Soft Randomized Machine Learning Procedure for Modeling Dynamic Interaction of Regional Systems

Yuri S. Popkov [1,2,3,4]

[1] Federal Research Center "Computer Science and Control" of Russian Academy of Sciences, 119333 Moscow, Russia; popkov@isa.ru
[2] Institute of Control Sciences of Russian Academy of Sciences, 117997 Moscow, Russia
[3] Department of Software Engineering, ORT Braude College, 216100 Karmiel, Israel
[4] Yugra Research Institute for Information Technologies, 628011 Khanty-Mansiysk, Russia

Abstract: The paper suggests a randomized model for dynamic migratory interaction of regional systems. The locally stationary states of migration flows in the basic and immigration systems are described by corresponding entropy operators. A soft randomization procedure that defines the optimal probability density functions of system parameters and measurement noises is developed. The advantages of soft randomization with approximate empirical data balance conditions are demonstrated, which considerably reduces algorithmic complexity and computational resources demand. An example of migratory interaction modeling and testing is given.

Keywords: soft randomization; entropy; entropy operator; migration; immigration; empirical balance; empirical risk

1. Introduction

The mutual influence of migratory processes in regional systems is a problem of growing significance in the modern world. The socioeconomic statuses of different regions demonstrate higher heterogeneity in response to rising political and military tension. All these factors cause an abrupt redistribution of migration flows and regional population variations, thereby increasing the cost of regional population maintenance [1–4]. Therefore, it is important to develop different tools (mathematical models, algorithms, and software) for forecasting the distribution of migration flows with adaptation to their dynamics considering available resources.

The authors of [5] suggested a dynamic entropy model for the migratory interaction of regional systems. In comparison with biological reproduction, migration mobility is a rather fast process [1,6]. Thus, the short-term dynamics of regional population size are described by the locally stationary state of a migratory process [7]. The latter can be simulated under the hypothesis that all migrants have a random and independent spatial distribution over interacting regional systems with given prior probabilities. The mathematical model of a locally stationary state is given by a corresponding entropy operator that maps the space of admissible resources into the space of migratory processes [8].

Mathematical modeling and analysis of interregional migration is considered in numerous publications. First, it seems appropriate to mention the monographs [9,10] that are dedicated to a wide range of interregional migration problems, including mathematical modeling of migration flows. Note that the problem of migration touches upon many aspects of socioeconomic, psychological and political status of the space of migratory movements. Thus, of crucial role is the structural analysis of inter- and intraregional migration flows [4] and motivations that generate them [2,11]. The results of structural and motivational analysis of migratory processes are used for computer simulation. There exist three directions of research in this field, each relying on some system of hypotheses. One of the

directions involves the stochastic hypothesis about the origin of migratory motivations [12], which is simulated using agent technologies [13,14]. This direction is adjoined by investigations based on the thermodynamic model of migration flows [3,8]. Of course, the short list above does not exhaust the whole variety of migration studies, merely outlining some topics of research.

This paper studies a stochastic version of the model in [5], in which random parameters and measurement noises are characterized by probability density functions (PDFs). These functions are estimated using retrospective information on the real dynamics of regional population size with "soft" randomized machine learning [15]. The learned model was implemented in the form of computer simulations, i.e., generation of an ensemble of random trajectories with the entropy-optimal PDFs of the model parameters and measurement noises. The resulting ensemble was used for testing of the model and also for short-term forecasting.

The method developed below is illustrated by an example of the randomized modeling and forecasting of the migratory interaction among three EU countries (Germany, France, and Italy—the system $\mathcal{GFI}$) and two countries as sources of immigration (Syria and Libya—the system $\mathcal{SL}$).

2. Randomized Model of Migratory Interaction

Consider the dynamic discrete-time model of migratory interaction with shared resource constraints that is presented in [5]. The first sub-model represents migration flows within the system $\mathcal{GFI}$ and is described by the dynamic regression equation

$$\mathbf{K}[(s+1)h] = (A-E)\mathbf{K}[sh] + \mathbf{F}(z[sh]), \qquad (\mathbf{K}, \mathbf{F}) \in R^N, \ s = \overline{0, K-1}, \tag{1}$$

where

$$A = h \begin{pmatrix} 1 & \alpha_2 a_{21} & \cdots & \alpha_N a_{N1} \\ \alpha_1 a_{12} & 1 & \cdots & \alpha_N a_{N2} \\ \cdots & \cdots & \cdots & \cdots \\ \alpha_1 a_{1N} & \alpha_2 a_{2N} & \cdots & 1 \end{pmatrix}, \tag{2}$$

$$E = h \operatorname{diag}[\alpha_n, \ n = \overline{1, N}]. \tag{3}$$

In these equations, $\mathbf{K}[sh]$ denotes the population distribution in the regional system $\mathcal{GFI}$ at a time sh.

At a time sh, the distribution of immigration flows from the regional system $\mathcal{SL}$ to the regional system $\mathcal{GFI}$ in terms of an entropy operator is modeled by the second sub-model, which can be described by a vector function $\mathbf{F}(z[sh])$ with the components

$$f_n[sh] = h \sum_{j=1}^{M} b_{jn}(z[sh])^{c_{jn}}, \qquad n = \overline{1, N}, \ s = \overline{0, K-1}, \tag{4}$$

The variable z, which is the exponential Lagrange multiplier in the entropy-optimal distribution problem of immigration flows, satisfies the equation

$$\sum_{k=1}^{M} \sum_{n=1}^{N} c_{kn} b_{kn}(z[sh])^{c_{kn}} = T[sh], \tag{5}$$

where $T[sh]$ is the amount of a shared resource used by all regions from the system $\mathcal{GFI}$ to maintain immigrants.

In this model, the input data are the amounts $T[0], T[h], \ldots, T[(K-1)h]$; and the output data are the regional population distributions $\mathbf{K}[0], \mathbf{K}[h], \ldots, \mathbf{K}[(K-1)h]$.

