STATISTICS IN TRANSITION new series, March 2024
Vol. 25, No. 1, pp. 93–107, https://doi.org/10.59170/stattrans-2024-006
Received: 03.04.2022; accepted: 15.05.2023

Inter-voivodship migration in Poland in the 2000–2020 period based on Markov chain analysis

Agnieszka Palma[1], Dorota Kałuża-Kopias[2]

Abstract

The paper presents the scale and directions of inter-voivodship migration in Poland in selected years of the 2000–2020 period. The study focused on permanent residence migration and aimed to identify areas of migration attractiveness and migration catchment voivodships. To study the stochastic nature of these migrations, a Markov chain model was used, in which the states were voivodships. An important aspect of the study involved determining the properties of the transition probability matrix as well as stationary probability in order to characterise the mechanism of inter-voivodship migrations in the years 2000, 2010 and 2020. Data obtained from Statistics Poland were used in the analysis. The transition probability matrix showed that the states were connected and irreducible to each other, while the stationary probability of migration to Dolnośląskie, Małopolskie, Pomorskie, and Wielkopolskie voivodships increased in 2020 compared to 2000. The analysis of the mechanism of migration in the years 2000, 2010, 2020 indicated that Mazowieckie Voivodship was still the main destination for migrants, with the highest stationary probability reaching 0.18 in 2010.

Key words: inter-voivodship migration, Markov chain, random transition count, transition probability matrix, stationary probability, mechanism of migration.

1. Introduction

Migration, after births and deaths, is the third basic factor influencing the population dynamics of an area. Its importance in influencing population growth and decline and in modifying the demographic characteristics of areas of origin and destination has long been obvious and recognised. The measurement and analysis of migration are important in preparing population estimates and projections for a nation or part of a nation.

The issue of the migration phenomenon is very broad (foreign migration, internal migration) and concerns every spatial scale, starting from the national scale, through the regional (provincial) level, other selected functional administrative units, up to complex settlement systems (Śleszyński et al. 2018). In addition, it deals with explaining the reasons for the variation in migration volumes. Certainly, many factors influence the intensity and direction of migration flows. One of them is the pace of socio-economic development of a given region, employment opportunities, housing or land prices, cost of living, availability of infrastructure, to a large extent family ties, and the age of migrants.

[1] Institute of Statistics and Demography, Faculty of Economics and Sociology, University of Lodz, Poland. E-mail: agnieszka.palma@uni.lodz.pl. ORCID: https://orcid.org/0000-0002-3558-1568.

[2] Institute of Statistics and Demography, Faculty of Economics and Sociology, University of Lodz, Poland. E-mail: dorota.kaluza@uni.lodz.pl. ORCID: https://orcid.org/ 0000-0001-5023-2596.

The topic of internal migration for permanent residence in Poland has been addressed in several works. Pietrzak (2013) described inter-regional migration using a gravity model. Roszko (2018) found a strong relationship between the standard of living of residents expressed by Gross Domestic Product per capita and the number of incoming new residents and the balance of interprovincial migration. Several works have addressed internal migration in a specific province: Rosner (2014), Ilnicki (2020), Józefowicz (2020), Kałuża-Kopias (2021), Ilnicki (2021).

To investigate the stochastic nature of inter-regional migration in Poland in the period 2000–2020, a Markov chain model was used. Markov chains describe many real-world processes, and they are used in many different fields such as physics, geography, chemistry, biology, medicine, music, economics and finance, game theory, sports, and more. Many examples of their application can be found in the literature, see for example: Clark (1960), Marble (1967), Iosifescu (2007), Privault (2018). Collins (1972) used Markov chains in forecasting industrial migration, while Berry (1971) in outlined a short-term model of neighbourhood turnover, Bourne (1976) suggested this method to monitor changes in Toronto's spatial structure, Azizah (2019) applied Markov chain to forecast rainfall data and Chu (2020) used a Markov chain model to forecast future land use change, Barra (2020) to count and model migraine attacks and Romeu (2020) to analyse covid-19 survival.

Markov chain models are useful and convenient tools for describing and analysing the nature of dynamic changes of a phenomenon or process of interest. They are an important tool for geographers who deal with mobility problems. These can be movements from one place to another, as well as movements from one state to another. State can be defined in different ways, it can refer to a class of provinces, municipalities, city size or income or type of land use or some other variable. They can also be used to forecast future changes: Sempewo (2016), Rahimipour (2018).

One of the applications of Markov chains is their use in migration studies, mainly to determine the dominant direction or rate of change, as well as the development of a system, for example, an urban system. Thanks to Markov chains we can determine which cities, provinces, or other territorial units have a tendency to increase in population and which to decrease. The study takes into account permanent migration movements between provinces to analyse the changes taking place.

In order to analyse the changes taking place in the years 2000–2020, the migration movements for permanent residence taking place between provinces were taken into account, and the probability matrices of the transition between states and the corresponding stationary distributions were determined. The empirical material used in the study was data from the Local Data Bank of Statistics Poland concerning the inflow and outflow of the population in individual voivodships. They allowed measuring the migration volume with the Markov chain model. It should be emphasised that data from public statistics do not allow for a reliable assessment of the migration situation in Poland. It results mainly from the fact that they are based on registration data, thus they do not register real permanent migration, which is not connected with checking out or registering. Despite the continuous improvement of current migration statistics both in terms of data collection methods and compilation techniques, some authors are critical of their quality (e.g. Jończy 2014; Śleszyński 2005, 2011). Theoretically, the number of previous and current places of registration for migrants should be equal, but in practice, the differences can be significant. The largest underestimation of internal migration concerns large cities and their suburban zones, as a large part of the inflow remains unregistered, Korcelii (1997). Although the source of information on migration is not perfect, it should be emphasised that data from registration registers, due to their continuity and updating, form the basis of migration reporting.

