Analysis and Forecast of the Social Sphere Development in the Arctic Regions of the Russian Federation within the Framework of Industry 4.0 Exemplified by Murmansk Region

S K Antipov, K N Kikkas, L K Bobodzhanova and E E Krasnozhenova
Peter the Great St. Petersburg Polytechnic University, St. Petersburg, Russia

Abstract. The given article considers the main strategies of the social attractiveness of the Russian Federation Arctic Zone in the context of Industry 4.0 concept presented on the example of the Murmansk region. The research is devoted to the main factors affecting the public interest in the region development as well as the factors that reflect technological development and the intensity of technology implementation. The relationships of these factors are analyzed and the model that describes obtained relationships is presented. Based on the obtained model, the forecast is made and the assessment of the development prospects of this region is provided.

1. Introduction
Considering the prospects for sustainable development of the Arctic regions in the Russian Federation, it should be noted that the possibility of its implementation will depend on maintaining the right balance between the main spheres of human activity. In the classical understanding of sustainable development, there are three areas, the interaction of which determines the achievement of an optimal result: economy, ecology, and society. However, taking into account the era of digital transformation and the Fourth Industrial Revolution (Industry 4.0), such approach will be insufficient. Big data has become an integral part of the contemporary world and the value of information is constantly growing. The development of technologies allows saving resources, replacing people at workplaces that are dangerous to their health and life. High-tech products facilitate both work and everyday life. Advanced information technologies make it possible to build accurate models that may predict the future with high accuracy. Neural networks are able to make operational decisions and control equipment. One of the key ideas of the new concept “Industry 4.0” is not just the introduction of technologies into most areas of human activity, but the creation of an integrated information technology environment that allows modeling and controlling global development as well as individual local areas. At the conceptual level, such integration is not a novelty; however, its implementation may become available only now since the general level of development of scientific and technological progress and engineering allows introducing these ideas. Cloud services aggregating data of different scales and levels are actively being developed and used by ordinary consumers and government agencies. Undoubtedly, this approach will be beneficial both for the population and for the authorities. It allows being considered at the national level as well as at the regional, municipal and territorial levels. Thus, while describing the concept of sustainable development, highlighting the fourth component that reflects the technological sphere, is a must. This approach involves developing the model of four pillars of sustainable development (Figure 1). Such
understanding provides completeness and integrity, which, in turn, makes it possible to conduct the high-quality and accurate research, and to provide reliable forecasts. All this will give a guarantee of making the right management decisions and determining the right vector in the field of national development.

![Figure 1. Model of four pillars of sustainable development](Source: composed by the authors)

Taking into account each of the areas, it can be noted that the development in the field of economics will allow achieving the most efficient use of resources and employment, which is particularly important in the context of increasing needs and population growth. Reducing unemployment and increasing welfare will make it possible to move to a higher level of quality of life and that will positively impact the development of social sphere. The tasks include preserving cultural heritage and observing human rights, eliminating any inequalities, and developing such society that will treat natural resources, consumption, and labor in a responsible way.

As for the level of ecology is concerned, this area is one of the most significant to achieve sustainable development. The general growth in the population as well as the growth of industrial capacities, strongly affects the climate, which is especially noticeable in the Far North and in the Arctic regions. In recent years, an increase in annual temperatures and a decrease in glacier-covered areas have been observed. It is necessary to understand the importance of environmental concern, both in terms of maintaining the integrity of the ecosystem, and in terms of the rational use of resources. The annually growing threat of global warming and critical climate change, especially in the Northern regions, manifests the importance of this area.

In technology, it is worth considering the impact of scientific and technological progress on people, assessing the factors that contribute to improving the quality of production, developing and using of advanced production technologies, and implementing innovations. Undoubtedly, in the era of the Fourth Industrial Revolution, this area dominates and allows achieving sustainable development. The widespread introduction of modern digital technologies and the formation of an integrated digital space makes is possible to combine all analyzed areas in real time and to get reliable and relevant information for separate factors as well as to observe their relationships in dynamics.

The implementation of sustainable development in the Arctic regions of the Russian Federation obeys general conceptual provisions; however, due to the special climate and geological conditions in the Arctic Zone and the special geopolitical interest in it, it has some subtle aspects. One of the key tasks aimed at achieving development is the formation of the most comfortable social environment. In other words, it makes sense to analyze and identify the factors that contribute to attracting the population to
the Arctic regions, stopping migration outflows, describing opportunities and prospects, compensating for significant difficulties and ensuring a proper standard of living. At the same time, advances in technology make it possible to increase the living standard, which was previously unavailable. For the Russian Federation, sustainable development is also the most promising way of conducting economic activity [1]. To a greater extent, this applies specifically to the regions of the Arctic Zone.