The dynamic model in Equations (1)–(5) contains the following parameters:

- $\alpha_n \in [0,1], \ n = \overline{1,N}$, as the shares of mobile population in system regions;
- $a_{in} \in [0,1], \ (i,n) = \overline{1,N}$, as the prior probabilities of individual migration in the system $\mathcal{GFI}$;

- b_{kn}, $k = \overline{1, M}$, $n = \overline{1, N}$, as the prior probabilities of individual immigration from region k of the system $\mathcal{SL}$ to region n of the system $\mathcal{GFI}$; and
- c_{kn}, $k = \overline{1, M}$, $n = \overline{1, N}$, as the normalized 1 specific generalized cost of immigration maintenance.

Normalization means that $0 < c_{kn} < 1$, $k = \overline{1, M}$, $n = \overline{1, N}$.

The parameters form three groups: mobility, migratory movements within the system $\mathcal{GFI}$, and immigratory movements from the system $\mathcal{SL}$ to the system $\mathcal{GFI}$. All these characteristics are specified by the regions of both systems. The dimensionality of the parametric space is reduced using the same approach as in [5]. The whole essence is to assign a relative regional differentiation of all parameters except for the weights b_1 (mobility) and b_2 (internal migration) of these groups, which are considered as model variables.

This approach leads to the parametric transformation

$$\alpha_n = b_1 m_n, \quad a_{in} = b_2 h_{in}, \tag{6}$$
$$(i, n) = \overline{1, N}; \; k = \overline{1, M},$$

where m_n and h_{in} are given parameters which characterize the relation of variables.

Then, the dynamic model of migratory interaction in Equations (1)–(5) takes the form

$$\mathbf{K}[(s+1)h] = (b_1 b_2 \tilde{A} - b_1 \tilde{E})\mathbf{K}[sh] + \tilde{\mathbf{F}}(z[sh]), \tag{7}$$

with the matrix

$$\tilde{A} = h \begin{pmatrix} 1 & m_2 h_{21} & \cdots & m_N h_{N1} \\ m_1 h_{12} & 1 & \cdots & m_N h_{N2} \\ \cdots & \cdots & \cdots & \cdots \\ m_1 h_{1N} & m_2 h_{2N} & \cdots & 1 \end{pmatrix} \tag{8}$$

and the diagonal matrix

$$\tilde{E} = h \, \mathrm{diag}[m_n, \; n = \overline{1, N}]. \tag{9}$$

The vector $\tilde{\mathbf{F}}(\mu, z)[sh]$ consists of the components

$$\tilde{f}_n(z[sh]) = h \sum_{k=1}^{M} q_{kn}(z[sh])^{c_{kn}}, \quad n = \overline{1, N}, \; s = \overline{0, K-1}. \tag{10}$$

For each time sh, the variable z satisfies the equation

$$\sum_{k=1}^{M} \sum_{n=1}^{N} c_{kn} q_{kn}(z[sh])^{c_{kn}} = T[sh], \qquad s = \overline{0, K-1}, \tag{11}$$

i.e., there exist K values $z = z^*[sh]$, $s = \overline{0, K-1}$.

The randomized version of this model is described by Equations (7)–(11) but some parameters (variables) have random character. These are two randomized parameters, b_1 and b_2, as well as the variable $z = b_3$, all of the interval type. More specifically, the parameters b_1 and b_2 belong to the intervals

$$\mathcal{B}_1 = [b_1^-, b_1^+], \; \mathcal{B}_2 = [b_2^-, b_2^+]. \tag{12}$$

The interval $\mathcal{B}_3$ of the variable b_3 is given by Equation (11).

Theorem 1. *Let the parameters b_{kn} and c_{kn} in Equation (11) be positive and $c_{kn} \in [0, 1]$. Then, the solution b_3^* of this equation belongs to the interval*

$$\mathcal{B}_3 = [b_3^-, b_3^+], \tag{13}$$

where

$$b_3^- = \left(\frac{T[sh]}{MN c_{max} b_{max}} \right)^{1/c_{max}} ; \quad b_3^+ = \left(\frac{T[sh]}{MN c_{min} b_{min}} \right)^{1/c_{min}} ; \tag{14}$$

$$c_{min} = \min_{kn} c_{kn}, \ c_{max} = \max_{kn} c_{kn}; \ b_{min} = \min_{kn} b_{kn}, b_{max} = \max_{kn} b_{kn}.$$

The proof is postponed to the Appendix A.

Therefore, the randomized dynamic model in Equations (7)–(11) includes three random parameters $\mathbf{b} = \{b_1, b_2, b_3\}$ of the interval type that are defined over the three-dimensional cube with faces (Equations (12) and (13)), i.e.,

$$\mathcal{B} = \bigotimes_{j=1}^{3} \mathcal{B}_j. \tag{15}$$

The probabilistic properties of the randomized parameters are described by a continuously differentiable PDF $W(\mathbf{b})$.

By assumption, real distributions of regional population sizes contain errors that are simulated by a random vector $\bar{\xi}[sh] \in R^N$ with the interval components

$$\bar{\xi}[sh] \in \Xi_s = \left[\bar{\xi}^-[sh], \bar{\xi}^+[sh] \right]. \tag{16}$$

The probabilistic properties of this vector are described by a continuously differentiable PDF $Q(\bar{\xi})$. The measured output of the randomized model has an additive noise,

$$\mathbf{v}[sh] = \mathbf{K}[sh] + \bar{\xi}[sh]. \tag{17}$$

3. Characterization of Empirical Risk and Measurement Noises

Construct a synthetic functional $J[W(\mathbf{b}), Q(\bar{\xi})]$ *that depends on the PDFs of the model parameters and measurement noises for assessing in quantitative terms the empirical risk (the difference between the regional population distribution generated by the model in Equations (7)–(11) and the real counterpart) and the guaranteed power of these noises.* The functional must have components characterizing an intrinsic uncertainty of randomized machine learning (RML) procedures, the approximation quality of empirical balances (the empirical risk) and the worst properties of the corresponding random interval-type noises.

