sustainability

MDPI

Article

Artificial Neural Networks as a Method for Forecasting Migration Balance (A Case Study of the City of Lublin in Poland)

Adam Gawryluk [1], Agnieszka Komor [2,*], Monika Kulisz [3], Patrycjusz Zarębski [4] and Dominik Katarzyński [4]

[1] Department of Landscape Studies and Spatial Management, Faculty of Agrobioengineering, University of Life Sciences in Lublin, Akademicka 13, 20-950 Lublin, Poland; adam.gawryluk@up.lublin.pl
[2] Department of Management and Marketing, Faculty of Agrobioengineering, University of Life Sciences in Lublin, Akademicka 13, 20-950 Lublin, Poland
[3] Department of Organisation of Enterprise, Management Faculty, Lublin University of Technology, Nadbystrzycka 38, 20-618 Lublin, Poland; m.kulisz@pollub.pl
[4] Department of Economics, Koszalin University of Technology, Kwiatkowskiego 6E, 75-343 Koszalin, Poland; patrycjusz.zarebski@tu.koszalin.pl (P.Z.); dominik.katarzynski@tu.koszalin.pl (D.K.)
* Correspondence: agnieszka.komor@up.lublin.pl

Abstract: Internal migration regulates both the size and structure of human resources and affects the labor market at different spatial scales. It therefore has not only a demographic dimension, but also a spatial one, which is why it can significantly affect development on both a local and regional scale. The main objective of this study was to examine the usefulness of artificial neural networks (ANN) for predicting the internal migration balance for the city of Lublin in Poland. Another objective was to develop an experimental neural network model for forecasting the internal migration balance for the city of Lublin (for one year ahead) based on selected economic and social factors. The study area included the city of Lublin and 14 municipalities located in the vicinity of the city and functionally connected to it (they form the Lublin Functional Area), i.e., a total of 15 spatial units. Data for the analysis covered the years 2005–2022 and were obtained from the Local Data Bank (BDL) of the Central Statistical Office (GUS). The number of input variables for the ANN model was reduced using principal component analysis (PCA), allowing for the inclusion of the most relevant demographic and economic features. These components can thus be considered reliable predictors of the migration balance for the city of Lublin. This suggests that artificial neural networks may be an effective tool in supporting decision-making processes for forecasting the migration balance of this city.

Keywords: artificial neural networks; suburbanization; migration prediction; population migrations

check for updates

Citation: Gawryluk, A.; Komor, A.; Kulisz, M.; Zarębski, P.; Katarzyński, D. Artificial Neural Networks as a Method for Forecasting Migration Balance (A Case Study of the City of Lublin in Poland). *Sustainability* **2024**, *16*, 11249. https://doi.org/10.3390/su162411249

Academic Editors: Stephen Drinkwater and Stephan Weiler

Received: 29 October 2024
Revised: 3 December 2024
Accepted: 20 December 2024
Published: 22 December 2024

Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

Population migration, which involves the movement of people between different locations, is an important demographic process that affects society and the economy in various ways, while simultaneously shaping changes in the functioning of places. Migration is "a major symptom of basic social change" [1]. Migration is one of the key components of demographic change (along with fertility and mortality). It is crucial not only to understand the current demographic situation in a given area, but also to be able to forecast the size and structure of the population in the future. This ability provides insight into numerous social, economic, spatial, and environmental consequences resulting from population changes, enabling effective development policy in a given area, as well as managing development at the regional and local levels, including strategic planning.

Ravenstein [2] is considered a pioneer in migration studies, having formulated seven laws of migration. He demonstrated that most migrants travel only short distances, while longer migration distances are typically associated with movement toward large commercial and industrial centers. There is an inverse relationship between the impact of

immigration and emigration centers and the distance involved. Every migration flow generates a counter-flow, and women as well as rural populations are more inclined to migrate. The determinants of migration processes are mainly economic factors (employment opportunities, income maximization), life cycle (life-course), and improvements in overall living circumstances and quality of life [3].

Two types of migration can be distinguished: internal migration and international migration. The focus of this article is on internal migration. In demography, the term "internal migration" refers to the movement of people between regions and labor markets within a given country. Factors influencing the internal migration process include regional differences in mobility, healthcare services, shopping opportunities, housing, labor market and education opportunities, as well as the attractiveness of a location related to architectural and landscape features [4]. One type of internal migration is the movement of people within metropolitan areas, related to the process of suburbanization, which is the subject of study in this work.

The process of suburbanization emerged in Central and Eastern European cities later than in North America and Western Europe. The intensity of suburbanization processes significantly increased following the political transformations that occurred in these areas at the end of the 20th century. Since then, suburbanization has been recognized as the most important spatial process in the development of post-socialist cities [5,6]. It leads to the deconcentration of people, capital, and various activities within the city, and it also impacts the character of suburban rural areas, ultimately leading to challenges related to sustainable environmental and social development [7]. In Poland, at the turn of the 20th and 21st centuries, changes in internal migration patterns were observed, with an increase in the movement of people from cities to rural areas (a result of the intensification of suburbanization processes) [8]. This shift was influenced, among other factors, by the growing affluence of certain segments of society, as well as the development of modern technologies that facilitated remote work, such as e-work.

This study utilized artificial neural networks to forecast migration balance using the example of the city of Lublin. Lublin is located in southeastern Poland and is the capital of one of the 16 Polish voivodeships, namely the Lublin Voivodeship. It is the largest academic, economic, and administrative center in eastern Poland. In 2022, the population of Lublin was 331,243, while the number of students was 58,196. The main economic specializations of the city include: the automotive and machinery industry, healthy society, smart networks and ICT, automated food processing and functional food, chemical processes and specialized chemical products, modern business services, and innovative logistics [9].

The issue of suburbanization in the suburban zone of Lublin is present in only a few studies, indicating a research gap. In the work of Miczyńska-Kowalska [10], an assessment of the suburbanization process in the suburban area of Lublin was conducted based on changes in the population of neighboring municipalities, changes in the level of migration from cities to neighboring municipalities, the number of entrepreneurs, the number of business entities operating in the municipalities, employment structure, and the number of housing units. Flaga [11] analyzed the demographic consequences resulting from suburbanization in selected cities in Poland (including Lublin), focusing mainly on population changes in urban and suburban areas, as well as the characteristics of demographic structures and natural growth. The issue of migration concerning Lublin is addressed in the works of Maleszyk [12,13], particularly regarding youth migration. Ferenc [14] tackled the issue of suburbanization in Lublin in the context of land use changes in suburban areas. However, there is a lack of scientific studies concerning the modeling of migration balance for the city of Lublin. Therefore, the research area encompassed the city of Lublin and the municipalities functionally linked to it, forming the Lublin Functional Area (LOF). In addition to the city of Lublin, this area includes 15 municipalities: Nałęczów, Niedrzwica Duża, Niemce, Piaski, Spiczyn, Strzyżewice, Miejska Świdnik, Wólka, Głusk, Jabłonna, Jastków, Konopnica, Lubartów, the city of Lubartów, and Mełgiew. The inclusion

of municipalities located relatively close to the city is related to research findings indicating that municipalities situated in the immediate vicinity of the city are more strongly connected to it. In the case of municipalities located further away from the city center, the intensity of these connections decreases. The size of the central city also plays an important role in determining the extent of suburbanization processes. For example, in the case of two Slovak metropolises (Bratislava and Košice), the suburbanization zone extends up to 20–30 km, while in other regional cities, it is approximately 10–15 km [15].

The main objective of the work was to examine the usefulness of artificial neural networks (ANN) for predicting the internal migration balance for the city of Lublin in Poland. Another objective was to develop an experimental neural network model for forecasting the internal migration balance for the city of Lublin (for one year ahead) based on selected economic and social factors. The paper attempted to answer the following research questions:

1. What is the significance of suburbanization processes for urban areas and suburban rural areas?
2. Are artificial neural networks an appropriate method for forecasting migration balances in comparison to other methods used?
3. What is the forecast migration balance for the city of Lublin in 2023?