In the described research, a quantitative assessment of the population in the Murmansk region is carried out to identify the main factors that affect the population dynamics. Revealing the main reasons for these changes will allow making the most accurate management decisions. Taking into consideration the strategic importance of the Arctic regions, making right decisions is very important, and therefore, only reliable data will minimize possible errors. All the above mentioned shows the necessity and relevance of the given study.

2. Literature review
Implementing sustainable development is now a matter of utmost importance. The goals adopted by the UN in 2015 focus mainly on social equality, poverty eradication, environmental care and refocusing industry towards environmental safety. It is also obvious that achieving these goals becomes impossible without the introduction of advanced technologies and innovations.

A number of studies are devoted to the theory of sustainable development. S. Baker explains the necessity for sustainable development and its quantitative assessment in his book “Sustainable Development” [2]. Prospects for sustainable development based on mathematical modeling are described by Antipov S.K., Didenko N.I. [3]. Khaidukov D.S. and Tasalov K.A. in their articles “Fundamentals of Ensuring Sustainable Development of Urban Agglomeration” and “Implementation of the Concept of Sustainable Development in Regional Management” assess the possibility of sustainable development in the regional context [4]. Skripnuk D. draws a particular attention to the socio-economic development of the Arctic regions in the Russian Federation in the article "Analysis of the Regional Modernization Processes in a Global Context with an Example of the Russian Northern Regions" [5].

Didenko N.I. [6] and Skripnuk D.F. [7] focus on modeling of sustainable development in their works. The issue of choosing quantitative indicators to assess the sustainable development of regions is described in the works of Didenko N. [8], Perelet P.A. [9], Didenko N. and Kunze K. [10]. Mensch G. [11], Schumpeter, J.A.and Drucker, P.F. [12, 13] discuss in their studies innovations in the field of economic and social development.

In almost all the works described above, the emphasis is on the analytics of the social sphere, the assessment of the impact of technology on humans and the environment. Michálle E. Mor Barak [14] and Djamilia S. [15] underline the importance of social development and the formation of new common good.

The assessment of the development prospects within the concept of Industry 4.0 is well described in the article written by Vasja Roblek, Maja Meško, Alojz Krapež [16]. Dominique Lepore, Francesca Spigarelli [17] and Klochkov Y., Klochkova E. [18] consider in detail regional development, taking into account technological and innovative changes.

3. Research methods and basic data
For a quantitative assessment of the prospects for the social sphere development of the Arctic regions in the Russian Federation in the framework of Industry 4.0, it is necessary to assess accurately the most significant factors influencing both society and scientific and technological progress, since thanks to them an increase in living standards may become possible. In addition, it is necessary to assess significant interconnections between these areas.

The procedure for analysis and forecasting techniques include the following sequential steps:
1. Identifying key factors
2. Selecting a region
3. Selecting and developing the model

4. Conclusion
The results of the study show the need for significant changes in the current state of the Arctic regions, taking into account the prospects for sustainable development. In the Arctic zone, actions are required to increase the quality of life and the development potential of the region. The implementation of Industry 4.0 technologies is a key factor in achieving these goals. The results of this research can be used as a basis for strategies and policies aimed at sustainable development of the Arctic regions.
4. Collecting statistical information
5. Training the model based on the received data
6. Forecasting

The choice of the model is one of the vital steps on which reliable prognostic results depend. Regression models or “random forest” models are better to be used as the models reflecting relationships between a large number of variables. The latter models are capable to provide more accurate results if a large set of test data is available. In other words, if there is an opportunity to obtain large statistical sets of data, then these models should be chosen. Analyzing the development of the Arctic regions, researchers encounter difficulties in finding these data; moreover, as a rule, the available statistics is extremely poor. This dictates the choice of regression models.