2. Migration between voivodships in Poland in dynamic terms

The last two decades have seen a decrease in the size of internal migration in Poland. Between 2000 and 2020, on average, the number of inter-regional migrations decreased by about 0.5% from year to year, (the calculated geometric mean value for 2000–2020 was 0.995). In spatial terms, the effect of these movements is the migration balance.

Table 1: Internal migration by voivodships

	Net migration			Net migration per 1000 people			Index of migration attractiveness		
voivodships	**2000**	**2010**	**2020**	**2000**	**2010**	**2020**	**2000**	**2010**	**2020**
Dolnośląskie	-573	1579	3295	-0,2	0,54	1,14	-0,04	0,1	0,22
Kujawsko-pomorskie	-407	-1443	-2112	-0,2	-0,69	-1,02	-0,04	-0,14	-0,21
Lubelskie	-2969	-4867	-4685	-1,35	-2,23	-2,23	-0,27	-0,41	-0,43
Lubuskie	-440	-474	-842	-0,44	-0,46	-0,83	-0,06	-0,07	-0,14
Łódzkie	-1107	-1757	-1812	-0,42	-0,69	-0,74	-0,09	-0,15	-0,19
Małopolskie	2376	3673	3412	0,74	1,1	1	0,16	0,24	0,23
Mazowieckie	8825	12687	10448	1,73	2,41	1,92	0,32	0,4	0,38
Opolskie	-88	-671	-747	-0,08	-0,66	-0,76	-0,01	-0,1	-0,13
Podkarpackie	-1730	-1973	-2226	-0,82	-0,93	-1,05	-0,2	-0,22	-0,25
Podlaskie	-1255	-1616	-1702	-1,04	-1,34	-1,45	-0,19	-0,26	-0,31
Pomorskie	1651	2749	3825	0,76	1,21	1,63	0,14	0,22	0,3
Śląskie	-1652	-3194	-3342	-0,35	-0,69	-0,74	-0,07	-0,16	-0,2
Świętokrzyskie	-2061	-2567	-2143	-1,58	-2	-1,74	-0,23	-0,31	-0,31
Warmińsko-Mazurskie	-2002	-2721	-2113	-1,4	-1,87	-1,49	-0,18	-0,26	-0,23
Wielkopolskie	1595	1706	1491	0,48	0,5	0,43	0,11	0,12	0,11
Zachodniopomorskie	-163	-1111	-747	-0,1	-0,64	-0,44	-0,01	-0,11	-0,09

Source: Authors' own calculations.

Most of the voivodships in Poland were characterized by a negative balance of inter-voivodship movements. In 2000, 12 voivodships had negative migration balances. The regions with the lowest urbanisation level - Świętokrzyskie and Lubelskie - had relatively the largest population migration losses exceeding 2,000 people (Table 1). The largest positive net migration volumes occurred in the Mazowieckie, Pomorskie, Wielkopolskie, and Małopolskie voivodships. In subsequent years, Dolnośląskie joined the group of voivodships with a positive migration balance. The regions characterized by permanent migration losses of the population were traditionally the areas of eastern and north-eastern Poland, which are mostly poorly urbanised agricultural areas. Relating migration balances to the population in individual voivodships, it turns out that in 2000, only in one region out of four with positive inter-voivodship migration balances, there was a migration increase exceeding 1 person per 1000 population (Table 1).

As regards the intensity of the migration loss of inhabitants, the northeastern voivodships had relatively the largest negative balances: Warmińsko-Mazurskie (-1.4 persons per 1,000 inhabitants) and Świętokrzyskie (-1.58). Over the years, there was a clear advantage in the size of positive migration balances (both in absolute and relative terms) of the Mazowieckie voivodship over other regions with higher inflows than outflows. Similarly, in the case of voivodships characterized by negative migration balances, in the subsequent years the migration loss of inhabitants assumed (in most regions) greater dimensions than in 2000. It should be noted, however, that - when assessing the size of migration - it is worth relating the size of migration balances to the overall migration turnover (inflow and outflow), as the same values of migration balances may be the result of a large inflow and outflow on the one hand or a much smaller size of these components (Obraniak 2007). As a measure of migration attractiveness (in demographic literature, the name "migration efficiency index" is also used), an index representing the ratio of migration balance to migration turnover was adopted (Table 1). Due to its construction and ease of interpretation, it is a quite commonly used measure (Potrykowska, Śleszyński 1999; Obraniak 2007; Kałuża-Kopias 2021). The calculated values of the index for 2000 indicate that the Mazowieckie voivodship was the most attractive in terms of migration (Table 1). The following places were occupied by agglomeration voivodships: Małopolskie, Pomorskie, Wielkopolskie. The list was closed by the voivodships of north-eastern, eastern, and central Poland - Lubelskie, Podlaskie, Warmińsko-Mazurskie, Podkarpackie, and Świętokrzyskie.

3. Mathematical model description and calculations

3.1. Markov chain model.

Let n and k be elements of $\mathbf{N}$, such that $n \geq 1$ and $k \geq 1$. Define $S = \{1, \ldots, k\}$. Consider a sequence of random variables $\{X_1, X_2, \ldots, X_n\}$ such that

$$p_{ij} = P(X_{d+1} = j | X_d = i) = P(X_{d+1} = j | X_1 = i_1,\ X_2 = i_2,\ \ldots,\ X_d = i)$$

is independent of d for all $i, j \in S$. Then, the sequence $\{X_1, X_2, \ldots, X_n\}$ is a first-order Markov chain with state space S and transition probabilities p_{ij} for $i, j \in S$.

The process starts in one of the states and moves successively from one state to another. Each move is called a step. If the chain is currently in state i, then it moves to state j at the next step with a probability denoted by p_{ij}, and this probability does not depend upon which state the chain was in before the current state.