2. Literature Review

2.1. Suburbanization Processes

Various research approaches are employed to analyze the settlement systems of large cities. One of these is the hierarchical functional approach, which emphasizes the existence of unequal functional connections within urban areas, resulting in the presence of a center and its subordinate peripheries [16]. This approach relates to understanding and analyzing the urbanization process through the lens of the stages of a city's development, which progresses from a growth phase to a phase of stability or decline as the significance of the area's original locational advantages changes [17,18]. According to the model framework, cities undergo four stages in their development: centralization, decentralization, a decrease in population in the core in favor of the ring, and a reduction in the metropolitan area as a whole [19]. The urban life cycle theory posits that cities progress through the following phases: urbanization, suburbanization, desurbanization, and reurbanization [20,21].

Suburbanization constitutes one of the stages of the urbanization process and is associated with an increase in the population of the outer zones of agglomerations (suburbs and peripheral areas) and development in these areas. In contrast, in central cities, population growth is no longer as dynamic as in the initial stage of urbanization and may even be negative. Particularly, if the growth rate of suburban areas is faster than that of the city center, we can speak of relative suburbanization. However, if the population of the city center begins to decrease while the suburbs continue to grow, this stage is referred to as absolute suburbanization [21].

The process of suburbanization generally results in depopulation in the core of the agglomeration, while there may be an increase in the number of residents in the peripheral districts of the city within its administrative boundaries. This is an intra-urban suburbanization process, which helps mitigate the negative effects of depopulation occurring in the city's core [22]. GIS methods are among the tools used to analyze intra-urban suburbanization processes [23].

It is important to note that suburbanization processes are associated, on the one hand, with the movement of city residents to its suburbs. On the other hand, interregional migrants (newcomers to the metropolitan area) may also settle in suburban areas instead of the city center [24]. The example of Chinese cities demonstrates that it is possible to maintain a high population density in the city center despite an intensive suburbanization process. This is linked to the simultaneous occurrence of two opposing processes: the suburbanization of city residents and the urbanization of migrants from rural areas to cities [25].

It is important to note that the process of suburbanization involves not only the movement of people, but also relates to changes in the location of businesses seeking more favorable conditions for conducting economic activities outside the city center in terms of:

- Access to resources (e.g., available land for development, commercial, production, storage, and office space);
- Lower costs (e.g., labor costs, land prices, rental or purchase prices for commercial, production, storage, and office space, tax incentives, utility costs, and better transport accessibility implying lower transportation costs);
- Market opportunities (e.g., market niches in the form of a lack or shortage of services offered in new residential developments).

2.2. The Significance of Suburbanization to Urban Areas

Currently, metropolitan areas are one of the most important factors in regional development, attracting investments, innovations, and know-how, as well as influencing migration processes [26]. This phenomenon should be examined from two aspects. First, these areas serve as attractive places to live due to the concentration of various functions and the easy access to heterogeneous public and urban goods, including diverse market and non-market services (such as educational, research, health, cultural, and sports services), extensive transport infrastructure, and business-supporting institutions that facilitate entrepreneurship development (e.g., technology parks, business incubators, technology transfer centers), as well as a receptive and diverse labor market. This results in migrations of people to cities on a national scale, thereby fostering the development of metropolitan areas. These areas are characterized by a relatively high level of socio-economic development and quality of life, measured, for example, by the value of disposable income per capita. The literature suggests that the primary factor influencing the intensification of migration processes is the income level of the population [27], which serves as a significant pull factor for migrants to cities. This is particularly relevant in cases of large regional disparities in unemployment levels and regional development between different regions.

On the other hand, within the metropolitan area itself, migratory movements related to the process of suburbanization are also observed, which is the subject of this study. These movements arise, on the one hand, from the accumulation of unfavorable phenomena in urban spaces, such as environmental pollution and ecological threats, noise, congestion, lifestyle diseases, social pathologies, and crime. Research indicates that among the important factors influencing the suburbanization process in relation to Budapest, congestion and traffic jams, environmental pollution (e.g., air pollution caused by vehicle traffic), low housing quality, and the increasing number of cars, as well as improving service levels in suburban areas, were noted. The possibilities for constructing homes or purchasing properties in suburban areas, as well as local government policies regarding land regulations and spatial planning, also played a significant role [28].

The process of suburbanization is often associated with the search for a higher quality of life, offered in suburban areas, by relatively affluent individuals, while still being close to all the benefits typical of large cities. Spórna [29] notes that the political and economic transformation in Poland, which has led to increased wealth and education levels among residents, has been significantly linked to the desire to improve quality of life. However, despite the modernization of housing resources in cities and the revitalization of post-industrial areas, a progressive process of suburbanization and the associated outflow of the population from urban areas is observed.

Suburbanization brings not only positive but also negative effects on the spatial development process. Among the negative effects of suburbanization are uncontrolled urban sprawl, chaotic development patterns, increased vehicle usage, traffic congestion on the access roads to cities (resulting in longer travel times and associated costs), increased land use, and higher energy consumption in households. It also leads to a decrease in tax revenues for cities and the need to service populations from outside the city [22]. A particularly important issue arising from the ongoing process of suburbanization is the phenomenon of

urban sprawl. For example, in Switzerland, it can be observed that to mitigate the negative external effects of uncontrolled urban development on the environment and public health, authorities must intervene more strongly in the spatial and housing market [30].

Among the negative effects of suburbanization on the natural environment are, among others, an unsustainable ratio between bioactive areas and built-up areas, as well as the creation of new residential zones in areas with high natural/ecological values [31]. Zuziak [32] notes that the suburbanization process also leads to the underdevelopment of services in the suburban area, disparities between the development of housing and technical infrastructure, landscape degradation, and the unattractiveness of buildings. Furthermore, suburbanization may contribute to the reduction in open spaces and recreational areas, as well as the destruction of landscapes. Increased demand for land in suburban areas also leads to rising land prices.

2.3. Forcasting Migration in Urban Areas

Forecasting migration is one of the most challenging elements of demographic forecasting due to the high complexity of this phenomenon and its dependence on various factors that are inherently characterized by low predictability [33]. These factors include economic, social, political, and environmental aspects, as well as trends and events related to these factors. Additionally, a significant issue is the quality and availability of statistical data, which often inadequately capture actual migration in the reported data [34]. In Poland, this problem is also relevant due to the issue of unregistered permanent migration [35,36], which poses a certain research limitation in this study.

Forecasting at the local level is more challenging than at the national level due to the difficulty in estimating historical trends related to demographic indicators that can be extrapolated into the future, as well as the generally fewer available observations. Therefore, estimation and demographic forecasting models are utilized to make use of the limited datasets, including the Rogers–Castro models, log-linear models, and Bayesian statistical methods [37]. The literature suggests that in the context of forecasting for relatively small populations (e.g., at the regional level or lower), estimating migration rates appears to be more advantageous than estimating gross or net migration flows [4]. Hence, this study employs an approach based on estimating migration rates.

Research has shown that the choice of internal migration model significantly impacts population projection results, generating substantial differences in total population, geographical distribution, and age–sex composition [38]. Given the complexity and multidimensional nature of migration issues, diverse methods and research approaches should be employed to comprehensively capture the specificity of this dynamically changing phenomenon [39,40]. A review of selected models of internal migration can be found in the works of Hajj [41], Wilson and Bell [38], and Fantazzini et al. [42].

Modeling internal migration can take two forms: a micro or macro approach [43]. The micro approach focuses on the behavior of individuals in the decision-making process regarding migration, taking into account the factors influencing migration. This is a behavioral approach based on the personal characteristics of the individual [44]. In contrast, the macro approach concentrates on the differences between specific places or locations (e.g., countries, regions, municipalities) that may serve as significant factors or barriers to migration. For macro-scale migration modeling, gravity models or systemic approaches are often employed [45].

The analysis and forecasting of internal migration utilize a variety of traditional and innovative econometric techniques, including panel data models, fixed/random effects models, probit/logit models, Bayesian statistics methods, formal demography, and microeconomics. Econometric modeling, particularly multiple regression models, is often used in research explaining internal migration within a socio-economic context, examining the relationship between a specific economic variable (e.g., unemployment rate) and the volume of migration [46].

Mapping methods are also employed for analyzing internal migration, with key indicators including population density, unemployment, distance to the nearest urban center, and size of administrative units [47,48]. Survey research methods are used to investigate respondents' migration intentions [49], as well as classification tree methods, which have limited predictive capabilities [39], gravity models [50,51], including Structural Equation Modeling (SEM) [26], and probit and logit models [52–55], as well as machine learning methods [56]. For short-term migration modeling, panel data (or longitudinal data) models are employed [57,58].

