

Constraints as factors reducing the entropy of distributions: an entropy-maximizing spatial interaction model as an example

*Ograniczenia jako czynniki redukujące entropię rozkładów:
przykład modelu maksymalizacji entropii przestrzennej interakcji*

JÁN PAULOV, ANTON BEZÁK

Comenius University, Faculty of Natural Sciences, Department of Regional Geography,
Protection and Planning of the Landscape
842 15 Bratislava 4, Mlynská dolina, Ilkovičova 6, Slovakia; anton.bezak@uniba.sk

Abstract. The aim of this paper is to provide empirical evidence for the statement that the constraints imposed on an objective function are able to reduce the entropy of the corresponding distributions produced by entropy-maximizing models. This idea is evaluated via an application to an entropy-maximizing spatial interaction model, as a typical representative of the family of entropy-maximizing models used in geography. Eleven versions of this spatial interaction model are fitted separately to six sets of data concerning interregional migration in Slovakia. For each model, the predicted flow distribution is derived, prior to calculation of the corresponding predicted entropy, and then comparison of the entropy values relating to all the models. The results obtained indicate very clearly that constraints imposed on an objective function reduce the initial maximum entropy successively, with this reduction depending on the number and nature of the constraints incorporated.

Key words: entropy of distribution, constraints, entropy-maximizing spatial interaction model, interregional migration flows, Slovakia.

Introduction

As is well-known, constraints used in the derivation of entropy-maximizing models are considered restrictions imposed on an objective function. However, Jaynes (1957, 1979), the author of the maximum entropy principle, simultaneously considers them information available. In that capacity, they are able to reduce the entropy of distributions produced by entropy-maximizing models. It

is this ability that is the main concern in this paper, and specifically empirical evidence in support thereof.

Conceptual background

Since the topic of this paper will “turn around” three closely-related basic concepts, viz. information, entropy and uncertainty, we will specify these briefly for the sake of convenience, notwithstanding their being rather well-known in principle.

If an amount of information contained in an event is $h = -\log p$, where p is the probability that the event would happen, then a mean amount of information (mathematical expectation) contained in a set of mutually exclusive and exhaustive events having the discrete probability distribution $P = \{p_1, p_2, \dots, p_n\}$ will be

$$H = -\sum_i p_i \log p_i. \quad [1]$$

Note, however, that relation [1] represents, not merely a mean amount of information, but also Shannon’s entropy relating to the probability distribution, which is in fact known as a measure of uncertainty (Wilson, 1970, p. 7; Yaglom and Yaglom, 1983, pp. 46-47).

The nature of relation [1] as a measure of uncertainty makes possible the derivation of a further concept, viz. information gain. Note that information gain is removed uncertainty, with the amount of information gained equal to the amount of uncertainty removed, i.e.

$$H_r = I_g = H_i - H_o, \quad [2]$$

where H_r denotes the amount of removed uncertainty, i.e. the information gain I_g , H_i denotes the initial uncertainty and H_o is the observed (actual or remaining) uncertainty.

It can be shown that maximum uncertainty, as measured in terms of maximum entropy, H_{\max} , appears when $p_i = 1/n = \text{const}$, where n denotes a number of mutually exclusive and exhaustive events (observations); in this case $H_{\max} = \log n$. If $H_i = H_{\max}$, then relation [2] assumes the form

$$H_r = I_g = H_{\max} - H_o = \log n + \sum_i p_i \log p_i = \sum_i p_i \log(np_i). \quad [3]$$

However, relation [3] represents a special case of the more-general Kullback information gain, $I_{g^{(k)}}$, defined as

$$I_{g^{(k)}} = \sum_i p_i \log(p_i/q_i), \quad [4]$$

where q_i denotes a *priori* probability and p_i a *posteriori* probability. If $q_i = 1/n = \text{const}$, then [4] converts to [3].