The transition probability matrix $P = [p_{ij}]_{i,j \in S}$ of finite Markov chains, is defined as follows:

$$P = \begin{bmatrix} p_{11} & p_{12} & \cdots & p_{1k} \\ p_{21} & p_{22} & \cdots & p_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ p_{k1} & p_{k2} & \cdots & p_{kk} \end{bmatrix}. \tag{1}$$

Since the elements of row i of this matrix represent the conditional probabilities for all possible state changes from state i, they must satisfy

$$\forall_{i,j \in S} \quad p_{ij} \geq 0, \quad \sum_{j=1}^{k} p_{ij} = 1 \text{ for each } i \in S. \tag{2}$$

A square matrix for which all elements are non-negative and the sum of the elements in each row is 1 is called a stochastic matrix. It follows from this definition that a Markov chain with known probability distribution of the initial state is completely characterized by a $k \times k$ matrix containing the transition probabilities p_{ij}.

Transition probability matrix P shows transition probabilities from one state to another in one step of Markov chain. Transition probabilities from state i to state j in n steps is noted as

$$p_{ij}^{(n)} = P(X_{d+n} = j | X_d = i) \text{ for } i, j \in S, \text{ and } d, n \in \mathbf{N}. \quad (3)$$

Transition probability matrix in n steps is $P^{(n)}$.

Throughout the paper we are dealing with a random transition count matrix M, which is defined in the usual way as

$$M = \begin{bmatrix} m_{11} & m_{12} & \dots & m_{1k} \\ m_{21} & m_{22} & \dots & m_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ m_{k1} & m_{k2} & \dots & m_{kk} \end{bmatrix}, \quad (4)$$

where m_{ij} denotes number of transitions from state i to state j, where $i, j \in S$.

Let $\sum_{j=1}^{k} m_{ij} = m_{i.}$, $\sum_{i=1}^{k} m_{ij} = m_{.j}$ and $\sum_{i,j=1}^{k} m_{ij} = m_{..}$.

Then,

Table 2: Random transition count

States	1	2	...	k	Total
1	m_{11}	m_{12}	...	m_{1k}	$m_{1.}$
2	m_{21}	m_{22}	...	m_{2k}	$m_{2.}$
⋮	⋮	⋮	⋱	⋮	⋮
k	m_{k1}	m_{k2}	...	m_{kk}	$m_{k.}$
	$m_{.1}$	$m_{.2}$	...	$m_{.k}$	$m_{..}$

We are interested in the estimation of the elements p_{ij} of matrix P; we denote them by $\hat{p}_{ij}$.

Using the maximum likelihood estimation method (Bhat and Miller, 2002; Billingsley, 1961) we obtain an estimate of the matrix P and denote it as

$$P = [\hat{p}_{ij}], \text{ where } \hat{p}_{ij} = \frac{m_{ij}}{m_{i.}} \quad \text{for each} \quad i, j \in \{1, \dots, k\}. \quad (5)$$

3.2. Markov chain's state space and calculations of migration

In our Markov chain model the states are voivodships. For the convenience of the reader we denote voivodships respectively $1, 2, \dots, 16$ and we give this below.

Table 3: States and corresponding voivodships

States	voivodships
1	Dolnośląskie
2	Kujawsko-Pomorskie
3	Lubelskie
4	Lubuskie
5	Łódzkie
6	Małopolskie
7	Mazowieckie
8	Opolskie
9	Podkarpackie
10	Podlaskie
11	Pomorskie
12	Śląskie
13	Świętokrzyskie
14	Warmińsko-Mazurskie
15	Wielkopolskie
16	Zachodniopomorskie

In this way we obtain the state space $S = \{1, 2, \ldots, 16\}$.

Input data were obtained from Statistics Poland. They are available in the Local Data Bank (BDL) and the Demography database. Data on migration volumes are presented in a matrix system. The matrix shows the internal migration of the population for permanent residence by voivodship of previous and current place of residence. The elements $m_{ij},\ i, j \in S$, of the matrix $M_i,\ i = 1, 2, 3$, denote the number of migrants who emigrated and stayed in another voivodship in Poland, respectively in the years 2000, 2010, 2020. Migrants are people who changed their place of permanent residence and moved to another province.

Based on the input data $M_i,\ i = 1, 2, 3$, after applying the maximum likelihood estimation defined by formula (5), we obtain an estimate of the probability of migration from i- th province to j-th province. In this way, the transition probability matrices $P_1,\ P_2,\ P_3$ for migration in Poland were determined for the years 2000, 2010, and 2020, respectively.

3.3. Test for first-order Markov chain

In this section, we will check whether the transition probability matrix for migration in 2000, 2010, and 2020 satisfies the assumption of a first-order Markov chain. To check the validity of the Markov chain model, the chi-square test of goodness of fit is used. The hypothesis to be tested is the null hypothesis that the collected observations are independent of the alternative hypothesis that the observed process is a first-order Markov chain. The hypothesis is $Ho : P = P_0$, where P_0 has identical rows under the assumption of independence. The χ^2 statistic to test independence against the first order Markov Chain (Bhat and Miller, 2002; Billingsley, 1961) is calculated from the following relationship:

$$\chi^2 = \sum_{i=1}^{k}\sum_{j=1}^{k} \frac{(m_{ij} - \frac{m_{i.}m_{.j}}{m_{..}})^2}{\frac{m_{i.}m_{.j}}{m_{..}}}$$

with the degrees of freedom $(k-1)^2 - d$, where d is number of zero cells. For the input data, χ^2 statistics were determined as well as Cramer's association coefficient. The results are given below:

Table 4: χ^2 test of independence for data migration in 2000, 2010, 2020

Migration data	χ^2 statistic	p-value	Cramer's V
2000	82103	<0.0000	0.234
2010	85064	<0.0000	0.238
2020	85425	<0.0000	0.251

Source: Authors' own calculations.

It follows from Table 4, that the assumption of independence can be rejected (p-value< 0.0000). Based on these results, it can be concluded that modelling the data as a first-order Markov chain is reasonable.