1. Uncertainty. In accordance with the general concept of RML, the first component among the listed ones is *an entropy functional* that describes the level of uncertainty:

$$\mathcal{H}[\mathbf{b}), Q(\bar{\xi})] = - \int_{\mathcal{B}} W(\mathbf{b}) \ln W(\mathbf{b}) d\mathbf{b} - \int_{\Xi} Q(\bar{\xi}) \ln Q(\bar{\xi}) d\bar{\xi}. \tag{18}$$

The two other functional components are constructed using Hölder's vector and matrix norms (The vector norm has the form $\|\mathbf{a}\|_\infty = \max_n |a_n|$; the matrix norm, the form $\|A\|_\infty = \max_{ij} |a_{ij}|$.) [16].

2. Approximate empirical balances. First, consider a characterization of *the empirical risk*. For the model in Equations (7)–(11), the deviation between the output and real data vectors is given by

$$\bar{\varepsilon}[sh] = \left(b_1 b_2 \tilde{A} - b_1 \tilde{E} \right) \mathbf{Y}[sh] + \mathbf{F}(b_3[sh]) - \mathbf{Y}[sh], \quad s = \overline{0, K-1}. \tag{19}$$

Using well-known inequalities for the matrix and vector norms, it is possible to write

$$\|\bar{\varepsilon}[sh]\|_\infty \leq \| \left(b_1 b_2 \tilde{A} - b_1 \tilde{E} \right) \|_\infty \|\mathbf{Y}[sh]\|_\infty + \|\mathbf{F}(b_3[sh])\|_\infty + \|\mathbf{Y}[(s+1)h]\|_\infty =$$
$$= \varphi(b_1, b_2, b_3, s), \quad s = \overline{0, K-1}. \tag{20}$$

Introducing the average matrix and vector norms over the observation interval,

$$\varphi(b_1, b_2, b_3) \leq h \left(\frac{1}{K} \sum_{s=0}^{K-1} \max_n y_n[sh] \right) \left(b_1 \max_n m_n + b_1 b_2 \max_{i,j} h_{ij} \right) +$$

$$+ \frac{1}{K} \sum_{s=0}^{K-1} \max_n y_n[(s+1)h] + MN c_{max} b_{max} (b_3)^{c_{max}}. \tag{21}$$

The parameters b_1 and b_2 take values within the intervals $\mathcal{B}_1$ and $\mathcal{B}_2$ (Equation (12)) while the parameter b_3 within the interval

$$\mathcal{B}_3 = \left[\left(\frac{T_{max}}{MN c_{max} q_{max}} \right)^{1/c_{max}}, \left(\frac{T_{max}}{MN c_{min} q_{min}} \right)^{1/c_{min}} \right], \tag{22}$$

where

$$T_{max} = \max_s T[sh]. \tag{23}$$

Denote

$$U_1 = h \left(\frac{1}{K} \sum_{s=0}^{K-1} \max_n y_n[sh] \right) \max_n m_n; \ U_2 = h \left(\frac{1}{K} \sum_{s=0}^{K-1} \max_n y_n[sh] \right) \max_{i,j} h_{ij};$$

$$U_3 = MN h c_{max} b_{max}; \ U_4 = \frac{1}{K} \sum_{s=0}^{K-1} \max_n y_n[(s+1)h]. \tag{24}$$

Then, the function $\varphi(b_1, b_2, b_3)$ takes the form

$$\varphi(b_1, b_2, b_3) = b_1 U_1 + b_1 b_2 U_2 + (b_3)^{c_{max}} U_3 + U_4. \tag{25}$$

Note that the coefficients $U_1, \ldots, U_4$ are determined by real data on regional population distributions and also by the characteristics of internal migration within the system $\mathcal{GFI}$ and immigration flows from the system $\mathcal{SL}$.

The equality in Equation (25) defines a function $\varphi(b_1, b_2, b_3)$ of random variables. Let its expectation be the characteristic of the empirical risk, i.e.,

$$r[W(\mathbf{b})] = \int_{\mathcal{B}} W(\mathbf{b}) \varphi(\mathbf{b}) d\mathbf{b}, \tag{26}$$

where $\mathcal{B} = \mathcal{B}_1 \otimes \mathcal{B}_2 \otimes \mathcal{B}_3$ and the intervals $\mathcal{B}_1$ and $\mathcal{B}_2$ have given limits. At the same time, the limits of the interval $\mathcal{B}_3$ are specified by the equalities in Equation (22).

Power of noises. The measurement noises are simulated by random vectors $\bar{\xi}[sh] \in R^N$, $s = \overline{0, K-1}$. The components of these vectors may have different domains (ranges of values) at different times $s = \overline{0, K-1}$. For each time, introduce the Euclidean norm $\|\bar{\xi}[sh]\|_N^2$ and its expectation

$$n_s[Q(\bar{\xi}[sh])] = \int_{\Xi} Q(\bar{\xi}[sh]) \|\bar{\xi}[sh]\|_N^2 d\bar{\xi}[sh]. \tag{27}$$

The average expectation of this norm over the time interval has the form

$$\bar{n}_s[Q(\bar{\xi}[sh])] = \frac{1}{K} \sum_{s=0}^{K-1} n_s[Q(\bar{\xi}[sh])]. \tag{28}$$

If the measurement noises are the same on the observation interval, then the noise power functional can be written as

$$\bar{n}_s[Q(\bar{\xi}[sh])] = n[Q(\bar{\xi})] = \int_\Xi Q(\bar{\xi})\|\bar{\xi}\|_N^2 d\bar{\xi}. \tag{29}$$

This formula involves the Euclidean norm for a quantitative characterization of the noise power. However, it is possible to choose other norms depending on problem specifics.