Application to entropy-maximizing models

Recall again that constraints, as items of information available, are able to reduce the uncertainty, i.e. the entropy as its measure, of the distributions produced by entropy-maximizing models. To support this view, let us quote the statement of Garrett, a physicist particularly concerned with the topic under discussion, who states: “The more constraints one has the smaller is the maximized entropy ...” (Garrett, 1991, p. 155), which in other words means the more uncertainty is removed. Similar formulations are also not unknown in the domains of geography or regional science. For example, a view identical to that from Garrett was expressed some years ago by Batty (2010). To be able to compare let us quote his statement: “This order-disorder continuum with respect to H is directly invoked if we consider that as we put more and more constraints on the form of a distribution we successively reduce the entropy” (Batty, 2010, pp. 397-398). An interesting assertion in line with the aim of our paper can also be found in the article of Roy and Thill (2004). These authors argue that, if maximization of entropy “is ... constrained by inducing the model flows to conform to certain aggregate base period quantities, [then] such constraints ... not only reduce the ‘entropy’ or ‘uncertainty’ of the final solution, but usually contribute to obtaining a ... solution closer to that observed at the base period” (Roy and Thill, 2004, p. 345).

The basic objective of this paper is to provide empirical evidence for the statement that the gradual imposition of constraints on an objective function, i.e. the addition of ever-more new constraints, really does lead to a gradual reduction in the entropy of a corresponding distribution. Note here also that the application of this procedure also allows for the ultimate derivation of a model producing such a distribution as is near to (in the ideal case as near as possible to) the one observed.

A spatial interaction model as an example

The objective stated above has been pursued by reference to an entropy-maximizing spatial interaction model, as a typical representative of the family of entropy-maximizing models used in geography. Recall that this kind of model was designed to predict or replicate the size of the interaction between a set of areal units. In the derivation of the spatial interaction model either Shannon’s entropy or the equivalent Boltzmann’s entropy is taken as an objective function that is maximized subject to certain constraints. We stress repeatedly that these constraints play a fundamental role in the entropy-maximization procedure. They represent whatever convenient information is available to the spatial-interaction modeller. The more information that is available, the greater will be the number and complexity of constraints. Consequently, the optimal flow

distribution in a particular case is strongly depended on the *number* and *nature* of the constraints incorporated.

As is well known, different versions of the spatial interaction model have been derived with respect to different constraints. The set of model versions selected in this paper includes, not only those considered standard and put to widespread use in interaction modelling, but also certain very simple versions used only rarely, if at all. Nevertheless, reduction of the entropy of corresponding distributions also arises in the case of these models, hence their non-emission for the sake of completeness. These simple models will be referred to as elementary spatial interaction models.

We saw no purpose in writing down the complete derivation of all models taken account of by this paper. The derivation is generally known, and contained in many studies, including the original works by Wilson (1967, 1970 and 1974). We therefore decided to derive only the elementary interaction models whose derivation is usually omitted. This work is offered in condensed form in the Appendix.

This paper examines eleven spatial interaction models, denoted by the Roman numerals I to XI. Model equations (including ones relating to balancing factors) and constraints used in their derivation are presented in Table 1. Models I to IV are the elementary spatial interaction models derived in the Appendix. Models V and VI are two singly constrained spatial interaction models with the exponential distance function. Model V is the production-constrained model, whereas model VI is the attraction-constrained model. Exactly the same principle of model specification is applied to derive the two singly constrained spatial interaction models with the power distance function denoted by the numerals VII and VIII. Model VII is related to model V and model VIII to model VI. Finally, three doubly constrained spatial interaction models are also considered. Model IX is one with the exponential distance function, while two other models can be produced by replacing the exponential distance function by the power distance function (model X) or Tanner's distance function (model XI).