3.4. Stationary distribution

As we progress through time, the probability of being in certain states more likely than others. Over the long run, the distribution will reach equilibrium with an associated probability of being in each state. This is known as the stationary distribution. A stationary distribution π is a vector whose entries are non-negative and sum to 1, is unchanged by the operation of transition matrix P on it, so it satisfied (Bhat 2002),

$$\pi_j = \sum_{i \in S} \pi_i \cdot p_{ij} \quad \text{and} \quad \forall_{j \in S}\ \pi_j \geq 0, \quad \sum_j \pi_j = 1.$$

In the matrix notation, it can be written as

$$\pi = \pi \cdot P_i \text{ for } i = 1,2,3,$$

where π is some distribution, which is a row vector with the number of columns equal to the states in the state space S and $P_i,\ i = 1,2,3$, is the transition probability matrix.

The stationary probability can also be found from limiting transition probability matrix

$$\lim_{n \to \infty} p_{ij}^{(n)} = \pi_j.$$

Comparison of stationary distribution between the years 2000, 2010, 2020 for individual voivodships is presented in Figure 1.

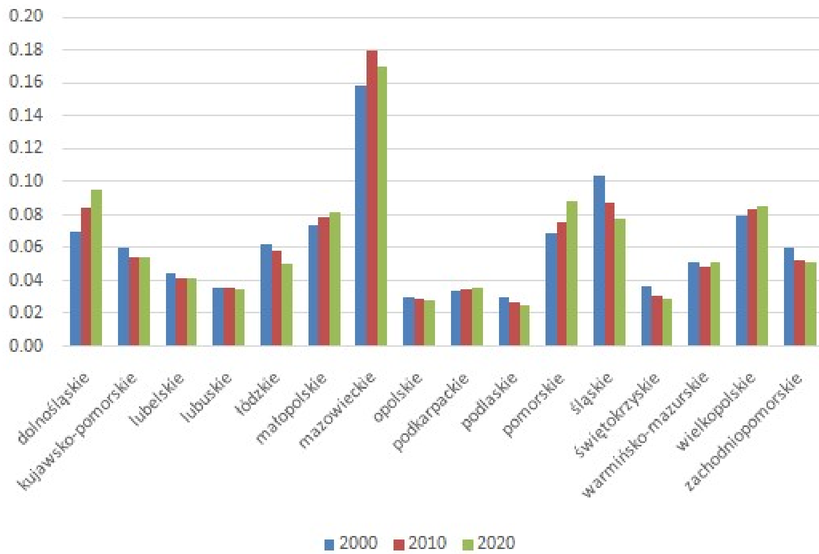

Figure 1: Comparison of stationary inter-voivodship migration distributions in Poland in 2000, 2010, 2020

Table 5: Stationary probability π

Transition matrix	π_1	π_2	π_3	π_4	π_5	π_6	π_7	π_8	π_9	π_{10}	π_{11}	π_{12}	π_{13}	π_{14}	π_{15}	π_{16}
P_1	0.07	0.06	0.04	0.04	0.06	0.07	0.16	0.03	0.03	0.03	0.07	0.1	0.04	0.05	0.08	0.06
P_2	0.08	0.05	0.04	0.04	0.06	0.08	0.18	0.03	0.04	0.03	0.08	0.09	0.03	0.05	0.08	0.05
P_3	0.1	0.05	0.04	0.04	0.05	0.08	0.17	0.03	0.04	0.03	0.09	0.08	0.03	0.05	0.09	0.05

Source: Authors' own calculations.

Figure 1 and Table 5 - the results of the analysis of stationary probability show that the variation for the following provinces: Dolnośląskie, Mazowieckie, Pomorskie, Śląskie is greater compared to the remaining provinces, for which the stationary distributions are differentiated, but the changes are relatively small. The values of stationary probabilities indicate that the probability of migration to Łódzkie decreased from 0.062 in 2000 to 0.05 in 2020, while for Śląskie from 0.103 to 0.077. On the other hand, the probability of migration to Dolnośląskie increased from 0.07 to 0.095, for the Pomorskie voivodship from 0.069 to 0.089, analogically, for the Podkarpackie and Wielkopolskie voivodships, the probabilities increased, although to a smaller extent. Based on these results, one may conclude that the Mazowieckie voivodship will continue to be the main direction of migration in the future. However, it is worth noting that in 2020 the probability of migration to the Mazowieckie voivodship decreased in comparison to 2010.

3.5. Mechanism of the inter-voivodship migration in Poland

In order to characterize the mechanism of interprovincial migration in Poland, one should first estimate the transition probability matrix using the procedure given in Bhat and Miller (2002) and Miall (1973) as follows:

$$P^0 = \begin{bmatrix} 0 & \frac{m_{.2}}{m_{..}-m_{.1}} & \frac{m_{.3}}{m_{..}-m_{.1}} & \cdots & \frac{m_{.k}}{m_{..}-m_{.1}} \\ \frac{m_{.1}}{m_{..}-m_{.2}} & 0 & \frac{m_{.3}}{m_{..}-m_{.2}} & \cdots & \frac{m_{.k}}{m_{..}-m_{.2}} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \frac{m_{.1}}{m_{..}-m_{.k}} & \frac{m_{.2}}{m_{..}-m_{.k}} & \frac{m_{.3}}{m_{..}-m_{.k}} & \cdots & 0 \end{bmatrix}. \tag{6}$$

To analyze the mechanism of the inter-voivodship migration in Poland, we find the estimated transition probability matrix P^0 by formula (6) and using M_1, M_2, M_3 - data from Statistics Poland for the 2000, 2010, 2020, respectively. The results of these calculations are presented, as an example, only for the matrix P_1^0. The matrices P_2^0, P_3^0 are determined analogically.