The regression model is capable to provide good modelling of the relationships of factors, but it is impossible to obtain forecasts based on the influencing factors. To solve this problem, a two-step modeling method should be used: the first step includes making one-dimensional forecasts in complying with the influencing factors, the second step includes assessing their complex influence on the resulting variables and making the final forecast. As one-dimensional models, it is best to adapt such models as ARMA (ARIMA) or SARMA (SARIMA). Formulas 1 and 2 show the general view of the regression model and the extrapolating models.

\[
y_t = a_0 + a_1 \cdot x_1^t + a_2 \cdot x_2^t + \cdots + a_n \cdot x_n^t (1)
\]

\[
\hat{y}_t = a_0 + a_1 \cdot \hat{x}_1^t + a_2 \cdot \hat{x}_2^t + \cdots + a_n \cdot \hat{x}_n^t (2)
\]

\(y_t\) – a resulting indicator
\(x_1^t, x_2^t, \ldots, x_n^t\) – influencing factors
\(\hat{y}_t\) – extrapolation of a key indicator based on the regression model
\(\hat{x}_1^t, \hat{x}_2^t, \ldots, \hat{x}_n^t\) – extrapolation of factors based on such model as ARMA or its derivatives
\(a_0, a_1, a_2, \ldots, a_n\) – regression model parameters

To obtain an adequate result, the model should include only those factors that are significantly linearly correlated with the resulting indicator and that correspond to stationary or integrated time series.

The final type of the prognostic model by factors will depend on the degree of autocorrelation of each time series, which is most conveniently estimated based on the autocorrelation function. The window of moving average is selected by comparing possible patterns and choosing the one that provides the best result that is measured by the Akaike Information Criterion (AIC).

One of the Arctic regions, namely the Murmansk region, has been chosen to test the proposed methodology and confirm the chosen concept. Social and technological spheres have been selected as the most significant. The key factor that assesses the social attractiveness of the region is the population size. It is evident that the region with a high level of wages, proper living standards and security, a favorable ecological situation will contribute to population growth due to increasing the migration balance, increasing life expectancy, reducing mortality and increasing birth rates. To assess the level of economic development and activity, it is best to choose the gross regional product. The development in the technological area is properly viewed from the point of the development and implementation of advanced production technologies. These indicators reflect how significant the potential for technological and innovative development in the particular region is. Figure 2 shows changes in the selected indicators. It can be seen that in recent years, despite the increase in the number of developed and implemented advanced production technologies, the population in the region is decreasing.
To determine the key factors affecting the population size, a number of indicators that should influence the population change directly or indirectly have been selected. The indicators have been divided into subgroups to interpret obtained results more vividly.

- **Social sphere**
  - **Income**
    - Average salary – s11
    - Share of the population with income below the minimum wage – s12
    - Average size of assigned pensions – s13
    - Real monetary income of the population – s14
  - **Costs**
    - Consumer spending per capita – s21
    - Share of expenses for housing and communal services – s22
    - Number of passenger cars – s23
  - **Labor**
    - Unemployment rate – s31
    - Labor market tensions – s32
    - Number of non-resident employees – s33
  - **Living conditions**
    - Number of dilapidated housing – s41
    - Number of commissioned apartments – s42
    - Total area of residential premises, per person – s43
    - Number of commissioned buildings – s44
  - **Demographic factors**
    - Mid-year population, per person – s51
    - Working-age mortality – s52
    - Average life expectancy – s53
    - Birth rate – s54
    - Migration rate – s55
    - Demographic load factor – s56
  - **Education**
    - Number of places in DOW – s61
    - Number of completed secondary vocational education – s62
    - Number of completed HEU – s63
  - **Health care**
    - Number of cases per 1000 people – s71
    - Capacity of outpatient and poly-x organizations – s72
    - Population per hospital – s73
  - **Leisure**
    - Number of museum visitors – s81
• Number of theatre visitors – s82
• Number of children camps – s83
• Volume of tourist service to the population – s84
  o Sport
    • Number of sports stadiums, over 1500 places – s91
    • Number of sports grounds and fields – s92
    • Number of sports gyms – s93
    • Number of swimming pools
  • Technology
    o Science and innovation
      • Innovation activity – t11
      • Innovative activity in the field of technological innovation – t12
      • Internal research and development costs – t13
      • Technological innovation costs – t14
      • Used advanced manufacturing technologies – t15
      • Developed advanced manufacturing technologies – t16
      • Organizations that perform research and development – t17
      • Issuance of utility model patents – t18
      • Grant of patents for inventions – t19
      • Number of researchers – t110

Statistical data have been collected for the selected groups within the period of 13 years (from 2008 to 2019).

4. Results
Taking into account the correlation between the indicators, it can be noted that the general hypothesis that concerns the choice of factors is correct. Figure 3 confirms the heat map of correlations.
In order to develop the most balanced model that would include the greatest number of directions and at the same time would not be overloaded, it is better to leave the most correlated factor from each group.