Table 1. Spatial interaction models used in the analysis

Model	Constraint(s)	Model Equation(s)
I	$\sum_{i=1}^m \sum_{j=1}^n T_{ij} = T$	$T_{ij} = T/mn$
II	$\sum_{j=1}^n T_{ij} = O_i \quad (i = 1, 2, \dots, m)$	$T_{ij} = O_i/n$
III	$\sum_{i=1}^m T_{ij} = D_j \quad (j = 1, 2, \dots, n)$	$T_{ij} = D_j/m$
IV	$\sum_{j=1}^n T_{ij} = O_i \quad (i = 1, 2, \dots, m)$ $\sum_{i=1}^m T_{ij} = D_j \quad (j = 1, 2, \dots, n)$	$T_{ij} = O_i D_j / T$
V	$\sum_{j=1}^n T_{ij} = O_i \quad (i = 1, 2, \dots, m)$ $\sum_{i=1}^m \sum_{j=1}^n T_{ij} c_{ij} = C$	$T_{ij} = A_i O_i \exp(-\beta c_{ij})$ $A_i = \left[\sum_{j=1}^n \exp(-\beta c_{ij}) \right]^{-1}$
VI	$\sum_{i=1}^m T_{ij} = D_j \quad (j = 1, 2, \dots, n)$ $\sum_{i=1}^m \sum_{j=1}^n T_{ij} c_{ij} = C$	$T_{ij} = B_j D_j \exp(-\beta c_{ij})$ $B_j = \left[\sum_{i=1}^m \exp(-\beta c_{ij}) \right]^{-1}$
VII	$\sum_{j=1}^n T_{ij} = O_i \quad (i = 1, 2, \dots, m)$ $\sum_{i=1}^m \sum_{j=1}^n T_{ij} \ln c_{ij} = C'$	$T_{ij} = A_i O_i c_{ij}^{-\beta}$ $A_i = \left[\sum_{j=1}^n c_{ij}^{-\beta} \right]^{-1}$
VIII	$\sum_{i=1}^m T_{ij} = D_j \quad (j = 1, 2, \dots, n)$ $\sum_{i=1}^m \sum_{j=1}^n T_{ij} \ln c_{ij} = C'$	$T_{ij} = B_j D_j c_{ij}^{-\beta}$ $B_j = \left[\sum_{i=1}^m c_{ij}^{-\beta} \right]^{-1}$
IX	$\sum_{j=1}^n T_{ij} = O_i \quad (i = 1, 2, \dots, m)$ $\sum_{i=1}^m \sum_{j=1}^n T_{ij} = D_j \quad (j = 1, 2, \dots, n)$ $\sum_{i=1}^m \sum_{j=1}^n T_{ij} c_{ij} = C$	$T_{ij} = A_i B_j O_i D_j \exp(-\beta c_{ij})$ $A_i = \left[\sum_{j=1}^n B_j D_j \exp(-\beta c_{ij}) \right]^{-1}$ $B_j = \left[\sum_{i=1}^m A_i O_i \exp(-\beta c_{ij}) \right]^{-1}$

Model	Constraint(s)	Model Equation(s)
X	$\sum_{j=1}^n T_{ij} = O_i \quad (i=1,2,\dots,m)$ $\sum_{i=1}^m T_{ij} = D_j \quad (j=1,2,\dots,n)$ $\sum_{i=1}^m \sum_{j=1}^n T_{ij} \ln c_{ij} = C'$	$T_{ij} = A_i B_j O_i D_j c_{ij}^{-\beta}$ $A_i = \left[\sum_{j=1}^n B_j D_j c_{ij}^{-\beta} \right]^{-1}$ $B_j = \left[\sum_{i=1}^m A_i O_i c_{ij}^{-\beta} \right]^{-1}$
XI	$\sum_{j=1}^n T_{ij} = O_i \quad (i=1,2,\dots,m)$ $\sum_{i=1}^m T_{ij} = D_j \quad (j=1,2,\dots,n)$ $\sum_{i=1}^m \sum_{j=1}^n T_{ij} c_{ij} = C$ $\sum_{i=1}^m \sum_{j=1}^n T_{ij} \ln c_{ij} = C'$	$T_{ij} = A_i B_j O_i D_j \exp(-\beta c_{ij}) c_{ij}^{-\alpha}$ $A_i = \left[\sum_{j=1}^n B_j D_j \exp(-\beta c_{ij}) c_{ij}^{-\alpha} \right]^{-1}$ $B_j = \left[\sum_{i=1}^m A_i O_i \exp(-\beta c_{ij}) c_{ij}^{-\alpha} \right]^{-1}$