In recent years, there have been attempts to use Google Trends for modeling migration in two Russian cities, Moscow and Saint Petersburg; however, the empirical analysis did not provide evidence of a relationship between the number of internet searches and the probability of migration [42]. Nevertheless, incorporating Google Trends data into the model results in smaller forecasting errors compared to models without such data.

2.4. Artificial Neural Networks

In Poland, population forecasts are presented by the Central Statistical Office (Główny Urząd Statystyczny) based on the cohort component method, which utilizes data on fertility, mortality, and migration, disaggregated by gender and single-year age groups. The Rogers–Castro model is employed to establish assumptions regarding migration [59].

This study proposes a relatively new methodological approach to forecasting internal migration balances using artificial neural networks. These networks represent one of the important techniques that can be utilized in modeling urban development [60]. However, so far, there are only a few cases in the global literature that indicate the use of artificial neural networks for forecasting international migration, such as studies on the residents of the Khmelnytskyi region [61] or the population of Fiji [62]. In Polish literature, Beim's [63] work presents an attempt to apply artificial neural networks to model the suburbanization process in the Poznań agglomeration. This tool was used to assess the attractiveness of settling in various areas of the agglomeration. In 2022, authors from Lublin developed an experimental model of an artificial neural network to predict migration balances for the city of Poznań for one year and two years into the future [64]. The conducted research concluded that it is possible to use artificial neural networks for accurately forecasting migration balances for Poznań. This study attempts to deepen the discussion on the application of artificial neural networks for forecasting migration balances.

3. Materials and Methods

The first stage of the analysis was to determine the input parameters for ANN modeling. For this purpose, PCA was used in the Statistica 13.1 environment (TIBCO Software Inc., Palo Alto, CA, USA). The technique of principal component analysis (PCA), a prevalent method for reducing dimensionality in multivariate data analysis and statistical studies, seeks to convert the initial high-dimensional data into a fresh series of orthogonal variables, termed as principal components. The objective is to preserve the maximum possible variance in the data. These principal components are arranged in a manner that the first component seizes the most substantial variance in the data, the second component seizes the second most substantial variance, and the sequence continues.

The fundamental objective of PCA is to streamline the representation of intricate datasets and to pinpoint the dominant patterns or trends inherent in the data. By diminishing the number of dimensions, PCA aids in alleviating the "dimensionality curse" issue, a potential problem when handling high-dimensional data, by concentrating on the most informative components.

In the process of executing a principal component analysis (PCA), the eigenvalues and eigenvectors of the covariance matrix are computed. The eigenvectors embody the principal components, while their corresponding eigenvalues quantitatively determine the magnitude of variance elucidated by each component. Subsequently, a selection of principal components is made. These principal components are those possessing the highest

eigenvalues. In the scrutinized studies, it was postulated that a number of components equivalent to what is necessary to elucidate approximately a minimum of 95% variance in the data would be chosen.

Data for the PCA covered the years 2005–2022 and were obtained from the Local Data Bank (BDL) of the Central Statistical Office (GUS). Chosen as input parameters for the PCA were indicators characterizing the housing stock (number and area of the existing dwellings, number of rooms in the dwelling) and the demographic situation (balance of internal migration, changes in the number of population per 1000 inhabitants, total population, broken down by women and men, and population density per 1 km^2) in 15 communes (Lublin, Lubartów, Głusk, Jabłonna, Jastków, Konopnica, Niedrzwica Wielka, Niemce, Strzyżewice, Wólka, Spiczyn, Nałęczów, Świdnik, Mełgiew and Piaski). The selection of parameters was guided by information from the Migration Survey [65], which shows that housing conditions were mentioned right after family matters as the second most important reason for migration. This is also confirmed by studies [66–68] which show that housing stock significantly influences the choice of place of settlement. On the other hand, research by Józefowicz [69] shows that population density is related to migration levels in a statistically significant way.

Based on the PCA, input parameters and their number will be selected for modeling the neural network. The migration balance for the city of Lublin for the subsequent year was modeled using a one output neuron. The Neural Net Fitting neural network library in the MatLab software (The MathWorks, Inc., Natick, MA, USA, Version R2023a) package was used to build artificial neural network (ANN) models. A shallow neural network with a single hidden layer was employed for the modeling process. The number of neurons was experimentally determined within a range of 2 to 15 and the Levenberg–Marquardt algorithm was used for training. The dataset was divided in a 75%:25% ratio (training data: validation data). This split was carried out randomly across the entire dataset to ensure the representativeness of the validation set and to avoid any bias linked to specific years. It is important to note that the data used for validation were not assigned to specific years but were randomly sampled from the full dataset. This approach was necessary given the limited size of the dataset, as it allowed the model to generalize across the available data. Due to the paucity of data, a test set was omitted [70]. To further prevent overfitting, the early stopping method was employed during the training process. Early stopping monitors the validation error during each training epoch and halts the training when the validation error stops decreasing for a predefined number of epochs. In this study, early stopping was critical in ensuring that the neural network did not overfit the training data, as shown in the training state plot. The training process concluded after five epochs, as the validation error plateaued, indicating that further training would not improve generalization. The quality of the network was determined based on the value of regression R (correlation coefficient), which measures the correlation between output and input data:

$$R(y',\ y^*) = \frac{cov(y, y')}{\sigma_{y'}\sigma_y} \quad (1)$$

where $\sigma_y{}'$—standard deviation of reference values, σ_y—standard deviation of predicted values.

The correlation coefficient (R) measures the strength and direction of the linear relationship between predicted and actual values. The value ranges from 0 to 1, with R = 0 indicating no correlation and R = 1 signifying a perfect linear correlation between output and input data. A higher R value indicates improved network performance, showcasing the model's capacity to accurately predict target values from the input data. This metric provides a straightforward and interpretable measure of the model's quality, especially in evaluating its suitability for forecasting tasks.

Following Kulisz and Kujawska [71], it was assumed that a regression R value > 0.85 indicates a good network fit and a high level of agreement of measurement points. In

addition to the regression coefficient R value, the Mean Squared Error (MSE) value was considered when selecting the best network, calculated as:

$$\mathrm{MSE} = \frac{1}{n}\sum_{i=1}^{n}\left(y_i - y_i'\right)^2 \quad (2)$$

where y_i is the actual value of the migration balance level for the city of Lublin and y_i' denotes the value of the migration balance level for the city of Lublin for the i-th observation obtained from the model.

4. Results

After comparing all the factors for all communes, 144 variables were obtained. The initial step necessitated the preparation of input data for modeling by subjecting the amassed data to standardization. This standardization process involved subtracting the mean of the variable from the analyzed data type, followed by dividing it by the standard deviation. As a result of such an operation, new data were obtained with an expected mean value of 0 and a standard deviation of 1. Subsequently, in the STATISTICA software (version 13.1), statistical processing was carried out on the appropriately prepared data.

To select parameters for modeling the neural network, PCA was used. The eigenvalues of the principal components were calculated; their number at the beginning was limited by the software itself to 17. As can be seen from Table 1, the cumulative % of variance for the eight principal components will achieve 95% of the variance of the data. This means that these eight principal components contain 95% of the information (measured by variance) that the original 144 features contained. This is also shown in Figure 1.

Table 1. The cumulative % of variance for the principal components.

Id	Eigenvalues	% of Total Variance	Cumulative Eigenvalues	Cumulative %
1	111.6297	77.52062	111.6297	77.5206
2	10.4295	7.24270	122.0592	84.7633
3	3.4310	2.38262	125.4902	87.1459
4	3.3374	2.31765	128.8276	89.4636
5	2.7387	1.90186	131.5662	91.3654
6	2.1623	1.50159	133.7285	92.8670
7	2.0716	1.43863	135.8002	94.3057
8	1.6446	1.14208	137.4448	95.4477
9	1.3237	0.91924	138.7685	96.3670
10	1.2181	0.84589	139.9865	97.2129
11	0.9724	0.67529	140.9589	97.8882
12	0.8303	0.57662	141.7893	98.4648
13	0.7178	0.49850	142.5071	98.9633
14	0.5553	0.38564	143.0625	99.3489
15	0.4187	0.29074	143.4811	99.6397
16	0.3024	0.21001	143.7835	99.8497
17	0.2165	0.15033	144.0000	100.0000

Source: own study.