4. Soft Randomized Estimation of Model Parameters

The model characteristics and measurement noises are estimated using a learning data collection: the real cost of immigrants maintenance $T[0], \ldots, T[(K-1)h]$ (input data) and the real distributions of regional population sizes $\mathbf{Y}[0], \ldots, \mathbf{Y}[(K-1)h]$ (output data).

In accordance with the general procedure of soft randomized machine learning [15], the optimal probability density functions $W(\mathbf{b})$ (model parameters) and $Q(\bar{\xi})$ (measurement noises) are calculated by the constrained minimization of the synthetic functional $J[W(\mathbf{b}), Q(\bar{\xi})]$ that contains the following functionals:

- the entropy

$$\mathcal{H}[W(\mathbf{b})] = -\int_\mathcal{B} W(\mathbf{b}) \ln W(\mathbf{b}) d\mathbf{b} - \int_\Xi Q(\bar{\xi}) \ln Q(\bar{\xi}) d\bar{\xi}; \tag{30}$$

- the average empirical risk over the observation interval

$$r[W(\mathbf{b})] = \int_\mathcal{B} W(\mathbf{b})(b_1 U_1 + b_1 b_2 U_2 + (b_3)^{c_{max}} U_3 + U_4) d\mathbf{b}; \tag{31}$$

and

- the average error norm

$$n[Q(\bar{\xi})] = \int_\Xi Q(\bar{\xi}) \sum_{i=1}^N \xi_i^2 d\bar{\xi}. \tag{32}$$

The soft randomized learning algorithm has the form

$$J[W(\mathbf{b}), Q(\bar{\xi})] = \mathcal{H}[W(\mathbf{b})] - r[W(\mathbf{b})] - n[Q(\bar{\xi})] \Rightarrow \max, \tag{33}$$
$$\int_\mathcal{B} W(\mathbf{b}) d\mathbf{b} = 1, \qquad \int_\Xi Q(\bar{\xi}) d\bar{\xi} = 1.$$

The solution of this problem is the optimal PDFs under maximal uncertainty, for the model parameters of the form

$$W^*(\mathbf{b}) = \frac{\exp\left(b_1 U_1 - b_1 b_2 U_2 - (b_3)^{c_{max}} U_3 - U_4\right)}{\mathcal{P}}, \tag{34}$$

where

$$\mathcal{P} = \int_\mathcal{B} \exp\left(b_1 U_1 - b_1 b_2 U_2 - (b_3)^{c_{max}} U_3 - U_4\right) d\mathbf{b}, \tag{35}$$

and for the measurement noises of the form

$$Q^*(\bar{\xi}) = \frac{\exp\left(-\sum_{i=1}^N \xi_i^2\right)}{\mathcal{Q}}, \tag{36}$$

where

$$\mathcal{Q} = \int_\Xi \exp\left(-\sum_{i=1}^N \xi_i^2\right) d\bar{\xi}. \tag{37}$$

In the case of soft randomization, there is no need for solving the empirical balance equations, which have high complexity and computational intensiveness due to the presence of integral components. Here, computational resources are required for the normalization procedure of the

resulting PDFs. On the other hand, the morphology of the optimal PDFs depends on a specific choice of the approximate data balancing criterion and a numerical characterization of the measurement noises.

5. Randomized Forecasting of Dynamic Migratory Interaction

Consider randomized forecasting of dynamic migratory interaction using the principle of soft randomization. Let $\mathcal{T}_{pr} = [s_0 h, s_{pr} h]$ be the forecasting interval and assume the initial state (the regional population distribution at the initial time $s_0 h$) coincides with the real distribution, i.e., $\mathbf{K}[s_0 h] = \mathbf{Y}[s_0 h]$. The shared cost of the system $\mathcal{GFI}$ to maintain immigrants is distributed in accordance with a given scenario. For each scenario, the value T_{max} and also the interval $\mathcal{B}_3$ in Equations (12), (22), and (23) are determined.

The forecasted trajectories are constructed using the randomized model in Equations (7), (10), and (11)

$$
\begin{aligned}
\mathbf{K}[(s+1)h] &= (b_1 b_2 \tilde{A} - b_1 \tilde{E})\,\mathbf{K}[sh] + \mathbf{F}[sh \mid b_3], \\
\mathbf{F}[sh \mid b_3] &= \{\sum_{k=1}^{M} b_{kn}(b_3)^{c_{kn}}, \quad n = \overline{1, N}\}, \\
s &= \overline{s_0, s_{pr}}, \quad \mathbf{K}[s_0 h] = Y[s_0 h].
\end{aligned}
\tag{38}
$$

The randomized parameters b_1, b_2, and b_3 take values within the corresponding intervals with the probability density function $W^*(\mathbf{b})$ (Equation (34)).

An ensemble of the forecasted trajectories for the model's output is obtained taking into account a random vector $\bar{\xi} \in \Xi$ with the PDF $Q^*(\bar{\xi})$ (Equation (36)):

$$
\mathbf{v}[sh] = \mathbf{K}[sh] + \bar{\xi}, \quad s = \overline{s_0, s_{pr}}.
\tag{39}
$$

For each scenario $T[s_0 h], \dots, T[s_{pr} h]$, an ensemble $\mathcal{K}$ of random forecasting trajectories is generated via sampling (the transformation of a PDF into a corresponding sequence of random vectors of length I) of the optimal PDFs of the model parameters and measurement noises for each time sh. The resulting ensemble allows deriving empirical estimates of different numerical characteristics as follows:

- the average trajectory

$$
\bar{\mathbf{K}}[sh] = \frac{1}{I} \sum_{i=1}^{I} \mathbf{K}^{(i)}[sh], \qquad s = \overline{s_0, s_{pr}};
\tag{40}
$$

- the variance trajectory

$$
\bar{\sigma}^2[sh] = \frac{1}{I-1} \sum_{i=1}^{I} \|\mathbf{K}^{(i)}[sh] - \bar{\mathbf{K}}[sh]\|^2, \qquad s = \overline{s_0, s_{pr}};
\tag{41}
$$