Notation in a migration context: m is the number of origin areas; n is the number of destination areas; T_{ij} is the number of migrations from origin area i to destination area j ; O_i is the known total outmigration from area i ; D_j is the known total immigration into area j ; T is the known total number of migrations in the system; c_{ij} is the distance between areas i and j ; C (or C') is the observed total migration length; α and β are distance decay parameters estimated in the model calibration; A_i and B_j are the balancing factors ensuring that the constraints on total outmigration and/or total immigration are satisfied.

All tables and figures in the paper are authors' own elaborations.

Empirical evaluation

This section presents empirical evidence as to this paper's basic statement that the constraints imposed on an objective function are able to reduce the entropy of the corresponding distributions produced by the entropy-maximizing models.

The theoretical arguments presented in the preceding section were evaluated through application to interregional migration flows in Slovakia. The evaluation procedure was relatively straightforward and similar to the above theoretical discussion. The spatial interaction models I to XI presented in Table 1 were fitted separately to each of the six sets of migration data. For each model the predicted $\{T_{ij}\}$ distribution was first derived, before T_{ij} 's were converted to P_{ij} 's and the corresponding predicted entropy, H_p (in bits), was calculated. The entropy of observed distribution, H_o , expressed by observed migration flows, was also computed. Finally, the entropy values corresponding to all the models were compared. The entire procedure was repeated for each of the six data-sets.

Migration data

As a spatial interaction model is fitted, the basic data requirements are for a matrix containing flows between a set of origins and destinations in some specific period, and (in the case of models V to XI) a corresponding matrix of inter-area distances. However, in the present context, certain qualifications must be applied to the observed interaction matrix, if computation of the entropy measure is to be made possible. In the first place, all entries of the interaction matrix must be known. This is to say that flows within regions must be known and taken into consideration. Finally, the interaction matrix must comprise non-zero values only. Clearly, the practical problems of collecting accurate interaction data can be severe in this context. In this paper, we use data on internal migration in Slovakia as the only kind of spatial interaction data that are simultaneously suitable and available.

The migration data used in our analysis were obtained from the records of the current registration of population in Slovakia. They relate to changes of permanent residence between two communes (the smallest units of local administration) in the country. It follows from the definition that data on internal migration in Slovakia are counts of moves rather than of transitions. Data on migration before 2000 were taken from the basic source *Pohyb obyvateľstva (Population Change)*, published annually by the Statistical Office of the Slovak Republic or the Federal Statistical Office of former Czechoslovakia, while all information concerning migration flows after 2000 was extracted from unpublished tabulations provided by the Statistical Office of the Slovak Republic. The annual migration data were consolidated into five-year sets to avoid the problems of sparse matrices which arise with data for single-year migration flows.

As already noted above, the evaluation procedure was applied to the six migration data-sets labelled A, B, C, D, E, and F, each with a different number of origin and destination regions. The data-sets A, B, and C relate to migration flows between functional urban regions (FURs) in Slovakia during the 2001-05 period. The FURs were derived by Bezák (2000a) using the journey-to-work data from the 1991 Census of Population and Housing. They were defined as spatially contiguous groups of communes which are cohesive internally and (relatively) self-contained in terms of daily commuting. To avoid the problem of sparse interaction matrices, it was decided to modify the original set of fifty-one functional urban regions as follows.

Data-set A relates to migration flows between the forty-four FURs modified by allocating each of seven original FURs with zero inflows and/or outflows to the adjacent FUR to which it sends the largest commuting flow. Data-set B includes only data on migration between the twenty-five largest FURs with population exceeding 80 thousand in 2001. To produce the data set C, the fifty-one original FURs were aggregated into twelve functional macro-regions, using

the INTRAMAX regionalization procedure (Masser and Scheurwater, 1978) and interregional commuting data from the 2001 Population Census. In all three cases, distances between and within regions were measured as mean straight-line distances between population-weighted centroids of their constituent communes, using the procedure described by Bell et al. (2002).