$$P_1^0 = \begin{bmatrix}
0 & 0.06 & 0.04 & 0.04 & 0.06 & 0.09 & 0.2 & 0.03 & 0.04 & 0.03 & 0.07 & 0.11 & 0.04 & 0.05 & 0.08 & 0.06 \\
0.07 & 0 & 0.04 & 0.04 & 0.06 & 0.09 & 0.19 & 0.03 & 0.04 & 0.03 & 0.07 & 0.11 & 0.04 & 0.05 & 0.08 & 0.06 \\
0.07 & 0.06 & 0 & 0.04 & 0.06 & 0.09 & 0.19 & 0.03 & 0.04 & 0.03 & 0.07 & 0.11 & 0.04 & 0.05 & 0.08 & 0.06 \\
0.07 & 0.05 & 0.04 & 0 & 0.06 & 0.09 & 0.19 & 0.03 & 0.04 & 0.03 & 0.07 & 0.11 & 0.04 & 0.05 & 0.08 & 0.06 \\
0.07 & 0.06 & 0.04 & 0.04 & 0 & 0.09 & 0.19 & 0.03 & 0.04 & 0.03 & 0.07 & 0.11 & 0.04 & 0.05 & 0.08 & 0.06 \\
0.08 & 0.06 & 0.04 & 0.04 & 0.06 & 0 & 0.2 & 0.03 & 0.04 & 0.03 & 0.07 & 0.11 & 0.04 & 0.05 & 0.09 & 0.06 \\
0.09 & 0.06 & 0.05 & 0.04 & 0.07 & 0.1 & 0 & 0.04 & 0.04 & 0.03 & 0.08 & 0.13 & 0.04 & 0.05 & 0.1 & 0.07 \\
0.07 & 0.05 & 0.04 & 0.04 & 0.06 & 0.09 & 0.19 & 0 & 0.04 & 0.03 & 0.07 & 0.11 & 0.04 & 0.05 & 0.08 & 0.06 \\
0.07 & 0.05 & 0.04 & 0.04 & 0.06 & 0.09 & 0.19 & 0.03 & 0 & 0.03 & 0.07 & 0.11 & 0.04 & 0.05 & 0.08 & 0.06 \\
0.07 & 0.05 & 0.04 & 0.04 & 0.06 & 0.09 & 0.19 & 0.03 & 0.04 & 0 & 0.07 & 0.11 & 0.04 & 0.05 & 0.08 & 0.06 \\
0.07 & 0.06 & 0.04 & 0.04 & 0.06 & 0.09 & 0.2 & 0.03 & 0.04 & 0.03 & 0 & 0.11 & 0.04 & 0.05 & 0.08 & 0.06 \\
0.08 & 0.06 & 0.04 & 0.04 & 0.06 & 0.09 & 0.2 & 0.04 & 0.04 & 0.03 & 0.08 & 0 & 0.04 & 0.05 & 0.09 & 0.06 \\
0.07 & 0.05 & 0.04 & 0.04 & 0.06 & 0.09 & 0.19 & 0.03 & 0.04 & 0.03 & 0.07 & 0.11 & 0 & 0.05 & 0.08 & 0.06 \\
0.07 & 0.06 & 0.04 & 0.04 & 0.06 & 0.09 & 0.19 & 0.03 & 0.04 & 0.03 & 0.07 & 0.11 & 0.04 & 0 & 0.08 & 0.06 \\
0.08 & 0.06 & 0.04 & 0.04 & 0.06 & 0.09 & 0.2 & 0.03 & 0.04 & 0.03 & 0.07 & 0.11 & 0.04 & 0.05 & 0 & 0.06 \\
0.07 & 0.06 & 0.04 & 0.04 & 0.06 & 0.09 & 0.19 & 0.03 & 0.04 & 0.03 & 0.07 & 0.11 & 0.04 & 0.05 & 0.08 & 0
\end{bmatrix},$$

In the next step, transition probability difference matrices is created, assuming that the transition probabilities are given by (5) and (6), so that the properties of the migration mechanism can be inferred. In this case, we compute the difference matrices

$$D_i = P_i - P_i^0 \quad \text{for each} \quad i \in \{1,2,3\}.$$

Below, we present only the matrix D_1, to illustrate the method used.

