

This formula allows the forecast to be decomposed into its components: once URGD is given, Y and S are uniquely determined. For the total population forecast, the exponential model $P(t) = 2603.4 \cdot e^{0.02086t}$ was adopted ($R^2 = 0.9985$). In the inertial scenario URGD is held at its 2015–2025 average, while in the target scenarios the URGD required to reach a given urbanization level is back-calculated. Calculations were performed in Python (NumPy).

Results and discussion. Urban and rural population growth. In 2009–2025 the urban population of the region increased from 1,140 thousand to 1,583.8 thousand (1.39-fold), with a compound annual growth rate (CAGR) of 2.08%. Over

the same period the rural population grew by 2.28% per year and the total population by 2.19%. As a result, the urbanization level fell from 43.41% to 42.61%. URGD was negative in 14 of the 16 years of 2010–2025; its 2015–2025 average was −0.17 percentage points. The direct cause of the decline in the urbanization level is the faster growth of the rural population driven by natural increase.

Quality of the trend models. Table 1 presents the indicators of the seven models for the whole period (2009–2025) and their forecasts for 2030.

Table 1.

Goodness-of-fit indicators of the trend models for the urban population (S) and urbanization level (Y) of Kashkadarya region, 2009–2025

Model	S: R^2	S: F	S: MAPE, %	S 2030, thous.	Y: R^2	Y: F	Y: MAPE, %	Y 2030, %
Linear	0.9985	9802.6	**0.46**	**1710.9**	0.361	8.47	0.47	42.59
Quadratic	0.9987	5529.9	1.05	1725.6	0.422	5.10	0.29	42.18
Cubic	0.9992	5580.6	0.80	1779.0	0.425	3.20	0.07	41.94
Quartic	0.9992	3875.1	5.19	1789.0	0.427	2.23	4.11	42.52
Power	0.8992	133.9	6.70	1546.9	0.260	5.27	0.59	42.80
Logarithmic	0.8703	100.7	7.01	1537.6	0.260	5.28	0.59	42.80
Exponential	0.9980	7533.4	0.82	1754.9	0.360	8.45	0.48	42.59

Note: *MAPE –forecast error for 2021–2025 of the model estimated on 2009–2020 data. Critical value of F (α = 0.05; m = 1, n = 17) ≈ 4.54. Source: author's calculations.*

For the urban population, all polynomial and exponential models satisfy the conditions $R^2 > 0.998$ and $F \geq F_{crit}$. However, in out-of-sample validation the linear model produced the smallest error (0.46%), whereas the error of the quartic polynomial rose to 5.19%. The cubic model has the lowest AIC (52.7), but it implies accelerating growth towards 2030. Therefore, the linear model was adopted as the baseline for the urban population and the exponential model as the upper bound:

$$S(t) = 1111.77 + 27.234\,t \quad (R^2 = 0.9985), \qquad S(t) = e^{7.0266 + 0.02016\,t} \quad (R^2 = 0.9980)$$

The urbanization level series is poorly approximated over the whole period: R^2 ranges from 0.26 to 0.43, and only the linear, quadratic, power, logarithmic and exponential models satisfy $F > F_{crit}$. This is caused by the high value in 2009 (43.41%) and the sharp drop in 2010 (42.50%). These years coincide with the period of statistical recalculation that followed the conversion of rural settlements into urban-type settlements. When re-estimated for 2011–2025, model quality improved sharply: linear $R^2 = 0.879$ ($F = 94.3$), quadratic $R^2 = 0.883$ ($F = 45.1$), cubic $R^2 = 0.926$ ($F = 46.0$), quartic $R^2 = 0.932$ ($F = 34.0$). The cubic model yields an accelerating decline to 41.48% by 2030, which is inconsistent with the decelerating dynamics of recent years. Therefore, the quadratic equation was selected as the baseline model:

$$Y(t) = 43.350 - 0.02862\,t - 0.00072\,t^2 \quad (2011–2025;\ R^2 = 0.883;\ F = 45.1)$$

According to this model, the urbanization level will be 42.60% in 2026, 42.49% in 2028 and 42.37% in 2030.