Data-set D contains the most recent data on migration between FURs during the 2011-15 period. However, to avoid zero values, only moves between twenty randomly selected FURs with non-zero inflows and outflows were used. Distances between and within regions were determined repeatedly through application of the procedure referred to above.

Finally, data-sets E and F refer to migration flows between former administrative districts that were defined in Slovakia prior to the 1996 administrative reform. Data-set E includes migration data for the 1981-85 period, while data-set F relates to those for the 1991-95 period. In order to obtain comparable regional units, each of the two urban districts of Bratislava and Košice was amalgamated with its surrounding rural district into one metropolitan area, with the effect that the number of districts used in analysis was reduced from thirty-eight to thirty-six. Distances between districts were measured as the road distances between district centres, while the average distances within districts were estimated as half of the radius of a circle of equal area, following the approach suggested by Rogerson (1990).

To complete the discussion, it should be noted that, in all six cases, migrations within regions (or districts) were defined as moves between their constituent communes.

Calibration and results

Spatial interaction models V to XI were calibrated using a maximum likelihood method based on the Newton-Raphson iterative routine for solving systems of non-linear equations (Batty and Mackie, 1972). The computer program used in this exercise was a modified version of one described in Stillwell (1983). The original program was updated, extended and rewritten into the C++ programming language by the second author. The entropy values, H_p (in bits), obtained by fitting the spatial interaction models to the six sets of migration data are as presented in Table 2.

Two points need stressing as entropy values contained in Table 2 are interpreted. First, the order of the models in Table 2 is consistent with their sequence in Table 1 and, as a result, this does not correspond exactly with the reduction in entropy values. Second, it needs to be emphasised that not every pair of consecutive models is comparable with respect to entropy reduction. For example, model IX is not comparable with model VIII on account of a different distance decay function resulting from a different migration length constraint (see Table 1). In

Table 2. Entropy values in bits for spatial interaction models fitted to the six sets of migration data

Model	Data set					
	A	B	C	D	E	F
I	10.920	9.288	7.170	8.644	10.340	10.340
II	10.675	9.121	7.131	8.439	10.228	10.223
III	10.623	9.072	7.105	8.279	10.153	10.198
IV	10.379	8.905	7.066	8.075	10.041	10.082
V	8.919	7.403	5.776	7.011	8.332	8.265
VI	8.873	7.352	5.743	6.820	8.242	8.233
VII	8.230	6.810	5.489	6.399	7.796	7.681
VIII	8.175	6.744	5.457	6.176	7.702	7.650
IX	8.636	7.215	5.706	6.657	8.195	8.188
X	7.945	6.620	5.408	6.042	7.671	7.622
XI	7.849	6.519	5.346	5.928	7.662	7.606
OBS	7.732	6.435	5.229	5.851	7.544	7.472

Note: The label OBS denotes the observed distribution of migration flows.

spite of this, Table 1 retains some value, as it allows us to compare entropy values corresponding with a particular model across different data-sets.

To facilitate correct comparison of entropy values among models, the said values for individual data-sets are as shown separately in Figs. 1-3. In these figures, each model is represented by a small rectangle, which is marked with the associate Roman numeral and simultaneously contains the corresponding entropy value. The elementary spatial interaction model I, with a single balancing constraint on total migration is placed at the top of the figure and the rectangle related to the observed distribution (labelled OBS) is located at the bottom of the figure. Other models can be found at several intervening levels, each of them referring to a different category of spatial interaction model. Note that models placed at the same level are distinguished only by the nature (or complexity) of the constraints imposed on an objective function. On the other hand, models placed at different levels are also distinguished by the number (or amount) of constraints incorporated.