$$D_1 = \begin{bmatrix}
0 & -0.03 & -0.01 & 0.08 & 0 & -0.02 & -0.08 & 0.09 & -0.01 & -0.01 & -0.04 & -0.01 & -0.01 & -0.03 & 0.07 & 0 \\
-0.04 & 0 & -0.03 & -0.01 & -0.02 & -0.07 & -0.01 & -0.03 & -0.03 & -0.02 & 0.13 & -0.05 & -0.03 & 0.03 & 0.13 & 0.03 \\
-0.03 & -0.04 & 0 & -0.02 & -0.03 & -0.02 & 0.27 & -0.02 & 0.05 & -0.01 & -0.02 & -0.03 & 0 & -0.03 & -0.06 & -0.02 \\
0.17 & -0.02 & -0.03 & 0 & -0.02 & -0.06 & -0.12 & -0.01 & -0.02 & -0.02 & -0.04 & -0.06 & -0.03 & -0.03 & 0.18 & 0.1 \\
-0.01 & -0.01 & -0.02 & -0.02 & 0 & -0.05 & 0.13 & 0 & -0.02 & -0.01 & -0.03 & 0.05 & 0 & -0.03 & 0.03 & -0.01 \\
-0.02 & -0.04 & -0.01 & -0.02 & -0.03 & 0 & -0.07 & -0.01 & 0.09 & -0.02 & -0.05 & 0.26 & 0.05 & -0.04 & -0.06 & -0.04 \\
-0.04 & 0.01 & 0.07 & -0.03 & 0.06 & -0.05 & 0 & -0.02 & -0.01 & 0.05 & 0 & -0.05 & 0.03 & 0.06 & -0.05 & -0.02 \\
0.2 & -0.04 & -0.03 & -0.01 & 0.01 & -0.03 & -0.1 & 0 & -0.01 & -0.02 & -0.06 & 0.2 & -0.02 & -0.03 & -0.03 & -0.03 \\
-0.02 & -0.04 & 0.08 & -0.02 & -0.03 & 0.22 & -0.03 & -0.02 & 0 & -0.02 & -0.05 & 0.02 & 0.03 & -0.03 & -0.05 & -0.03 \\
-0.05 & -0.03 & -0.01 & -0.03 & -0.03 & -0.06 & 0.26 & -0.03 & -0.02 & 0 & 0.01 & -0.06 & -0.03 & 0.15 & -0.06 & -0.03 \\
-0.04 & 0.11 & -0.02 & -0.01 & -0.02 & -0.06 & -0.02 & -0.02 & -0.02 & 0 & 0 & -0.07 & -0.02 & 0.09 & 0.01 & 0.11 \\
0.01 & -0.02 & 0 & -0.02 & 0.02 & 0.14 & -0.09 & 0.05 & 0.01 & -0.01 & -0.04 & 0 & 0.03 & -0.01 & -0.04 & -0.02 \\
-0.03 & -0.05 & 0 & -0.03 & 0.01 & 0.09 & 0.08 & -0.02 & 0.04 & -0.02 & -0.04 & 0.11 & 0 & -0.04 & -0.06 & -0.04 \\
-0.05 & 0.04 & -0.02 & -0.02 & -0.03 & -0.07 & 0.09 & -0.02 & -0.03 & 0.11 & 0.14 & -0.05 & -0.03 & 0 & -0.04 & -0.02 \\
0.07 & 0.11 & -0.02 & 0.08 & 0.03 & -0.06 & -0.09 & -0.01 & -0.03 & -0.02 & -0.01 & -0.05 & -0.03 & -0.03 & 0 & 0.06 \\
0 & 0.04 & -0.01 & 0.06 & -0.01 & -0.06 & -0.05 & -0.02 & -0.02 & -0.02 & 0.09 & -0.05 & -0.02 & -0.02 & 0.1 & 0
\end{bmatrix},$$

The positive elements of the matrices D_1, D_2, D_3 represent those transitions which have higher probability of occurrence. This makes it possible to characterize the mechanism of the process of interregional migration in the years $2000, 2010, 2020$.

The stationary distribution allowed finding voivodships that are migration catchment areas, while the values of the matrix D_i, $i = 1,2,3$, indicate areas attractive to migration for individual voivodships.

Table 6: Positive values of matrix D_i, $i = 1,2,3$, i.e. mechanism of migration between voivodships in the years 2000, 2010, 2020

state	Dolnośląskie	Kujawsko-Pomorskie	Lubelskie	Lubuskie	Łódzkie	Małopolskie	Mazowieckie	Opolskie	Podkarpackie	Podlaskie	Pomorskie	Śląskie	Świętokrzyskie	Warmińsko-Mazurskie	Wielkopolskie	Zachodnio-pomorskie	year
Dolnośląskie				0.08 0.10 0.08	 0.01			0.09 0.08 0.09				 0.01 0.01			0.07 0.08 0.08	 0.01	2000 2010 2020
Kujawsko-Pomorskie											0.13 0.16 0.14			0.03 0.01 0.02	0.13 0.16 0.15	0.03 0.01 0.01	2000 2010 2020
Lubelskie							0.27 0.32 0.33		0.05 0.04 0.05								2000 2010 2020
Lubuskie	0.17 0.15 0.20														0.18 0.20 0.21	0.10 0.08 0.10	2000 2010 2020
Łódzkie	 0.01 0.04						0.13 0.15 0.12					0.05 0.03 0.03			0.03 0.02 0.03		2000 2010 2020
Małopolskie									0.09 0.10 0.11			0.26 0.26 0.24	0.05 0.03 0.04				2000 2010 2020
Mazowieckie		0.01 0.01 0.01	0.07 0.08 0.09		0.06 0.07 0.05					0.05 0.05 0.05			0.03 0.02 0.02	0.06 0.05 0.06			2000 2010 2020
Opolskie	0.20 0.27 0.31				0.01							0.20 0.17 0.15					2000 2010 2020
Podkarpackie			0.08 0.05 0.05			0.22 0.24 0.29						0.02 0.01	0.03 0.03 0.02				2000 2010 2020
Podlaskie							0.26 0.28 0.29				0.01			0.15 0.13 0.12			2000 2010 2020
Pomorskie		0.11 0.11 0.11												0.09 0.01 0.12	0.01	0.11 0.10 0.10	2000 2010 2020
Śląskie	0.01 0.01 0.02				0.02 0.01 0.02	0.14 0.16 0.17		0.05 0.05 0.06	0.01 0.01				0.03 0.03 0.02				2000 2010 2020
Świętokrzyskie			 0.01 0.01		0.01 0.01 0.01	0.09 0.11 0.14	0.08 0.09 0.10		0.04 0.04 0.04			0.11 0.04 0.02					2000 2010 2020
Warmińsko-Mazurskie		0.04 0.03 0.03					0.09 0.07 0.06			0.11 0.07 0.07	0.14 0.20 0.25						2000 2010 2020
Wielkopolskie	0.07 0.09 0.11	0.11 0.09 0.08		0.08 0.06 0.06	0.03 0.03 0.03											0.06 0.07 0.06	2000 2010 2020
Zachodnio-pomorskie		0.04 0.02 0.01		0.06 0.08 0.08							0.09 0.08 0.08				0.10 0.11 0.14		2000 2010 2020

Source: Authors' own calculations.