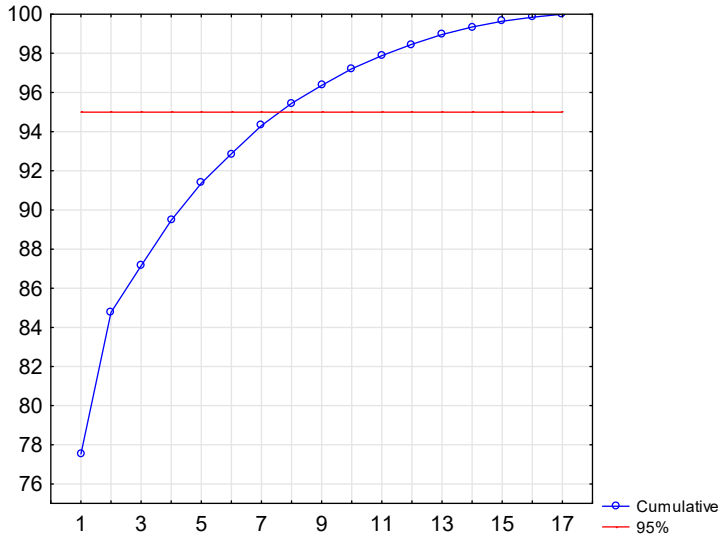


Figure 1. Cumulative % of variance for the principal components. Source: own study.

To identify the features that have the most influence on each principal component, we can examine the PCA components' attributes. The component attributes of the PCA object gives us the coefficients of each feature in the principal components. The higher the absolute value of a coefficient, the more influence that feature has on the principal component. Table 2 shows the variable that has the greatest impact on the principal component.

Table 2. Variable with the greatest impact on principal component.

Principal Components	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**
Factor name	a74	a59	a67	a85	a139	a40	a76	a58
Factor coordinates of the variables (absolute value)	0.9992	0.7496	0.6305	0.8121	0.5898	0.6181	0.5486	0.4583

Source: own study.

Therefore, for modeling the migration balance for the city of Lublin, such input parameters were selected as follows:

a40—internal migration balance in Jablonna;
a58—internal migration balance in Konopnica;
a59—birthrate per 1000 population in Konopnica;
a67—internal migration balance in Niedrzwica Duża;
a74—number of rooms in the dwelling in Niemce;
a76—internal migration balance in Niemce;
a85—internal migration balance in Strzyżewice;
a139—internal migration balance in Piaski;

and the variable balance of internal migration in Lublin was also added.

The optimal modeling results were obtained using a network consisting of nine neurons. This result was obtained after 5 epochs (Figure 2). The outcomes derived for the optimal neural network, along with the performance metrics associated with this network are shown in Table 3, where training stops after six consecutive increases in validation error (or no error reduction), and the optimal results are derived from the iteration with the minimum validation error.

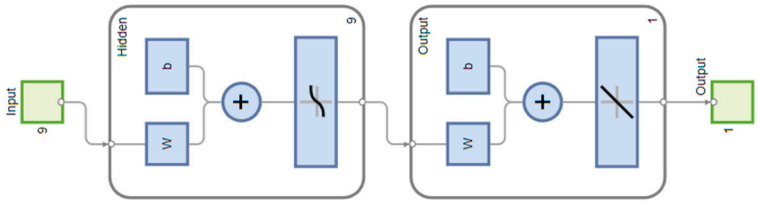


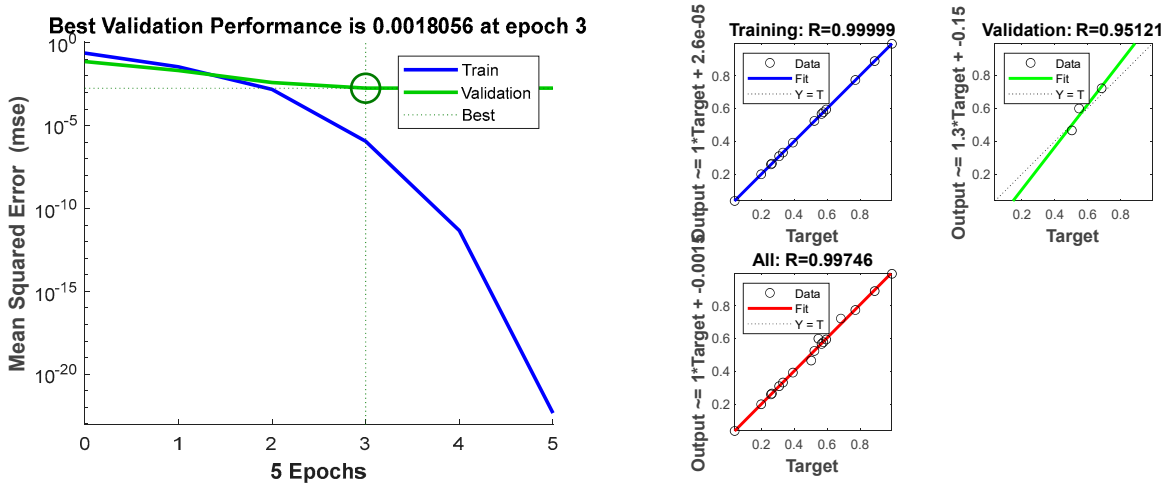
Figure 2. A diagrammatic depiction of a synthetic neural network. Source: own study.

Table 3. The outcomes derived for the optimal neural network, along with the performance metrics associated with this network.

Training Algorithm	Levenberg–Marquardt
Epoch	11
Performance (MSE)	4.56×10^{-23}
Best validation performance	0.0018056 at epoch 3
Gradient	4.216×10^{-12}
R(all)	0.99746
MSE	3.1957×10^{-4}

Source: own study.

The Mean Squared Error (MSE) metric was used to evaluate the quality effectiveness of artificial neural networks (ANNs). The results showed that optimal validation was achieved during the third epoch. This is shown in Figure 3a. Regression plots for training, validation, and all datasets are shown in Figure 3b.



(**a**) (**b**)

Figure 3. (**a**) Network validation performance. (**b**) Network regression statistics for individual sets and the total set. Source: own study.

For each subset, the results are as follows: the training set yields an R value of 0.99999, while the validation set results in an R value of 0.95121. The overall correlation amounted to R = 0.99746, indicating the degree of overlap between the measurement points and the fitting line with the ideal prediction line Y = T. These outcomes surpass the value of R > 0.85, testifying to the network's excellent fit and the high level of agreement among the measurement points.

The juxtaposition of empirical data and the data obtained as a result of the prediction, as well as the forecast of migration balance for the city of Lublin for the one year ahead is presented in Figure 4. The predicted data, in many cases, closely align with the actual

values, demonstrating the high accuracy of the ANN model. This alignment is further evidenced by the fact that the predicted points on the graph are located directly beneath the actual data points, indicating minimal deviation between the two. Such a high degree of agreement reflects the robustness of the model in capturing the underlying trends and relationships within the migration data, enabling reliable forecasts.

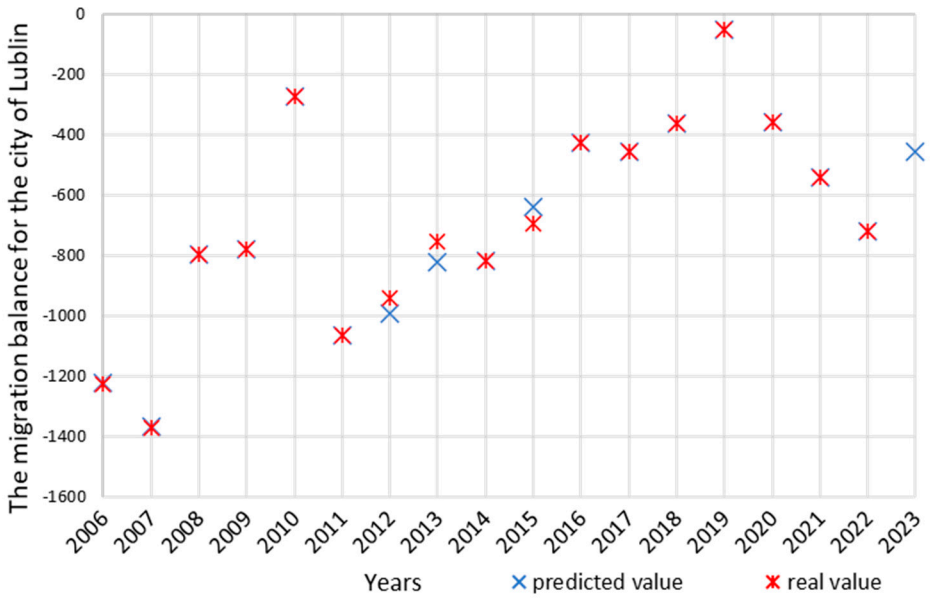


Figure 4. The migration balance for the city of Lublin. Source: own study.