URGD scenarios. Table 2 presents the 2030 results calculated under different URGD values (total population in 2030: 4,120.6 thousand).

The inertial scenario (42.40%) almost coincides with the quadratic trend model (42.37%) and with the cubic-model forecast in the dissertation (42.2%) [15]. The agreement of three independent approaches increases the reliability of the forecast. Raising the region's urbanization level to 60% by 2030 would require a URGD of +14.1 percentage points per year, which could only be achieved through administrative reclassification.

Table 2.

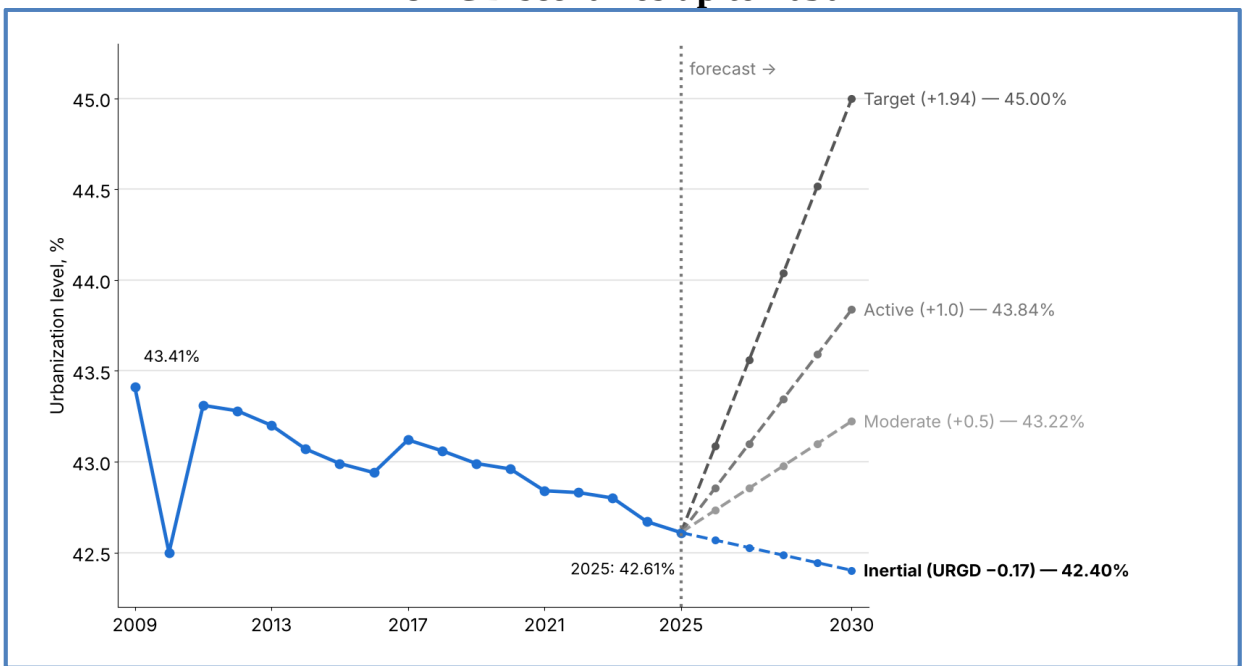
Scenario forecast of urbanization in Kashkadarya region up to 2030

Scenario	URGD, pp per year	Y 2030, %	S 2030, thousand people
Inertial (2015–2025 average)	−0.17	42.40	1747
Share stabilization	0.00	42.61	1756
Moderate acceleration	+0.50	43.22	1781
Active urban development policy	+1.00	43.84	1807
Target: 45%	+1.94	45.00	1854

Source: *author's calculations*.