Figure 2 presents graphically the mechanism of inter-voivodship migration in 2020. Individual voivodships have been marked with appropriate colours, depending on the values of probabilities $\pi_i,\ i \in \{1,2,\cdots,16\}$ of the stationary distribution P_3 in Table 5. The arrows between provinces are of different thickness, corresponding to the values of the matrix D_3 and showing the intensity of migration between provinces. For the convenience of the reader, the arrows coming out of different provinces are marked with different colours.

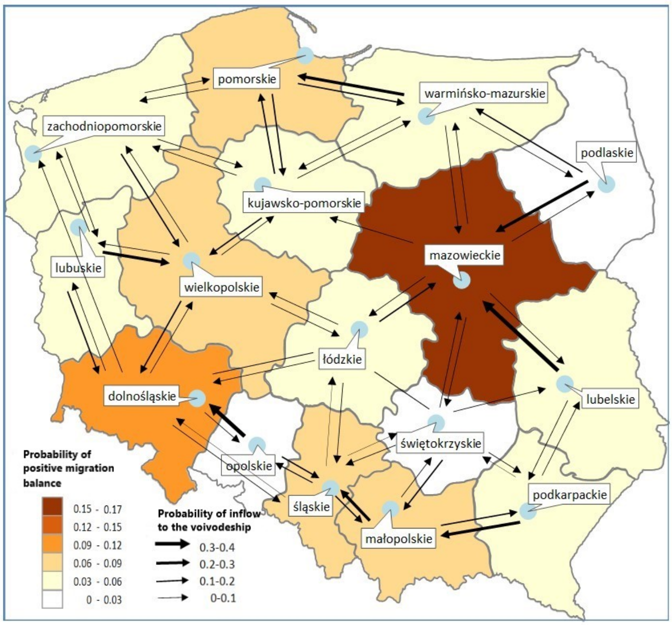

Figure 2: The mechanism of the inter-voivodship migration in Poland in 2020

4. Conclusions

The aim of this study was to analyse the scale and spatial range of inter-voivodship migrations of Poland's population in the years 2000, 2010, 2020. The application of the Markov chain model allows for a meticulous evaluation of the population flow between individual voivodships. The results of the study indicate that the most favourable situation remains in the Mazowieckie voivodship, which is the most attractive area of settlement for people from other regions of the country, mainly from the Lubelskie, Podlaskie, and Łódzkie voivodships, and to a lesser extent from Świętokrzyskie and Warmińsko-Mazurskie voivodships. However, during the analysed period, the inflow from the Łódzkie and Warmińsko-Mazurskie voivodships decreased, while for the remaining voivodships the inflow increased. The Mazowieckie voivodship is the most populous region in Poland and the largest in terms of area, and also very spatially diversified in terms of socio-economic development (Struzik 2007). The Dolnośląskie voivodship is ranked next. It is an attractive area for migrants from the Opolskie, Lubuskie, and Wielkopolskie voivodships. The inflow to the Dolnośląskie voivodship from those voivodships increased in the analysed period. In Figure 2, Wielkopolskie, Śląskie, Małopolskie, and Pomorskie voivodships are marked with the same colour and their probability of stationary distribution is at the level of 0.09-0.12. Migrants come to Wielkopolskie voivodship from Lubuskie, Zachodniopomorskie, Kujawsko-Pomorskie, Dolnośląskie, and to a small extent from Łódzkie. The weakest regions are Opolskie, Świętokrzyskie, and Podlaskie, which are not attractive in terms of migration. Migratory movement of people between voivodships is characterized by an inflow of people mainly from the areas of the neighbouring voivodships.

In reality, the inflow to the most attractive voivodships may be much higher. Due to imperfections in the analysed statistical data, the probabilities of inter-voivodship migration calculated on their basis reflect only the main trends in migration movement in the analysed period.

Certainly, many factors influence the intensity and direction of interprovincial movements. One of them is the pace of social and economic development of a given voivodship, more precisely, the possibility of employment, housing or land prices, cost of living, accessibility to infrastructure, to a large extent family ties, as well as the age of the migrants. According to Kałuża-Kopias (2021), we find that the most mobile group is made up of people aged 25-29, who have the largest spatial extent of migration. The least mobile group includes those aged 35-39 (people who have achieved stabilisation in the labour market and in their family life, and are presumably moving in order to improve their quality of life and rather over short distances) and 65-69 year-olds (retired people).

Another impacting factor for the scale of inter-voivodship flows is migration policy, which is the responsibility of the central government. Local governments have little influence on the government's migration policy instruments (Gońda 2021, Leśniewska, Matuszczyk 2018). However, they can impact the migration decisions of residents by introducing appropriate measures to encourage arrivals and settlement. In the case of all provinces, references to the issue of migration can be found in strategic documents on regional development [3]. In these strategies, the problem of migration is discussed only in the context of the demographic situation in the region as a factor accelerating population ageing and depopulation in the voivodship. Undoubtedly, conducting an active and conscious population policy taking into account the problem of population migration should become an important element of the regional development policy.

Rakowska (2009) points out that since the subject of the study is inter-voivodship migrations, the scale of population inflow will also depend on the size of the voivodship area, the number of its inhabitants, the number of rural communes and towns, its geographical location in relation to other voivodships, as well as socio-economic ties with the nearest surroundings and the level of development and the rate of economic growth.

These subjects are another challenge in migration research.

[3] https://strateg.stat.gov.pl//strategie/wojewodzkie

References

Azizah, A., Welastika, R., Nur Falah, A., Ruchjana, B., Abdullah, A., (2019). An Application of Markov Chain for Predicting Rainfall Data at West Java using Data Mining Approach, IOP Conf. Ser.: Earth Environ. Sci. 303 012026.

Barra, M., Dahl, F. A., Vetvik, K.G. et al., (2020). A Markov chain method for counting and modelling migraine attacks. Sci Rep 10, 3631.

Bilingsley P., (1961). Statistical Methods in Markov Chain , *Ann. Math. Stat*, 32, pp. 12–40.