- the variance pipe, i.e., the set of random trajectories that almost surely (since an ensemble consists of a finite number of trajectories, the matter concerns not probability but its empirical estimate) belong to the domain

$$
\mathcal{D} = \{\mathbf{K}[sh] : \bar{\mathbf{K}}[sh] - \bar{\sigma}^2[sh] \le \mathbf{K}[sh] \le \bar{\mathbf{K}}[sh] + \bar{\sigma}^2[sh], \quad s = \overline{s_0, s_{pr}}\};
\tag{42}
$$

- the empirical probability distribution and its dynamics on the forecasting interval

$$
\mathbb{P}\left(\mathbf{K}[sh] \le \Delta, \; s = \overline{s_0, s_{pr}}\right) = \frac{I_\Delta}{I},
\tag{43}
$$

where I_Δ denotes the number of vectors $\mathbf{K}[sh]$ whose components are smaller than Δ; and

- the median trajectory $\hat{\mathbf{K}}[sh], s = \overline{s_0, s_{pr}}$, which satisfies the equation

$$\mathbb{P}(\mathbf{K}[sh]) = 0,5; \;\; s = \overline{s_0, s_{pr}}. \tag{44}$$

The ensemble $\mathcal{K}$ can be used to calculate other characteristics, e.g., α-quantiles, confidence probabilities, etc.

6. Example

The appearance of territories with low economic status always causes the growth of immigration. The early 2000s were remarkable for the formation of several such territories in Northern and Central Africa, the Near East, Afghanistan, etc. As a result, tens of millions of migrants moved to the EU as the level of life in these territories dropped below the subsistence minimum. The EU countries have to allocate considerable financial resources for their filtering and accommodation, which are often unacceptable. An example below illustrates the use of soft randomization for estimating and forecasting of immigration flows from Syria (1) and Libya (2) (the system $\mathcal{SL}$) to Germany (1), France (2), and Italy (3) (the system $\mathcal{GFI}$).

1. Randomized model, parameters, measurement errors, time intervals, and real data collections. Choose the randomized mathematical model (Equation (25)) with the normalized variables

$$p_n[sh] = \frac{K_n[sh]}{K_{max}}, \qquad n = \overline{1,3}. \tag{45}$$

This gives

$$
\begin{aligned}
p_n[(s+1)h] &= (1 - b_1 m_n)p_n[sh] + hb_1 b_2 \sum_{i=1, i \neq n}^{3} m_i h_{in} p_i[sh] + h f_n[sh], \\
f_n[sh] &= \sum_{i=1}^{M} b_{in} b_3^{c_{in}}, \qquad n = \overline{1,3}, \\
T[sh] &= \sum_{n=1}^{3} \sum_{i=1}^{2} c_{in} b_{in} b_3^{c_{in}}.
\end{aligned} \tag{46}
$$

The state variables of the system $\mathcal{GFI}$ and also the immigration flows from the system $\mathcal{SL}$ are normalized, i.e.,

$$0 \leq p_n[sh] \leq 1, \quad 0 \leq f_n[sh] \leq 1, \qquad n = \overline{1,N}. \tag{47}$$

The variable z^* characterizes the entropy operator of the immigration process and satisfies the last equation in Equation (46). The values of the parameters $m_i, h_{in}, b_{in},$ and c_{in} are combined in Table 1, where columns are different values of corresponding parameter. Recall that the two lowest rows of Table 1 indicate the values of the parameters c_{in}. By assumption, the regions of both systems have the same specific cost.

Table 1. Values of relative parameters.

m_n	0.43	0.50	0.40
h_{1n}	0	0.3	0.3
h_{2n}	0.3	0	0.3
h_{3n}	0.5	0.4	0
b_{1n}	0.4	0.3	0.3
b_{2n}	0.3	0.1	0.4
c_{1n}	0.4	0.4	0.3
c_{2n}	0.4	0.4	0.3

In accordance with this table, $m_{max} = 0.5, h_{max} = 0.5, b_{min} = 0.3, b_{max} = 0.4$, and $c_{max} = c_{min} = c = 0.5$. The measurement errors of population sizes $\bar{\xi}[sh] \in R^3$ (in normalized units) belong to the intervals

$$\bar{\xi}[sh] \in \Xi = [\bar{\xi}_-, \bar{\xi}_+], \qquad \xi_n^{\pm} = 0.01, \tag{48}$$

and by assumption they have the same limits for times sh.

The normalized observation (model output) has the form

$$\mathbf{v}[sh] = \mathbf{p}[sh] + \bar{\xi}[sh]. \tag{49}$$

The random parameter model in Equation (46) was employed for estimating parameter characteristics and testing on corresponding time intervals with step $h = 1 year$:

- $\mathrm{T}_{est} = 2009$–2013 as the estimation interval; and
- $\mathrm{T}_{tst} = 2014$–2018 as the testing interval.

2. Entropy estimation of PDFs of model parameters and measurement noises (interval T_{est}). This problem was solved using available data on regional population distribution for Germany ($n = 1$), France ($n = 2$), and Italy ($n = 3$) and also on the shared cost of immigrants maintenance on the estimation interval (see Table 2 and UNdata service at https://data.un.org/).

Table 2. Input and output data collections.