5. Discussion

5.1. Comparison of Econometric Models and Artificial Neural Networks in Urban Migration Forecasting

The analysis of urban migration forecasting has evolved significantly over the years, with various methodologies being employed to predict migration patterns. Among these, artificial neural networks (ANNs) have emerged as a promising alternative to traditional econometric models. This comparative analysis aims to highlight the advantages of ANNs in forecasting urban migration by contrasting them with conventional econometric approaches, which often rely on linear assumptions and aggregate data.

Traditional econometric models, such as regression analysis, have been widely used to study migration patterns. These models typically assume linear relationships between variables, which can oversimplify the complexities of migration dynamics. Laajimi and Gallo [72] emphasize that while traditional macroeconomic variables remain significant in explaining migration flows, they often fail to capture the asymmetric effects and interactions between different factors. This limitation is particularly evident when considering the multifaceted nature of migration, where push and pull factors can vary significantly across regions and populations. In contrast, ANNs are capable of modeling non-linear relationships and interactions among variables, allowing for a more nuanced understanding of migration dynamics.

Moreover, traditional econometric models often rely on aggregate data, which can obscure important variations at the individual or household level. Zhu and Luo [73] point out that households with better-educated labor are more likely to migrate, a factor that may not be adequately captured in aggregate models. ANNs, on the other hand, can leverage disaggregated data, such as individual-level characteristics, to improve the accuracy of migration forecasts. Ling and Longley [74] highlight the potential of new georeferenced data sources to enhance migration research, enabling more precise modeling of migration patterns at a granular level.

Another significant advantage of ANNs is their ability to incorporate a wide range of variables without the need for strict assumptions about their relationships. Traditional econometric models often require researchers to specify the functional form of the relationships between variables, which can introduce bias if the true relationships are misrepresented. In contrast, ANNs can learn complex patterns from data without requiring explicit specification of the underlying relationships. This flexibility allows ANNs to adapt to the unique characteristics of different migration contexts, as noted by Bigotte et al. [75], who argue for the importance of considering various factors beyond traditional economic indicators in migration studies.

Furthermore, the predictive performance of ANNs has been shown to surpass that of traditional econometric models in various contexts. Li and Kockelman [76] conducted a comparative analysis of machine learning techniques, including ANNs, against conventional econometric methods, revealing that machine learning approaches often yield higher prediction accuracies. This finding is particularly relevant for migration forecasting, where accurate predictions are crucial for urban planning and policy formulation. The ability of ANNs to handle large datasets and identify patterns that may not be apparent through traditional methods further enhances their utility in this field.

In addition to their predictive capabilities, ANNs also offer advantages in terms of handling missing data and noise. Traditional econometric models can be sensitive to missing or outlier data points, which can skew results and lead to unreliable conclusions. ANNs, however, are generally more robust to such issues, as they can learn from the available data and make predictions, even in the presence of incomplete information. This resilience is particularly beneficial in migration studies, where data quality and availability can vary significantly across regions and time periods.

Moreover, the integration of ANNs with other advanced data analysis techniques, such as Geographic Information Systems (GIS), can further enhance migration forecasting. The combination of spatial analysis with the predictive power of ANNs allows researchers to visualize migration patterns and identify hotspots of migration activity, as demonstrated by Wu and Wu [77]. This spatial dimension is often overlooked in traditional econometric models, which may treat migration as a purely economic phenomenon without considering the geographical context.

The application of ANNs in migration forecasting also aligns with the growing trend towards data-driven decision-making in urban planning. As cities face increasing pressures from population growth and migration, the need for accurate forecasting tools becomes paramount. ANNs provide urban planners and policymakers with actionable insights that can inform strategies for managing migration and its impacts on urban infrastructure and services. By leveraging the strengths of ANNs, cities can better anticipate and respond to the challenges posed by migration, ultimately leading to more sustainable urban development.

5.2. Broader Implications of Applying Artificial Neural Networks in Migration Forecasting

This study employed principal component analysis (PCA) to reduce dataset dimensionality and identify the most significant variables for modeling purposes. PCA was implemented on the complete dataset prior to its division into training and validation subsets to reduce variability in the chosen principal components, thereby maintaining consistency in the input features throughout the analysis. These selected components can therefore be considered dependable indicators for predicting the migration balance of the city of Lublin. This suggests that artificial neural networks may be an effective tool in supporting decision-making processes for forecasting the migration balance of this city.

The forecast for internal migration balance for the city of Lublin in 2023 indicates a reduction in the negative migration balance; however, the declining trend is expected to continue throughout 2023. At the same time, neighboring municipalities are predicted to have a positive migration balance, though this too will show a downward trend.

The application of artificial neural networks (ANN) in migration forecasting, as demonstrated in the case of Lublin, holds the potential for implementing this method in other

Polish cities as well as on the international stage. Cities with varying levels of development, including medium-sized regional centers and large metropolises, can benefit from the opportunities offered by ANN. Thanks to the flexibility of the ANN model, it is possible to account for complex, non-linear relationships between factors influencing internal migration, which are difficult to capture using traditional methods. Research conducted by Ovchynnikova et al. [61] on the example of residents of the Khmelnytskyi region in Ukraine and Qiokata and Khan [62] in Fiji indicates that ANN can be effectively applied in migration forecasting across diverse geographical and socio-economic contexts.

At the same time, it is important to note that migration modeling using artificial neural networks (ANN) may encounter certain limitations, particularly in situations involving sudden socio-political changes, such as shifts in migration policy, economic crises, or armed conflicts. Due to their unpredictability, these factors can lead to significant deviations from predicted migration trends [42]. Therefore, it may be beneficial to enhance ANN models by incorporating additional variables related to unstable political conditions or crisis events that could have a substantial impact on migration.

One of the limitations of the current study is the potential impact of unregistered permanent migrations on the accuracy of the model, particularly in situations involving sudden socio-political or economic disruptions. To improve the model's adaptability and robustness, future research could incorporate additional variables that reflect political and environmental dynamics. These could include indices of political stability, policy changes, or environmental crises, which are likely to influence migration processes. Integrating such factors would make the model more responsive to abrupt changes and better suited to forecasting in volatile conditions.

In the literature on internal migration forecasting, various models are used, including traditional econometric methods such as regression models or cohort-component models [4]. These models are widely applied; however, their limitations stem from assumptions about the linearity of relationships between variables. Artificial neural networks offer an advantage in modeling non-linear relationships, which is particularly important for migration forecasting, where the connections between demographic, economic, and spatial factors are often non-linear in nature [40].

The study by Pietrzak, Wilk, and Matusik [50], which employed a gravity model to analyze internal migration in Poland, highlights the need to incorporate various methods in migration research, including more advanced technologies such as artificial neural networks (ANN). Compared to gravity models, which rely on basic assumptions about distance and population size, ANN allows for the inclusion of a larger number of variables and better adaptation to complex data structures.

This work addressed the issue of extrapolation by employing a conventional data division, allocating 75% for training and 25% for validation, with data randomly picked to assure representativeness across the dataset. This method was considered appropriate due to the restricted data provided, and early halting was employed to avert overfitting. Early stopping observed validation error during training and terminated the process when additional iterations ceased to enhance the model's performance, hence ensuring strong generalization. While this methodology proved effective, it is acknowledged that more advanced validation procedures, such as k-fold cross-validation with a temporal framework, could provide additional insights into the model's performance and generalizability. Future research could explore these methodologies to better address the challenges of extrapolation and ensure the model's robustness in scenarios with temporal dependencies and dynamic migration patterns.