Bhat U. N., Miller, K., (2002). Elements of Applied Stochastic Processes.*New York: John Wiley and Sons.*

Constant A., Zimmermann K., (2012). The Dynamics of Repeat Migration: A Markov Chain Analysis. *International Migration Review*, Vol. 46, No. 2, The Center for Migration Studies of New York, Inc., pp. 362-88.

Collins L., (1975). An introduction to Markov chain analysis. *Geo Abstracts,* Norwich, CT.

Gońda M., (2021). Czy imigracja to remedium dla wyludniających się regionów? Cudzoziemcy w dokumentach strategicznych województw łódzkiego i śląskiego, *Polityka Społeczna*, t. XLVIII, No. 9, pp. 15–24.

Ilnicki, D., (2020). Wielkość i kierunki migracji na pobyt stały w województwie wielkopolskim w latach 2002–2017. *Rozwój Regionalny i Polityka Regionalna*, (52), 141–160. https://doi.org/10.14746/rrpr.2020.52.09

Ilnicki, D., Szczyrba, Z., (2021). Migracje wewnętrzne na pobyt stały w makroregionie południowo-zachodnim w latach 2002–2017. *Studia Miejskie*, 36, pp. 25–44. https://doi.org/10.25167/sm.1534

Iosifescu, M., (2007). Finite Markov Processes and Applications. *Dover, New York.*

Jończy R., (2014). Problem nierejestrowanej emigracji definitywnej (emigracji zawieszonej) w badaniu procesów społeczno-gospodarczych na obszarach wiejskich, *Prace Naukowe Uniwersytetu Ekonomicznego we Wrocławiu*, 360, pp. 11–18.

Józefowicz, K., (2020). Atrakcyjność migracyjna miast i obszarów wiejskich województwa wielkopolskiego. *Space – Society – Economy*, pp. 213–227.

Kałuża-Kopias, D., (2021). Ruch wędrówkowy ludności w województwie łódzkim po 2002 r., *Space – Society – Economy*, (32), pp. 61–81. https://doi.org/10.18778/1733-3180.32.03

Korcelli P., (1997). Alternatywne projekcje zmian demograficznych i migracji w aglomeracjach miejskich, [w:] Korcelli P,.(red.), Aglomeracje miejskie w procesie transformacji, *Zeszyty IGiPZ PAN 45*, Warszawa, pp. 5–22.

Lesińska, M., Matuszczyk, K., (2018). Działania samorządów wobec migracji w kontekście zmian demograficznych. Przykład trzech polskich województw. *Studia regionalne i lokalne*, No. 3(77), pp. 64–82

Privault N.,(2018). Understanding Markov chains: Examples and applications. *Springer.*

Rahimipour Sheikhani Nejad, M.A., Nasiri Jan Agha, F. and Khatami, S. S., (2018). Monitoring and predicting land cover changes in the coastal areas for optimal land allocation (Case study: Chaf and Chamkhaleh, Guilan, Iran), *International Journal of Development and Sustainability*, Vol. 7 No. 3, pp. 973–985.

Rakowska, B., Rakowski W., (2009). Województwo mazowieckie jako obszar napływu i odpływu ludności. *Rocznik Żyrardowski*, 7, pp. 343–369.

Rogers, A., (1968). Matrix Analysis of Interregional Population Growth and Distribution. *University of Kalifornia.*

Romeu, J., (2020). A Markov Chain Model for Covid-19 Survival Analysis. Syracuse University. https://www. researchgate. net/profile/Jorge–Romeu.

Rosner, A., (2014). Migracje wewnętrzne i ich związek z przestrzennym zróżnicowaniem rozwoju społeczno-gospodarczego wsi, *Wieś i Rolnictwo*, 1 (162), pp. 63–79.

Roszko-Wójtowicz, E., (2018). Migracje międzywojewódzkie w Polsce w latach 2010–2016 a jakość życia, *Ekspertyzy i Opracowania*, 75, pp. 1–13.

Sempewo, J., Kyokaali, L., (2016). Prediction of the Future Condition of a Water Distribution Network Using a Markov Based Approach: A Case Study of Kampala Water, *Procedia Engineering*, Vol.154, pp. 374–383.

Sojka, E., (2018). Odległość geograficzna i miernik rozwoju społeczno-gospodarczego a wielkość migracji w województwie śląskim, *Studia Ekonomiczne. Zeszyty Naukowe Uniwersytetu Ekonomicznego w Katowicach*, 353, pp. 73–88.

Struzik, A., (2007). Współczesne problemy rozwoju województwa mazowieckiego, *Rocznik Żyrardowski*, tom V, WSRL, Żyrardów.

Śleszyński, P., (2005). Różnice liczby ludności ujawnione w Narodowym Spisie Powszechnym 2002, *Przegląd Geograficzny*, 77(2), pp. 193–212.

Śleszyński, P., (2011). Oszacowanie rzeczywistej liczby ludności gmin województwa mazowieckiego z wykorzystaniem danych ZUS, *Studia demograficzne*, 2, pp. 35–58.

Śleszyński, P., Heffner, K., Solga, B., Wiśniewski, R. (2018). Perspektywa geografii i studiów regionalnych w badaniach nad migracjami. [W:] A. Horolets, M. Lesińska, M. Okólski (red.), Raport o stanie badań nad migracjami w Polsce po 1989 roku. *Komitet Badań nad Migracjami PAN*, pp. 174–210.

Usman, M. E., Faiz, M.B., (2015). Markov chains analysis and mechanism of migration in Indonesia in the period 1980-2010, *Journal of Engineering and Applied Sciences*, 10 (22). pp. 17256–17264.

Willekens, F., (1999). Probability models of migration: complete and incomplete data. *Journal of Demography* 7(1).

Venkatesan, G., Sasikala, V., (2018). Statistical analysis on migrants using Markov chain model. *The International Journal of Creative Research Thoughts,* Vol. 6(1).