Year	2009	2010	2011	2012	2013
s	0	1	2	3	4
$Y_1[s]$	81.90	81.77	80.27	80.42	80.64
$y_1[s]$	1.00	0.998	0.980	0.982	0.985
$Y_2[s]$	62.47	62.80	63.11	63.41	63.70
$y_2[s]$	0.762	0.767	0.771	0.774	0.778
$Y_3[s]$	59.39	59.53	59.63	59.71	59.75
$y_3[s]$	0.725	0.727	0.728	0.729	0.726
$T[s]$ (*billion*)	0.093	0.094	0.095	0.096	0.097

In this model, the random parameters b_1, b_2, and b_3 take values within the intervals

$$b_1 \in \mathcal{B}_1 = [1.0, 2.5]; \quad b_2 \in \mathcal{B}_2 = [0.5, 1.8], \quad b_3 \in \mathcal{B}_3 = [0.3, 1.5]. \tag{50}$$

In accordance with Equation (24),

$$U_1 = 0.5; \; U_2 = 0.5; \; U_3 = 1.2; \; U_4 = 0.986. \tag{51}$$

Then, the soft RML procedure yields the following optimal PDFs of the model parameters and measurement noises:

$$
\begin{aligned}
W^*(\mathbf{b}) &= \frac{\exp\left(-0.5b_1 - 0.5b_1 b_2 - 1.2b_3^{0.5} - 0.986\right)}{\mathcal{W}}, \\
Q^*(\bar{\xi}) &= \frac{\exp\left(-\sum_{n=1}^{3} \xi_n^2\right)}{\mathcal{Q}},
\end{aligned}
\tag{52}
$$

where

$$
\begin{aligned}
\mathcal{W} &= \int_{\mathcal{B}_1} \int_{\mathcal{B}_2} \int_{\mathcal{B}_3} \exp\left(-0.5b_1 - 0.5b_1 b_2 - 1.2b_3^{0.5} - 0.986\right) db_1 db_2 db_3, \\
\mathcal{Q} &= \prod_{n=1}^{3} \int_{-0.01}^{0.01} \exp(-\xi^2) d\xi.
\end{aligned}
\tag{53}
$$

The two-dimensional sections of the three-dimensional PDFs of the model parameters are shown in Figure 1a–c, while the graphs of the PDFs of the measurementnoises in Figure 2.
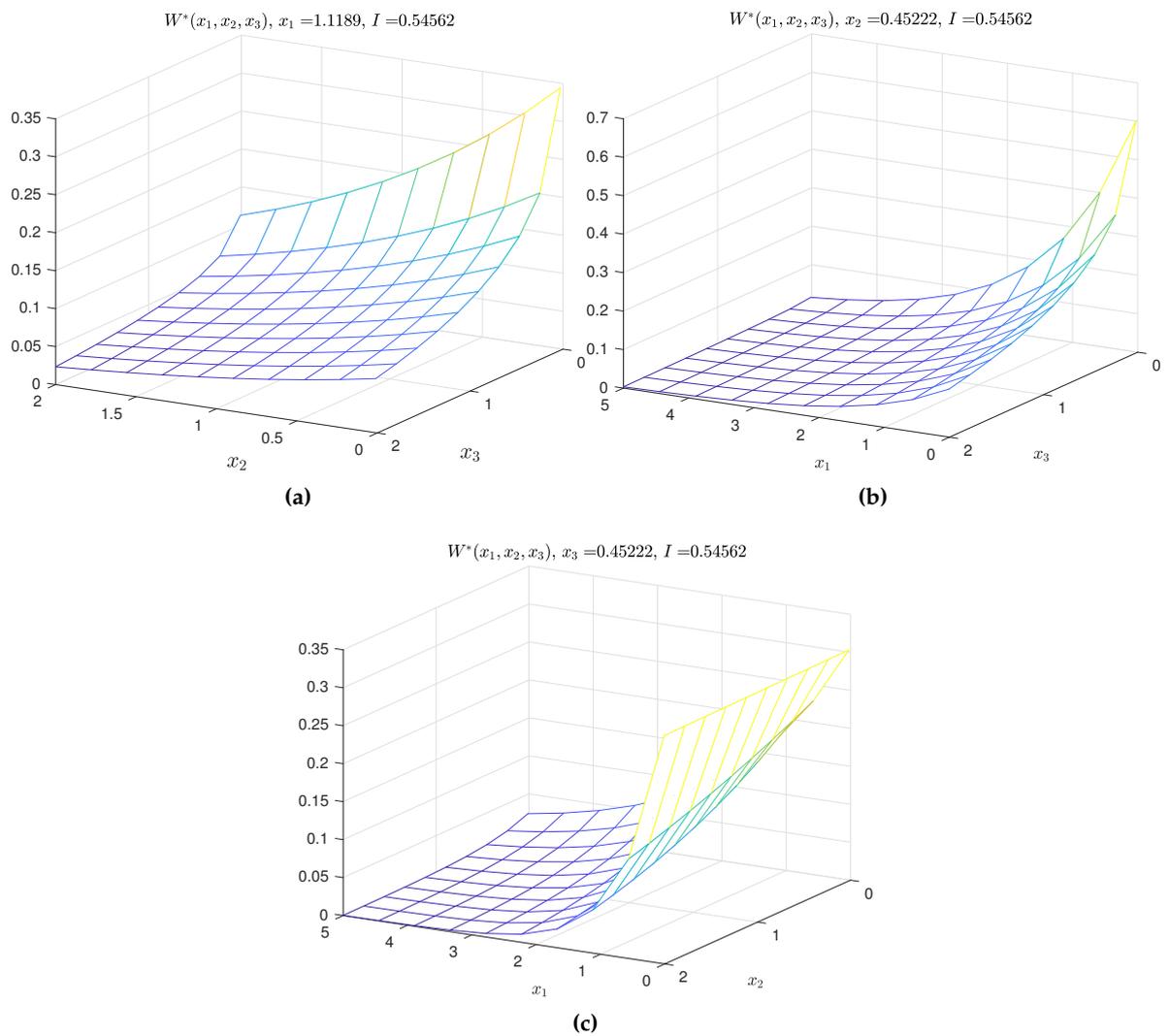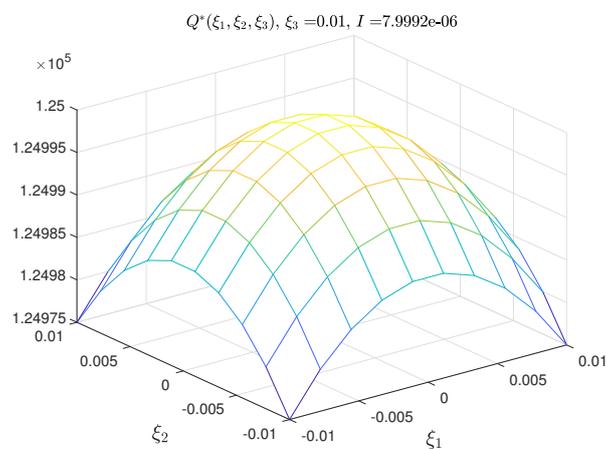

(a)

(b)

(c)

Figure 1. 2-dimensional section of W.