Factor model. The results of the multiple regression built in STATA 14 for the dissertation confirm this conclusion [15]. Urban population size (+0.0371; $p < 0.001$), population density (+0.664; $p = 0.097$) and urban labour resources (+0.0037; $p = 0.091$) have a positive effect on urbanization, whereas total population has a negative effect (−0.0389; $p = 0.024$). Natural increase, net migration and the employment rate proved statistically insignificant ($p > 0.1$).

Figure 1.
Dynamics of the urbanization level in Kashkadarya region (2009–2025) and URGD scenarios up to 2030



Source: *Table 3.9 of the dissertation; Kashkadarya Regional Statistics Department (2025, 2026); author's calculations.*

The results show that Kashkadarya is experiencing "urban growth without urbanization": the urban population grows by about 27 thousand people per year on average, yet its share is declining. Such a pattern is typical of regions with high natural increase in rural areas and weak migration to cities [2; 10]. In 2024–2025 the region's net migration remained negative (−6.7 and −6.4 thousand, respectively), while natural increase amounted to 84–85 thousand per year [16; 17]. Hence, rural-to-urban migration is not the main source of urban population growth.

Two methodological conclusions are important. First, a high R^2 does not guarantee forecast accuracy. Although the quartic polynomial has the highest R^2 for the urban population, its out-of-sample error was 11 times larger than that of the linear model. Second, the value $R^2 = 0.9639$ reported in the dissertation for the urbanization level does not match the calculation in Appendix 24 ($R^2 = 0.2916$) [15]. Recalculation confirmed the appendix value (0.26–0.43 for the whole period); a high fit is achieved only by excluding 2009–2010. Therefore, rather than regressing share indicators directly on time, modelling them through their components (S and R), i.e. the URGD approach, is better founded [13; 14].

It is also useful to compare the forecast results with planning documents. Under Resolution No. 105 of the Cabinet of Ministers of 24 February 2025, the population of Karshi city is planned to reach 470 thousand by 2042 and its area to double [8]. If this plan is implemented, URGD could rise to the moderate-scenario level (+0.5) owing to the Karshi agglomeration. However, reaching 45% at the regional level would require the urban population to grow by about 4.1% per year, twice the current rate.

Limitations of the study: (1) the series is short (17 years), which limits the forecast horizon to 5 years; (2) administrative changes in the network of urban

settlements are not accounted for in the model; (3) the factor model is based on aggregate regional data. Future research should use district-level panel data and spatial regression.

Conclusions and recommendations. By 2030 the urban population of Kashkadarya region will increase to 1.71–1.76 million, while the urbanization level will remain at 42.3–42.4% if current policy is maintained.

- The linear model describes urban population dynamics most reliably (R^2 = 0.9985; MAPE = 0.46%). Higher-order polynomials are superior in terms of R^2 but unstable in forecasting.

- For the urbanization level, the quadratic model based on 2011–2025 (R^2 = 0.883) and the URGD inertial scenario give the same result: 42.37% and 42.40%, respectively, in 2030.

- The main causes of the decline in urbanization are the faster growth of the rural population (2.28% versus 2.08%) and negative net migration. The core problem is the low migration attractiveness of cities.

- To reach 45% by 2030, the urban population of the region must grow on average 1.94 percentage points per year faster than the rural population. The national target of 60% cannot be achieved through natural processes alone.

Recommendations:

1. Set urbanization targets in urban development programmes in terms of URGD rather than the urban share, and monitor them annually.

2. Bring net migration close to zero by expanding employment and affordable housing in the Karshi agglomeration (Karshi, Koson, Beshkent and satellite towns) and in the Shahrisabz–Kitob system.

3. Make an inventory of urban-type settlements that lack an established town-forming economic base, and develop industry and services in them.