Moreover, as demonstrated in the research by Gawryluk, Kulisz, and Komor [64], artificial neural networks (ANN) can generate more accurate short-term forecasts than traditional regression models. Specifically, in studies concerning Poznań, ANN proved to be more effective in predicting migration balance compared to models based on historical trend analysis. Including a comparative analysis in this article could highlight how ANN

performs relative to other forecasting methods, especially in terms of short-term predictions and the complex migration conditions in medium-sized cities like Lublin.

6. Conclusions

The main objective of the work, which was to examine the usefulness of artificial neural networks for predicting the internal migration balance for the city of Lublin, was achieved. Its implementation allowed the development of an experimental neural network model for predicting the internal migration balance (for 1 year ahead), the correlation coefficient of which was R = 0.99746, indicating excellent network fit and a high level of data agreement obtained as a result of prediction.

Another objective was to develop an experimental neural network model for forecasting the internal migration balance for the city of Lublin (for one year ahead) based on selected economic and social factors. The development of the artificial neural network model in the research process was two-stage. In the first stage, a principal component analysis (PCA) was conducted to reduce the initial dataset of 144 variables by identifying the most informative principal components, which together captured approximately 95% of the variance in the data. The PCA identified eight principal components based on eigenvalues, which were then used to select the key input variables for modeling. In the second stage, the mean square error (MSE) metric was used to assess the qualitative effectiveness of artificial neural networks (ANN). The results showed that optimal validation was achieved in the third epoch.

The research allowed us to answer the research questions posed in the paper. Suburbanization processes are determined by a number of economic, social, environmental, and political factors. The study shows that suburbanization, as a process of an economic and spatial nature, plays an important role in local and regional development. Suburbanization causes changes in the distribution and spatial concentration of people, capital, and various activities in urban space. Suburbanization processes also have negative effects, e.g., in the form of urban sprawl, congestion of the city's infrastructure (e.g., transport), a decrease in the city's tax revenue, higher costs associated with serving the population outside the city. Suburbanization also affects changes in the character of suburban rural areas, e.g., the emergence of new housing zones in areas with valuable natural assets or inadequate infrastructure, landscape degradation, underdevelopment of services (including public services) in relation to the needs of residents. Hence, it is important to be able to forecast changes in the size and structure of the population in the future, which provides the opportunity for effective management and local and regional development policies in a given area.

The study shows that artificial neural networks have many advantages over other methods used to forecast migration balances. While traditional econometric models have played a significant role in the study of urban migration, their limitations in capturing the complexities of migration dynamics highlight the need for alternative approaches. ANNs offer several advantages, including the ability to model non-linear relationships, incorporate disaggregated data, and improve predictive accuracy. As urban migration continues to evolve, the integration of ANNs into migration forecasting represents a promising avenue for enhancing our understanding of this critical phenomenon.

The results of the study showed that the forecast of the internal migration balance for the city of Lublin in 2023 indicates a decrease in the negative migration balance, but the downward trend is expected to continue throughout 2023. At the same time, the neighboring municipalities are projected to have a positive migration balance, although this too will show a downward trend.

The application of artificial neural networks (ANN) in migration forecasting, as shown in the example of Lublin, creates the potential to implement this method in other Polish cities, as well as internationally. Artificial neural networks provide urban planners and decision-makers with practical insights that can be used to develop strategies to manage migration and its impact on urban infrastructure and services. By exploiting the strengths of neural networks, cities can better anticipate and respond to the challenges of migration,

ultimately leading to more sustainable urban development, as well as areas located in close proximity to cities.

Among the directions for further research are continued studies on the use of artificial neural networks as a tool for forecasting migration balances, including, i.e., in relation to other Polish cities, cities located in other EU countries or outside the EU. However, further research is needed to specify the model.

A certain research limitation in the present study is the frailty of the available statistical data, mainly related to the problem of unregistered permanent migration. Limitations associated with modeling migration using artificial neural networks also arise from the occurrence of sudden socio-political changes (e.g., in migration policy, economic crises or armed conflicts), which can be mitigated by including additional variables characterizing these problems in ANN models.

While our study demonstrates the potential of using PCA and ANN for migration balance forecasting, it is important to acknowledge certain limitations and challenges associated with these tools. One significant drawback is the computational complexity of ANN models, which require substantial computational resources and expertise, particularly during the model tuning phase. The process of selecting hyperparameters, such as the number of neurons or layers, can be time-consuming and may involve trial-and-error approaches. Additionally, the interpretability of ANN results remains a challenge. Unlike traditional statistical models, which provide straightforward relationships between variables, ANN operates as a 'black box', making it difficult to explain how input variables influence the output. This can pose challenges for decision-makers, who require transparent and interpretable models for practical application.

The methodology presented in this study provides a replicable framework for forecasting migration balances in other regions. While each region differs in its socio-economic, demographic, and spatial characteristics, the step-by-step approach—starting with data preparation, dimensionality reduction using PCA, and building an ANN model—allows for adaptation to local contexts. This study emphasizes the importance of tailoring the input variables and model parameters to regional specificities, as a universal model is unlikely to capture the diversity of migration processes. Future applications of this framework in other regions could refine and validate its generalizability.

Migration balance forecasts generated by artificial neural networks (ANN) can be used for strategic planning in economic, social, and environmental sectors, helping local authorities and businesses adapt to demographic changes and future challenges. In the economic sector, modeling results can be used in local and regional planning to make decisions, such as implementing economic incentives in areas struggling with a negative migration balance, to attract businesses and retain residents in these regions. In the social sector, migration balance forecasts can identify areas where a growing population (due to migration) may increase demand for public services, potentially straining existing infrastructure (schools, nurseries, kindergartens, hospitals, and other public services). This allows local authorities to prepare for future changes by building schools, hospitals, and other public facilities, ensuring appropriate infrastructure (buildings) and the necessary personnel for the proper functioning of specific institutions. In terms of infrastructure development, predicted population changes can be used in decision-making regarding the expansion of public transport systems, roads, and other infrastructure. In the environmental sector, migration balance forecasts can support effective natural resource management by identifying areas with a growing population that require the development of water, energy, and waste management systems.

Author Contributions: Conceptualization, A.G., A.K. and M.K.; data curation, A.G. and M.K.; formal analysis, A.G., A.K., M.K., P.Z. and D.K.; funding acquisition, A.G., A.K., M.K., P.Z. and D.K.; investigation, A.G., A.K., M.K., P.Z. and D.K.; methodology, A.G., A.K. and M.K.; project administration, A.G., A.K. and M.K.; resources, A.G., A.K. and M.K.; software, M.K.; supervision, A.G., A.K. and M.K.; validation, A.G. and M.K.; visualization, M.K.; writing—original draft, A.G.,

A.K., M.K., P.Z. and D.K.; writing—review and editing, A.G., A.K., M.K., P.Z. and D.K. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable. This study did not require ethical approval.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data presented in this study are openly available in Statistics Poland: https://bdl.stat.gov.pl/bdl/start (accessed on 1 June 2024).