For example, elementary spatial interaction models II and III, each with one balancing constraint on total outmigration or total immigration, are placed at the second level. However, elementary spatial interaction model IV, derived by using both constraints, is located at the third level. The fourth level includes four singly constrained spatial interaction models also with balancing constraints on total outmigration or total immigration. In addition, a migration-length constraint is imposed on the migrations in these models. It needs to be empha-

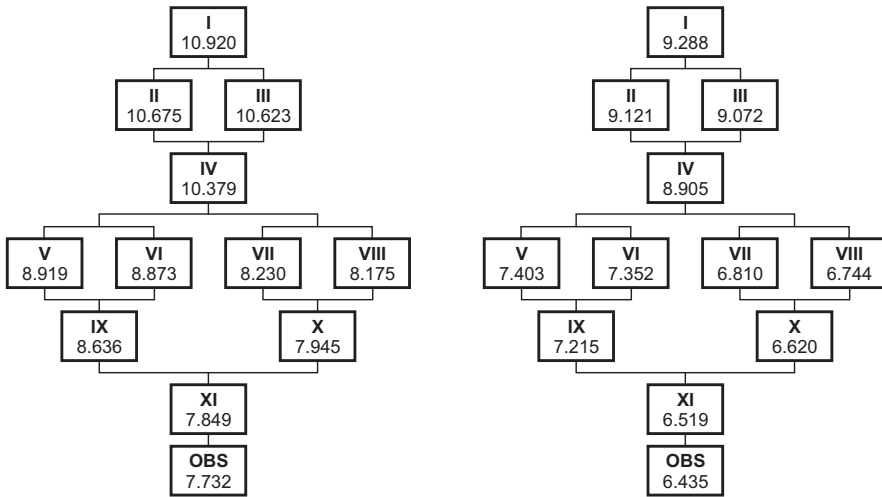


Fig. 1. Entropy values in bits for data-sets A and B

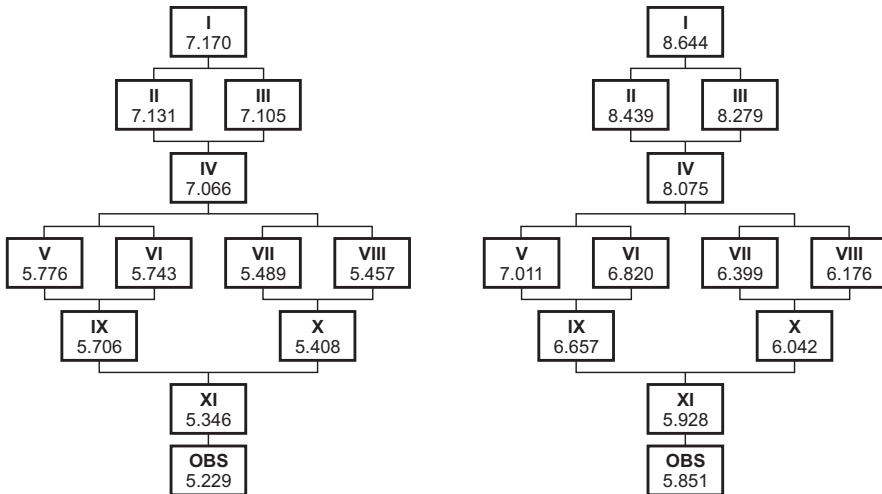


Fig. 2. Entropy values in bits for data-sets C and D

sised, however, that the modification of the migration-length constraint leads in this case to two versions of the spatial interaction models with different distance functions. It can be supposed, therefore, that the initial maximum entropy, H_{max} , corresponding with elementary spatial interaction model I, is successively reduced from the top to the bottom of the figure just along lines connecting the