4. Make AIC and out-of-sample validation (MAPE) mandatory criteria, alongside R^2 and F, when selecting forecast models.

References

1. Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, *19*(6), 716–723. https://doi.org/10.1109/TAC.1974.1100705
2. Bocquier, P. (2005). World urbanization prospects: An alternative to the UN model of projection compatible with the mobility transition theory. *Demographic Research, 12*(9), 197–236. https://doi.org/10.4054/DemRes.2005.12.9
3. Draper, N. R., & Smith, H. (1998). *Applied regression analysis* (3rd ed.). Wiley.
4. Lappo, G. M. (1997). *Geografiya gorodov [Geography of cities]*. VLADOS. (In Russian)
5. Statistics Agency under the President of the Republic of Uzbekistan. (2025). *Demographic data*. https://stat.uz
6. Hyndman, R. J., & Athanasopoulos, G. (2021). *Forecasting: Principles and practice* (3rd ed.). OTexts. https://otexts.com/fpp3/
7. Decree of the President of the Republic of Uzbekistan. (2023, September 11). *On the "Uzbekistan – 2030" Strategy* (No. PF-158). https://lex.uz
8. Cabinet of Ministers of the Republic of Uzbekistan. (2025, February 24). *On approval of the master plan of Karshi city* (Resolution No. 105). https://lex.uz

9. Northam, R. M. (1979). *Urban geography* (2nd ed.). John Wiley & Sons.
10. Henderson, V. (2003). The urbanization process and economic growth: The so-what question. *Journal of Economic Growth*, *8*(1), 47–71. https://doi.org/10.1023/A:1022860800744
11. United Nations, Department of Economic and Social Affairs, Population Division. (2019). *World urbanization prospects: The 2018 revision* (ST/ESA/SER.A/420). United Nations.
12. Pivovarov, Yu. L. (2001). *Osnovy geourbanistiki: Urbanizatsiya i gorodskie sistemy [Fundamentals of geourbanistics: Urbanization and urban systems]*. VLADOS. (In Russian)
13. United Nations. (1974). *Manual VIII: Methods for projections of urban and rural population* (Population Studies No. 55). United Nations.
14. United Nations, Department of Economic and Social Affairs, Population Division. (2018). *World urbanization prospects 2018: Methodology* (Working Paper No. ESA/P/WP.252). United Nations.
15. Qurbonov, P. R. (2025). *Janubiy O‘zbekistonda urbanizatsiya jarayonlarining hududiy tarkibini takomillashtirish* [Improving the territorial structure of urbanization processes in Southern Uzbekistan] (Doctoral (DSc) dissertation in Geography). Tashkent. 254 p. (In Uzbek)
16. Kashkadarya Regional Department of Statistics. (2025). *Demographic situation as of 1 January 2025* [Press release; in Uzbek]. https://qashstat.uz/uploads/press24/DM-PR-UZ-12-2024.pdf
17. Kashkadarya Regional Department of Statistics. (2026). *Demographic situation as of 1 January 2026* [Press release; in Uzbek]. https://qashstat.uz/uploads/press25/DM-PR-UZ-12-2025.pdf
18. Qurbonov P. R. IQLIM O’ZGARISHI SHAROITIDA TEKISLIK MINTAQASI SHAHARLARINING RIVOJLANISH XUSUSIYATLARI (JANUBIY O’ZBEKISTON MISOLIDA) //Экономика и социум. – 2023. – №. 10 (113)-1. – С. 210-214.
19. Qurbonov P. R., Qo’ldosheva G. M. Ecological-geographical problems in the residential areas of kashkadarya region //Экономика и социум. – 2025. – №. 5-1 (132). – С. 617-619.
20. Qurbonov P. O ‘zbekistonda urbanizatsiya jarayonlari rivojlanishining asosiy bosqichlari //Farg'ona davlat universiteti. – 2023. – №. 3. – С. 165.
21. Qurbonov P. R. Janubiy O ‘zbekiston yer usti tuzilishining urbanizatsiya jarayonlari va shaharlar taraqqiyotiga ta’siri //SCIENCE and SOCIETY. – 2025. – №. 5-1. – С. 40-42.