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Bogue, D.J. Internal migration. In *The Study of Population*; Hauser, P.M., Duncan, O.D., Eds.; University of Chicago Press: Chicago, IL, USA, 1959.
2. Ravenstein, E.G. The laws of migration. *J. R. Stat. Soc.* **1885**, *48*, 167–235. [CrossRef]
3. Boyle, P.; Halfacree, K.; Robinson, V. *Exploring Contemporary Migration*; Routledge: London, UK; New York, NY, USA, 2014.
4. Vanella, P.; Hellwagner, T.; Deschermeier, P. Parsimonious stochastic forecasting of international and internal migration on the NUTS-3 level—An outlook of regional depopulation trends in Germany. *Vienna Yearb. Popul. Res.* **2023**, *21*, 1–55. [CrossRef]
5. Enyedi, G. Urbanisation in East Central Europe: Social processes and societal responses in the state socialist systems. *Urban Stud.* **1992**, *29*, 869–880. [CrossRef]
6. Kubeš, J. European post-socialist cities and their near hinterland in intra-urban geography literature. *Bull. Geogr. Socio-Econ. Ser.* **2013**, *19*, 19–43. [CrossRef]
7. Vasárus, G.L.; Lennert, J. Suburbanization within city limits in Hungary—A challenge for environmental and social sustainability. *Sustainability* **2022**, *14*, 8855. [CrossRef]
8. Kałuża-Kopias, D. Trafność prognozowanego poziomu migracji w pracach GUS. *Space-Soc. -Econ.* **2014**, *13*, 75–90. [CrossRef]
9. Kijek, T.; Bukalska, E.; Czerniak, J.; Kijek, A.; Komor, A.; Matras-Bolibok, A.; Miszczuk, A.; Pylak, K.; Szczypa, G. *Specjalizacje Gospodarcze Lublina–w Nowym Europejskim Układzie Współzależności Ekonomicznych (w Perspektywie do 2030 r.)*; Urząd Miasta Lublin: Lublin, Poland, 2021.
10. Miczyńska-Kowalska, M. Suburbanizacja strefy podmiejskiej Lublina (Suburbanization of the suburban zone of Lublin). *Ann. Agric. Econ. Rural. Dev.* **2019**, *106*, 72–86. [CrossRef]
11. Flaga, M. Demographic consequences of suburbanization in selected towns in the eastern borderlands of Poland. *Bull. Geogr. Socio-Econ. Ser.* **2010**, *14*, 5–22. [CrossRef]
12. Maleszyk, P. Outflow of talents or exodus? Evidence on youth emigration from EU's peripheral areas. *Region* **2021**, *8*, 33–51. [CrossRef]
13. Maleszyk, P. Migracje młodzieży i ich znaczenie dla potencjału demograficznego miast. Przykład Lublina. In *Konsekwencje Zmian Demograficznych*; Markowski, K., Ed.; Główny Urząd Statystyczny: Warsaw, Poland, 2022; pp. 68–86.
14. Ferenc, M. Land use changes in suburban areas—Case study of Lublin. *Acta Agrar. Debreceniensis* **2012**, *49*, 43–46. [CrossRef]
15. Trembošová, M.; Dubcová, A.; Nagyová, L.; Cagáňová, D.; Kóňová, T. Suburbanization process of regional cities in Slovakia. *Transp. Logist.* **2022**, *22*, 21–31.
16. Zborowski, A. *Przemiany Struktury Społeczno-Przestrzennej Regionu Miejskiego w Okresie Realnego Socjalizmu i Transformacji Ustrojowej (na Przykładzie Krakowa)*; Instytut Geografii i Gospodarki Przestrzennej Uniwersytetu Jagiellońskiego: Kraków, Poland, 2005.
17. Birch, D.L. Towards a stage theory of urban growth. *J. Am. Inst. Plan.* **1971**, *37*, 78–87. [CrossRef]
18. Rust, E. *No Growth: Impacts on Metropolitan Areas*; Lexington Books: Lexington, MA, USA, 1975.
19. Hall, P. Spatial structure of metropolitan England and Wales. In *Spatial Policy Problems of the British Economy*; Chisholm, M., Manners, G., Eds.; Cambridge University Press: Cambridge, UK, 1971; pp. 96–125.
20. Klaassen, L.H.; Paelinck, J.H.P. The future of large towns. *Environ. Plan. A Econ. Space* **1979**, *11*, 1095–1104. [CrossRef]
21. van den Berg, L.; Drewett, R.; Klaassen, L.H.; Rossi, A.; Vijverberg, C.H.T. *Urban Europe: A Study of Growth and Decline*; Pergamon Press: Oxford, UK, 1982.
22. Szmytkie, R. Suburbanisation processes within and outside the city: The development of intra-urban suburbs in Wrocław, Poland. *Morav. Geogr. Rep.* **2021**, *29*, 149–165. [CrossRef]
23. Hegedűs, L.D.; Túri, Z.; Apáti, N.; Pénzes, J. Analysis of the intra-urban suburbanization with GIS methods. *Folia Geogr.* **2023**, *65*, 23–39.
24. White, M.J.; Lindstrom, D.P. Internal migration. In *Handbook of Population*; Poston, D.L., Micklin, M., Eds.; Handbooks of Sociology and Social Research; Springer: Boston, MA, USA, 2005; pp. 311–346. [CrossRef]
25. Qin, X.; Wei, Y.D.; Yu, Z.; Xiong, N. Urbanization, suburbanization, and population redistribution in urban China: A case study of Nanjing. *J. Urban Plan. Dev.* **2022**, *148*, 05022034. [CrossRef]
26. Pietrzak, M.B.; Żurek, M.; Matusik, S.; Wilk, J. Application of structural equation modeling for analyzing internal migration phenomena in Poland. *Przegląd Stat.* **2012**, *4*, 487–503. [CrossRef]

27. Fachin, S. Long-run trends in internal migrations in Italy: A study in panel cointegration with dependent units. *J. Appl. Econom.* **2007**, *22*, 401–428. [CrossRef]
28. Kok, H.; Kovács, Z. The process of suburbanization in the agglomeration of Budapest. *Neth. J. Hous. Built Environ.* **1999**, *14*, 119–141. [CrossRef]
29. Spórna, T. The suburbanization process in a depopulation context in the Katowice conurbation, Poland. *Environ. Socio-Econ. Stud.* **2018**, *6*, 57–72. [CrossRef]
30. Lerch, M. The end of urban sprawl? Internal migration across the rural–urban continuum in Switzerland, 1966–2018. *Popul. Space Place* **2023**, *29*, e2621. [CrossRef] [PubMed]
31. Korwel-Lejkowska, B. Changes in the spatial structure of land use as a result of suburbanization processes in rural areas surrounding the Tri-city agglomeration. *Acta Sci. Pol. Adm. Locorum* **2022**, *21*, 395–413. [CrossRef]
32. Zuziak, Z. Strefa podmiejska w architekturze miasta. W stronę nowej architektoniki regionu miejskiego. In *Problem Suburbanizacji*; Lorens, P., Ed.; Biblioteka Urbanisty: Warszawa, Poland, 2005; Volune 7, pp. 17–32.
33. Zborowski, A. Prognoza demograficzna dla Krakowa–metoda badań. In *Człowiek i Przestrzeń*; Kortus, B., Ed.; Instytut Geografii i Gospodarki Przestrzennej Uniwersytetu Jagiellońskiego: Kraków, Poland, 2001; pp. 89–97.
34. Rogers, A.; Little, J.; Raymer, J. *The Indirect Estimation of Migration: Methods for Dealing with Irregular, Inadequate, and Missing Data*; Springer Science & Business Media: Berlin, Germany, 2010; Volume 26.
35. Śleszyński, P. Znaczenie migracji wewnętrznych w rozwoju społeczno-gospodarczym i terytorialnym Polski. In *Migracje a Sytuacja Demograficzna Polski*; Hrynkiewicz, J., Witkowski, J., Potrykowska, A., Eds.; Rządowa Rada Ludnościowa: Warszawa, Poland, 2019; pp. 60–73.
36. Jonczy, R. Migracje a sytuacja demograficzna Polski. In *Migracje a Sytuacja Demograficzna Polski*; Hrynkiewicz, J., Witkowski, J., Potrykowska, A., Eds.; Rządowa Rada Ludnościowa: Warszawa, Poland, 2019; pp. 37–59.
37. Zhang, J.L.; Bryant, J. Bayesian disaggregated forecasts: Internal migration in Iceland. In *Developments in Demographic Forecasting*; The Springer Series on Demographic Methods and Population Analysis; Mazzuco, S., Keilman, N., Eds.; Springer: Cham, Switzerland, 2020; Volume 49, pp. 193–215. [CrossRef]
38. Wilson, T.; Bell, M. Comparative empirical evaluations of internal migration models in subnational population projections. *J. Popul. Res.* **2004**, *21*, 127–160. [CrossRef]
39. Matusik, S.; Pietrzak, M.B.; Wilk, J. Ekonomiczno-społeczne uwarunkowania migracji wewnętrznych w Polsce w świetle metody drzew klasyfikacyjnych. *Stud. Demogr.* **2012**, *162*, 3–28. [CrossRef]
40. Pietrzak, M.B.; Wilk, J. *The Impact of Metropolitan Areas on Internal Migrations in Poland: The Case of Southern Regions*; Institute of Economic Research Working Papers, No. 24/2013; Institute of Economic Research (IER): Toruń, Poland, 2013.
41. Hajj, H.M. Internal migration models and population forecasts. *J. Urban Plan. Dev. Div.* **1975**, *101*, 201–215. [CrossRef]
42. Fantazzini, D.; Pushchelenko, J.; Mironenkov, A.; Kurbatskii, A. Forecasting internal migration in Russia using Google Trends: Evidence from Moscow and Saint Petersburg. *Forecasting* **2021**, *3*, 774–803. [CrossRef]
43. Cadwallader, M.T. *Migration and Residential Mobility: Macro and Micro Approaches*; University of Wisconsin Press: Madison, WI, USA, 1992.
44. Woods, R. *Theoretical Population Geography*; Longman: London, UK, 1982.
45. Etzo, I. *Internal Migration: A Review of the Literature*; Munich Personal RePEc Archive, MPRA Paper No. 8783; University of Cagliari: Munich, 2008.
46. Ghatak, S.; Mulhern, A.; Watson, J. Inter-regional migration in transition economies: The case of Poland. *Rev. Dev. Econ.* **2008**, *12*, 209–222. [CrossRef]
47. Kupiszewski, M.; Berinde, D.; Teodorescu, V.; Durham, H.; Rees, R. *Internal Migration and Regional Population Dynamics in Europe: Romanian Case Study*; Working Paper 97/07; University of Leeds: Leeds, UK, 1997.
48. Kupiszewski, M.; Durham, H.; Rees, P. Internal migration and urban change in Poland. *Eur. J. Popul./Rev. Eur. Démographie* **1998**, *14*, 265–290. [CrossRef]
49. Fassmann, H.; Hintermann, C. Potential East-West migration: Demographic structure, motives and intentions. *Czech Sociol Rev.* **1998**, *1*, 59–72. [CrossRef]
50. Pietrzak, M.B.; Wilk, J.; Matusik, S. *Gravity Model as the Tool for Internal Migration Analysis in Poland in 2004–2010*; Institute of Economic Research Working Papers, No. 28/2013; Institute of Economic Research: Toruń, Poland, 2013.
51. Garcia, A.J.; Pindolia, D.K.; Lopiano, K.K.; Tatem, A.J. Modeling internal migration flows in sub-Saharan Africa using census microdata. *Migr. Stud.* **2015**, *3*, 89–110. [CrossRef]
52. Yüksel, S.; Eroğlu, S.; Özsarı, M. An analysis of the reasons for internal migration in Turkey with logit method. *Bus. Manag. Horiz.* **2016**, *4*, 34–45. [CrossRef]
53. Ackah, C.; Medvedev, D. Internal migration in Ghana: Determinants and welfare impacts. *Int. J. Soc. Econ.* **2012**, *39*, 764–784. [CrossRef]
54. Pham, N.H.; Tran, M.D.; Le, A.D.; Le, T.L. Determinants influencing the decision of internal migration in the context of an emerging country. *Corp. Gov. Organ. Behav. Rev.* **2021**, *5*, 90–98. [CrossRef]
55. Kamninga, T.M. Spillover effects of education on internal migration in Malawi. *Migr. Dev.* **2020**, *9*, 238–252. [CrossRef]
56. Bogusz, H.; Winnicki, S.; Wójcik, P. Assessing the role of internal migration in demographic processes in Poland after 1989. *Bull. Geogr. Socio-Econ. Ser.* **2019**, *44*, 27–36. [CrossRef]