It should be noted that multicollinearity is revealed between some of the selected factors, however, due to its purely statistical nature, none of them will be excluded. Thus, Formula 3 presents the general view of the regression model for the Murmansk region.

\[
y_1 = a_0 + a_1 \cdot s_{11} + a_2 \cdot s_{21} + a_3 \cdot s_{33} + a_4 \cdot s_{43} + a_5 \cdot s_{55} + a_6 \cdot s_{63} + a_7 \cdot s_{72} + a_8 \cdot s_{94} + a_9 \cdot s_{92} + a_{10} \cdot t_{13} + a_{11} \cdot t_{16}
\]  

(3)

To illustrate the choice of factors, paired regression models between them and the resulting indicator of the population size have been built (Figure 4).
Figure 4. Paired linear regression models due to the most correlated factors for social and technological spheres

Based on the obtained final model (Formula 4), it may be concluded that the population size is mostly affected by the total area of living quarters, the number of people who graduated from higher educational institutions, the capacity of outpatient organizations and the number of developed advanced production technologies.

\[
y_1 = 4351.38 - 1.97 \cdot s_{11} + 8.40 \cdot s_{21} + 11.08 \cdot s_{33} + 84180.23 \cdot s_{43} - 995.77 \cdot s_{55} + 12490.00 \cdot s_{63} - 29750.10 \cdot s_{72} + 5.64 \cdot s_{84} - 558.46 \cdot s_{92} + 8.73 \cdot t_{13} + 2603.40 \cdot t_{16} - 4940.61 \cdot t_{17}
\]  

(4)

The obtained model has the coefficient of determination of 0.98, and the coefficients are significant at the level of 0.05. It indicates the adequacy of the model and the possibility of its further application for forecasting.

The second part of the modeling is to build one-dimensional autoregressive models for forecasting based on factors. Initially, the assessment of autocorrelation and partial autocorrelation functions was carried out in order to identify the components of the time series.

Figure 5 shows the construction of ACF and PACF for the indicator "average wages".
Figure 5. Autocorrelation and partial autocorrelation functions for the indicator "average wages" in the Murmansk region (2008-2019)

To make a factorial forecast, it is necessary to clearly identify the shape of the selected components of the series and determine the optimal parameters of the ARIMA (SARIMA) model. Figure 6 shows the decomposition of the time series "average wages" starting from 2013.

Figure 6. Decomposition by the components of the time series "average wages" in the Murmansk region

Based on the comparison of the results for autocorrelation models till lag 20, it was found that the model with the lag bias parameters 1, windows of moving average 2, and the cycle period 12 has the lowest Akaike Information Criterion (AIC) which is equal to 875.06. Figure 7 shows the comparison of
predicted and actual data, as well as the extrapolation for a year. The graphs also show the corridors of confidence intervals for the obtained model.

![Graphs showing predicted and actual data with confidence intervals.]

**Figure 7.** Comparison of simulated and actual values (above) and the forecast due to the obtained model for one year (below)

In a similar way, the models and the extrapolation for all other factorial values have been built. After that they have been substituted into the obtained regression model. Thus, the predicted value of the population for 2020 is equal to 736753, which is 1.07% lower than in 2019, and the predicted value for 2021 is equal to 730808 people, which is 0.81% lower than for 2020.

5. **Discussion**

Analyzing the obtained data, it is worth mentioning that despite the fact that the population will continue to decrease, the rate of change will also become lower, which indicates a positive trend. The factors under consideration really have a significant impact on the population size, and in order to improve the situation in the region, a particular attention should be paid to housing conditions and the health sector. In addition, it is necessary to develop the education sector by increasing the number of higher educational institutions or increasing the capacity of existing ones as well as achieving competitiveness in relation to educational institutions in other regions. Moreover, a key issue is the development of advanced production technologies, since within the conditions of severe climate these factors are able to increase living standards significantly.

A lot of scientists express such opinions, however, to a greater extent, their assessments are rather of a qualitative nature and they are based on their expertise and logic. The developed model gives a clear and quantitative description of the main influencing factors.

What is more, the considered two-step model of the development of the social attractiveness in the Murmansk region provides positive prognostic results and it can be used to predict values in the short term. Undoubtedly, like in any other mathematical model, the predicted values depend on more factors than those included in the model. It is worth noting that the obtained forecast, including 2021, may deviate from real values, mainly due to the factors associated with the current epidemiological situation of COVID-19. For a more reliable prediction of changes, the considered model should include the indicators characterized by up-to-date data based on the level of morbidity and mortality in a pandemic.