Figure 2. 2-dimensional section of Q.

3. Model testing. The randomized model in Equation (49) with the optimal PDFs in Equations (52) and (53) was tested using the above data on regional population sizes from the UNdata service

(see Table 3). This table also presents the testing results in terms of the ensemble-average trajectories $\bar{p}_1[sh]$, $\bar{p}_2[sh]$, and $\bar{p}_3[sh]$.

Table 3. Input and output data collections.

Year	2014	2015	2016	2017	2018
s	0	1	2	3	4
$Y_1[s]$	81.489	81.707	82.063	82.386	82.674
$y_1[s]$	0.985	0.988	0.993	0.996	1.000
$\bar{p}_1[sh]$	0.986	0.615	0.743	0.639	0.999
$Y_2[s]$	64.190	64.457	64.791	65.134	65.484
$y_2[s]$	0.721	0.472	0.564	0.529	0.708
$\bar{p}_2[sh]$	0.722	0.695	0.707	0.691	0.715
$Y_3[s]$	59.585	59.504	59.504	59.509	59.516
$y_3[s]$	0.775	0.609	0.562	0.699	0.650
$\bar{p}_3[sh]$	0.776	0.617	0.607	0.705	0.628
$T[s]$ (*billion*)	0.097	0.097	0.097	0.098	0.098

Testing was performed via sampling of the randomized interval parameters with the PDFs in Equations (52) and (53) and construction of the corresponding trajectories by Equation (49). Figure 3a–c shows ensembles of such trajectories $v_1[sh]$, $v_2[sh]$, $v[sh]$ as well as the ensemble-average trajectories $\bar{v}_1[sh]$, $\bar{v}_2[sh]$, $\bar{v}_3[sh]$ (Graph 1); the real trajectories $y_1[sh]$, $y_2[sh]$, $y_3[sh]$ of regional population sizes (Graph 2); and the limits of the variance pipes $\bar{p}_1^*[sh] \pm \sigma_1$, $\bar{p}_2^*[sh] \pm \sigma_2$, $\bar{p}_3^*[sh] \pm \sigma_3$ (Graph 3).
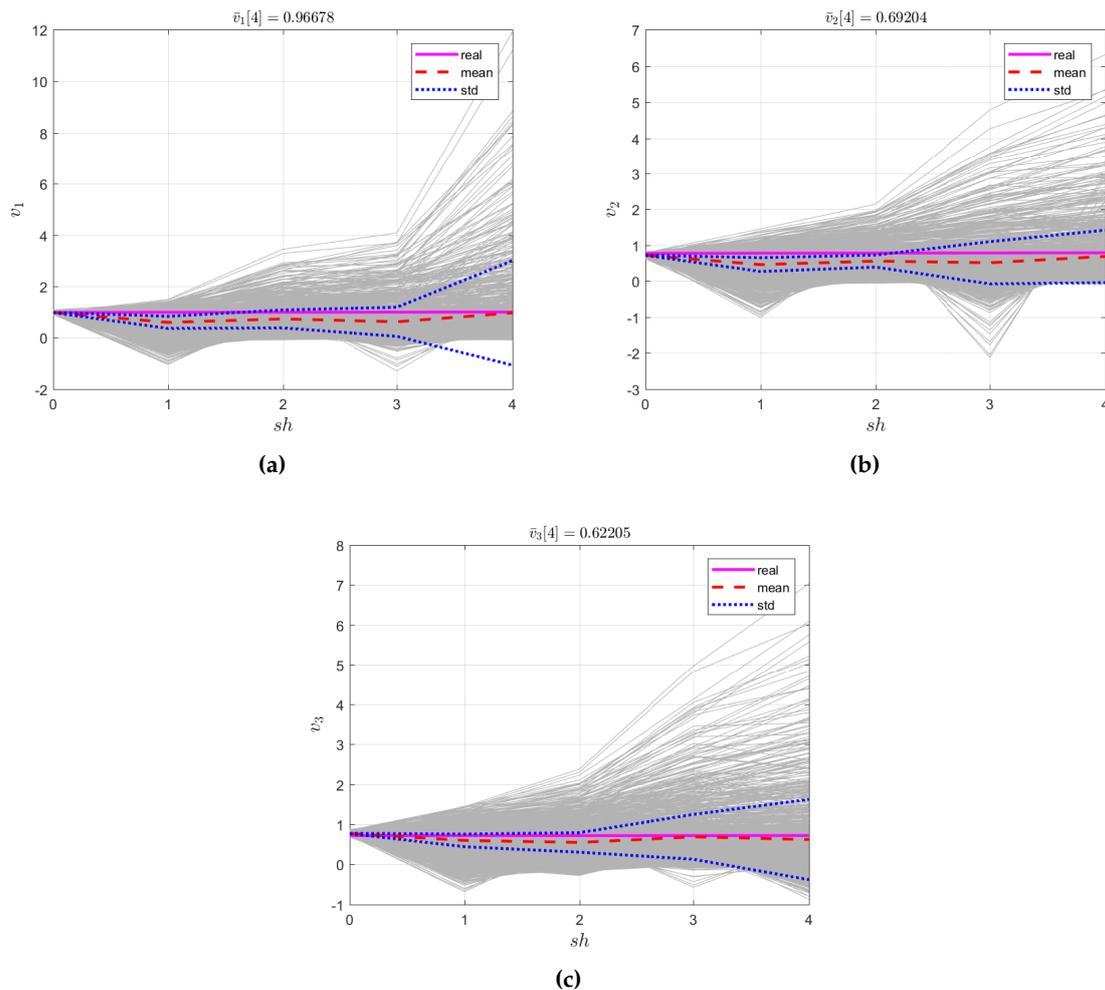

(a)

(b)

(c)

Figure 3. (**a**) $\bar{v}_1[4]$, (**b**) $\bar{v}_2[4]$, (**c**) $\bar{v}_3[4]$.