57. Frees, E.W. Forecasting state-to-state migration rates. *J. Bus. Econ. Stat.* **1992**, *10*, 153–167. [CrossRef]
58. Frees, E.W. Short-term forecasting of internal migration. *Environ. Plan. A* **1993**, *25*, 1593–1606. [CrossRef]
59. *Prognoza Ludności na Lata 2023–2060 (Population Projection 2023–2060)*; Główny Urząd Statystyczny (Statistics Poland): Warszawa, Poland, 2023.
60. Radło-Kulisiewicz, M. Przegląd wybranych podejść w zakresie prognozowania rozwoju obszarów miast. *Arch. Fotogram. Kartogr. I Teledetekcji* **2015**, *27*, 109–122. [CrossRef]
61. Ovchynnikova, O.; Belovsky, C.; Khan, O. Neural Network Forecasting of International Population Migration. In Proceedings of the 2021 11th International Conference on Advanced Computer Information Technologies (ACIT), Deggendorf, Germany, 15–17 September 2021; IEEE: New York City, NY, USA, 2021; pp. 147–152. [CrossRef]
62. Qiokata, V.; Khan, M.G.M. Modeling emigration of Fiji's population using artificial neural network. In Proceedings of the 2015 2nd Asia-Pacific World Congress on Computer Science and Engineering (APWC on CSE), Nadi, Fiji, 2–4 December 2015; IEEE: New York City, NY, USA, 2015; pp. 1–8. [CrossRef]
63. Beim, M. *Modelowanie Procesu Suburbanizacji w Aglomeracji Poznańskiej*; Bogucki Wydawnictwo Naukowe: Poznań, Poland, 2009.
64. Gawryluk, A.; Kulisz, M.; Komor, A. Modelowanie salda migracji dla miasta Poznania z wykorzystaniem sztucznych sieci neuronowych. *Rozw. Reg. I Polityka Reg.* **2022**, *62*, 87–109. [CrossRef]
65. Kostrzewa, Z.; Stańczak, J. *Migracje Długookresowe Ludności w Latach 1989–2002 na Podstawie Ankiety Migracyjnej 2002*; Główny Urząd Statystyczny (Central Statistical Office): Warszawa, Poland, 2004. Available online: https://stat.gov.pl/files/gfx/portalinformacyjny/pl/defaultaktualnosci/5738/15/1/1/nsp2002_migracje_dlugookresowe_ludnosci_w_latach_1989-2002.pdf (accessed on 1 September 2024).
66. Kałuża, D. Migracje wewnętrzne a poziom rozwoju społeczno-gospodarczego wybranych największych miast Polski. *Acta Univ. Lodziensis. Folia Oeconomica* **2010**, *237*, 29–41. Available online: https://dspace.uni.lodz.pl/xmlui/handle/11089/382 (accessed on 1 September 2024).
67. Harańczyk, A. Procesy suburbanizacji w Krakowskim Obszarze Funkcjonalnym. *Stud. Miej.* **2015**, *18*, 85–102. Available online: https://www.studiamiejskie.uni.opole.pl/wp-content/uploads/2016/05/S_Miejskie_18_2015-Haranczyk-Procesy.pdf (accessed on 1 September 2024). [CrossRef]
68. Sasin, M. Czynniki różnicujące migracje stałe w województwie zachodniopomorskim. *Zesz. Nauk. Wydziału Nauk. Ekon.* **2021**, *1*, 131–144.
69. Józefowicz, K. Atrakcyjność migracyjna miast i obszarów wiejskich województwa wielkopolskiego. *Space Soc. Econ.* **2020**, *31*, 213–227. [CrossRef]
70. Zagórski, I.; Kulisz, M.; Kłonica, M.; Matuszak, J. Trochoidal Milling and Neural Networks Simulation of Magnesium Alloys. *Materials* **2019**, *12*, 2070. [CrossRef] [PubMed]
71. Kulisz, M.; Kujawska, J. Application of artificial neural network (ANN) for water quality index (WQI) prediction for the river Warta, Poland. *J. Phys. Conf. Ser.* **2021**, *2130*, 012028. [CrossRef]
72. Laajimi, R.; Gallo, J. Push and pull factors in Tunisian internal migration: The role of human capital. *Growth Chang.* **2022**, *53*, 771–799. [CrossRef]
73. Zhu, N.; Luo, X. The impact of migration on rural poverty and inequality: A case study in China. *Agric. Econ.* **2010**, *41*, 191–204. [CrossRef]
74. Ling, T.; Longley, P. An individual level method for improved estimation of ethnic characteristics. *Int. Reg. Sci. Rev.* **2022**, *46*, 328–353. [CrossRef]
75. Bigotte, J.; Antunes, A.; Krass, D.; Berman, O. The relationship between population dynamics and urban hierarchy. *Int. Reg. Sci. Rev.* **2014**, *37*, 149–171. [CrossRef]
76. Li, W.; Kockelman, K. How does machine learning compare to conventional econometrics for transport data sets? A test of ML versus MLE. *Growth Change* **2021**, *53*, 342–376. [CrossRef]
77. Wu, H.; Wu, Q. The geography of migrant workers' income in China: Evidence from a migrants dynamic survey in 2017. *Growth Change* **2021**, *52*, 1826–1851. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.