Model development according to the proposed methodology is considered to be a completely universal method both to identify the key factors affecting the resulting indicators, and to obtain highly accurate prognostic results that can make it possible to form the most holistic and accurate understanding of changes in the studied processes or phenomena. Further research will be devoted to the construction of a three-step model, in which it would be reasonable to include not only correlation and autocorrelation
connections, but also to evaluate cross-correlations. This will provide an even higher accuracy of the results as well as broaden the understanding of possible factorial influences with lagged delays. In addition, the linear approach is not always capable to give an exhaustive description of complex economic and social processes, which makes it possible to form more flexible nonlinear models.

References

[1] Report on Human Development in the Russian Federation for 2016 /ed. S.N. Bobyleva and L.M. Grigoriev. - M. : Analytical Center for the Government of the Russian Federation, 2016. 298 p.(in Rus).

[2] Baker, S. 2006. Sustainable Development. New York, NY: Routledge

[3] Antipov, S.K., Didenko, N.I., Lobatyuk, V.V., Sosnina, M.N. Comparing Models of Sustainable Development for the Arctic Region of the Russian Federation IOP Conference Series: Earth and Environmental Science, 2020,539(1), 012102

[4] Khaidukov D.S., Tasalov K.A. Fundamentals of Ensuring Sustainable Development of Urban Agglomeration // Materials of the XIV International Conference of Moscow State University "Public Administration in the XXI Century", - M. : "University book", 2017. 783 p.—789ISBN 978-5-91304-707-6 (in Rus).

[5] Skripnuk D.: Analysis of the Regional Modernization Processes in a Global Context with an Example of the Russian Northern Regions. Int. J. Syst. Assur. Eng. Manag. 11, pp 100–110 (2020). https://doi.org/10.1007/s13198-019-00895-6.

[6] Didenko N. (2020) Modeling the Global Nickel Market with a Triangular Simultaneous Equations Model. International Journal of Systems Assurance Engineering and Management 11, pp 119-129.

[7] Skripnuk D., Kikkas K., Romashkina E. (2019) Sustainable Development and Environmental Security in the Countries of the Circumpolar North. E3S Web of Conferences. 110, 02037.

[8] Didenko N.I. Concept and Mechanism for Managing the Digital Transformation of Arctic Target Subspaces and Spheres of Vital Activity. IOP Conference Series: Materials Science and Engineering, 2020, 940 (1), 012121.

[9] Perelet R.A. Revealing Indicators of Sustainable Development // Problems of the Environment and Natural Resources. VINITI //1995.– №6 (in Rus).

[10] Didenko N., Kunze K. The Relationship Between Energy Policies and Quality of Life in Russia. WIT Transactions on Ecology and the Environment, 2014, 186, pp. 3–11.

[11] Mensch G. Stalemate in Technology: Innovation Overcome the Depression. Cambridge (Mass), 1979, p. 14-17.

[12] Schumpeter, J.A. The Theory of Economic Development. – Moscow: Directmedia Publishing, 2008. – 401 p

[13] Drucker, P.F. Innovation and Entrepreneurship/ edited by K. S. Golovinsky . – Moscow; SPb; Kiev: Williams, 2007. – 423 p. - ISBN 5-84591-195-7.

[14] Michalle E. Mor Barak The Practice and Science of Social Good: Emerging Paths to Positive Social Impact Research on Social Work Practice, vol. 30, 2: pp. 139-150.,First Published January 3, 2018.

[15] Djamila S., Kuzava A., Glushkova A. (2020) Formation of a Database of Indicators and Analysis of the Environmental and Socio-Economic Vital Activity Spheres of Russian Federation Arctic Zone Municipalities. International Journal of Systems Assurance Engineering and Management. 11, pp. 19-43.

[16] Vasja Roblek, Maja Meško, Alojz Krapež A Complex View of Industry 4.0 SAGE Open, vol. 6, 2, First Published June 1, 2016.

[17] Dominique Lepore, Francesca Spigarelli Integrating Industry 4.0 Plans Into Regional Innovation Strategies Local Economy, vol. 35, 5: pp. 496-510.,First Published July 8, 2020.

[18] Klochkov Y., Klochkova E., Didenko N., Frolova E., Vlasova N. (2018) Development of Methodology for Assessing Risk of Loss of a Consumer through the Fault of an Outsourcer.
International Conference on Infocom Technologies and Unmanned Systems: Trends and Future Directions, ICTUS 2017, pp. 719-724.