The testing accuracy was estimated in terms of the relative root-mean-square error

$$\delta_n = \frac{\sqrt{\sum_{s=0}^{4}(\bar{p}_n[sh] - y_n[sh])^2}}{\sqrt{\sum_{s=0}^{4}(\bar{p}_n[sh])^2} + \sqrt{\sum_{s=0}^{4}(y_n[sh])^2}}. \tag{54}$$

In the example under study, it constituted 4.6% (Region 1), 3.5% (Region 2), and 2.6% (Region 3).

7. Conclusions

This paper has developed a mathematical model for dynamic migratory interaction of regional systems with locally stationary states described by corresponding entropy operators. The model incorporates random parameters, and their probabilistic characteristics—the probability density functions of system parameters and measurement noises—have been calculated using soft randomized machine learning. An example of migratory interaction modeling and testing has been given.

Conflicts of Interest: The author declares no conflict of interest.

Appendix A

Proof of Theorem 1. Consider the function

$$\varphi(z) = \mu \sum_{k=1}^{M} \sum_{n=1}^{N} c_{kn} q_{kn}(z[sh])^{c_{kn}}, \tag{A1}$$

which appears on the left-hand side of Equation (19). Taking advantage of the obvious inequalities,

$$\varphi_-(z) = MN\mu c_{min} q_{min}(z)^{c_{min}} < \varphi(z) < MN\mu c_{max} q_{max}(z)^{c_{max}} = \varphi_+(z). \tag{A2}$$

The variables are $0 < c_{min} < 1$, $0 < c_{max} < 1$, $c_{min} < c_{max}$, and $c_{min} < c_{kn} < c_{max}$. Consider the equations

$$\varphi_-(z) = T[sh], \quad \varphi(z) = T[sh], \quad \varphi_+(z) = T[sh]. \tag{A3}$$

The functions $\varphi_-(z)$, $\varphi(z)$, and $\varphi_+(z)$ are strictly convex. Therefore, the solutions of these equations has the relationship

$$z_- < z^* < z_+, \tag{A4}$$

which concludes the proof of Theorem 1. $\square$

References

1. Bilecen, B.; Van Mol, C. Introduction: International academic mobility and inequalities. *J. Ethic Migr. Stud.* **2017**, *43*, 1241–1255. [CrossRef]
2. Black, R.; Xiang, B.; Caller, M.; Engberson, G.; Heering, L.; Markova, E. Migration and development: Causes and consequences. In *The Dynamics of International Migration and Settlement in Europe: A State of the Art*; Penninx, R., Berger, M., Kraal, K., Eds.; Amsterdam University Press: Amsterdam, The Netherlands, 2006; pp. 41–63.
3. Imel'baev, S.S.; Shmul'yan, B.L. Modeling of stochastic communication systems. In *Entropy Methods for Complex Systems Modeling*; Wilson, A.G., Ed.; Nauka: Moscow, Russian, 1975; pp. 170–234.
4. Van der Knaap, G.A.; Steegers, W.F. Structural analysis of interregional and intraregional migration patterns. In *Demographic Research and Spatial Policy*; Heide, H., Willekens, F., Eds.; Academic Press: Voorburg, The Netherlands, 1984.
5. Popkov, Y.S. Dynamic entropy model for migratory interaction of regional systems. *Tr. Inst. Sist. Analiz. Ross. Akad. Nauk* **2018**, *2*, 3–11.
6. Zelinsky, W. The hypothesis of the mobility transition. *Geogr. Rev.* **1971**, *46*, 219–249. [CrossRef]

7. Popkov, Y.S. *Mathematical Demoeconomy: Integrating Demographic and Economic Approaches*; De Gruyter: Berlin, Germany, 2014.

8. Wilson, A.G. *Entropy in Urban and Regional Modelling*; Routledge: London, UK, 1970.

9. Rogers, A.; Willekens, F.; Raymer, J. Modelling interregional migration flows: continuity and change. *J. Math. Popul. Stud.* **2001**, *9*, 231–263. [CrossRef]

10. 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.

11. Volpert, V.; Petrovskii, S.; Zincenko, A. Interaction of human migration and wealth distribution. *Nonlinear Anal.* **2017**, *150*, 408–423. [CrossRef]

12. Pan, J.; Nagurney, A. Using Markov chains to model human migration in a network equilibrium framework. *Math. Comput. Model.* **1994**, *19*, 31–39. [CrossRef]

13. Klabunde, A.; Willekens, F. Decision-making in agent-based models of migration: State of the art and challenges. *Eur. J. Popul.* **2016**, *32*, 73–97. [CrossRef] [PubMed]

14. Klabunde, A.; Zinn, S.; Willekens, F.; Leuchter, M. Multistable modelling extended by behavioural rules. An application to migration. *Popul. Stud.* **2017**, *71*, 61–67. [CrossRef] [PubMed]

15. Popkov, Y.S. Soft randomized machine learning. *Doklady Math.* **2018**, *98*, 646–647. [CrossRef]

16. Voevodin, V.V.; Kuznetsov, Y.A. *Matrices and Calculations*; Nauka: Moscow, Russian, 1984. (In Russian)