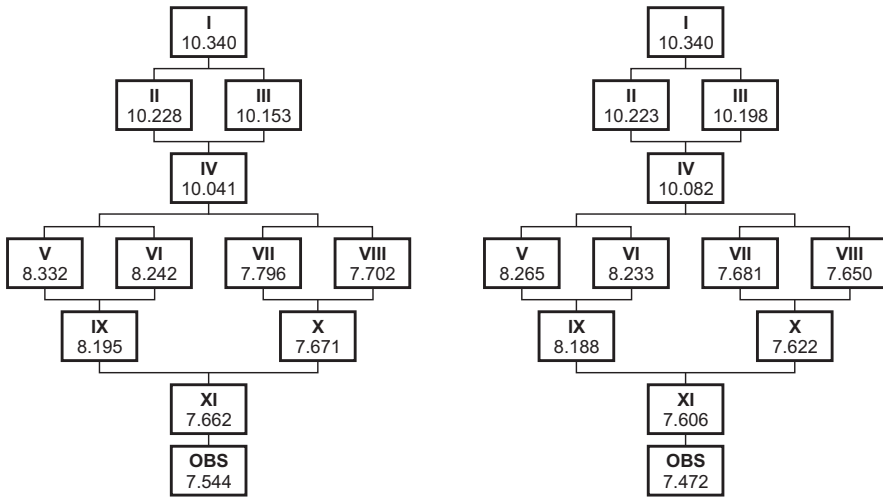


Fig. 3. Entropy values in bits for data-sets E and F

individual models. Moreover, it is clear that only the models connected with these lines are comparable with respect to the entropy reduction.

The main results derived from these figures can be summarised as follows. In the first place, it is evident that constraints imposed on an objective function reduce the initial maximum entropy. As expected, the reduction depends on the number and nature of constraints incorporated. All figures show clearly that the spatial interaction models with a distance decay function contribute far more to entropy reduction than do elementary interaction models I to IV. In turn, the doubly constrained models IX to XI with the complete system of constraints prove more successful in reducing the initial entropy than all comparable versions of singly constrained model. Another interesting point is that the reductions in initial entropy achieved by attraction-constrained models VI and VIII are greater than those generated by the corresponding production-constrained models V and VII.

As regards the form of the distance decay function used, it is evident that the negative power function in any case provides for a greater reduction of the initial entropy than does the negative exponential function. This result is consistent with the broad consensus that the power function is more appropriate for analysing longer-distance interactions such as migration flows (cf. Fotheringham and O'Kelly, 1989, pp. 12-13), as confirmed also by the study on interregional migration in Slovakia (Bezák, 2000b). Furthermore, it is clear that the best approximation of the predicted entropy, H_p , to the entropy of observed distribution, H_o , has been achieved in the case of doubly constrained model XI with Tanner's distance decay function. We can conclude, therefore, that from among the set

of models considered, model XI is the most appropriate spatial interaction model according to the methodology discussed in this paper.

Conclusion

The principal conclusion to be drawn from this study is that the constraints imposed on an objective function are able to reduce the entropy of the corresponding distributions produced by the entropy-maximizing spatial interaction models. It should be emphasised that this reduction is gradual, and depends on the number and nature of the constraints imposed. In addition, we have arrived at certain conclusions concerning the kinds of constraints most successful in reducing initial maximum entropy. Finally, we have demonstrated that the doubly constrained spatial interaction model with Tanner's distance function can be regarded as the best model from among the set considered, as far as the suggested methodology is concerned.

We can thus conclude that the basic statement of this paper is now supported by some empirical evidence. However, we are far from any generalization based on these conclusions, because our examination was based on only a limited number of cases. For this reason, further examinations in the same direction would be welcome.

*

Acknowledgments. The authors are grateful to Professor Gejza Wimmer of the Mathematical Institute of the Slovak Academy of Sciences in Bratislava for his very useful comments on an earlier version of this paper. Thanks are extended to RNDr. Zuzana Podmanická from the Statistical Office of the Slovak Republic in Bratislava, for providing the unpublished migration data used in our analysis. The constructive comments of three anonymous referees are also appreciated.

The research reported in this paper was supported, in part, by the Scientific Agency of the Ministry of Education, Science, Research and Sport of the Slovak Republic and the Slovak Academy of Sciences, under Grant 1/0540/16 – "Social, economic and environmental determinants of regional development and transformation: a regional geographic approach".

References / Piśmiennictwo

- Batty M., 2010, *Space, scale, and scaling in entropy maximizing*, Geographical Analysis, 42, 4, pp. 395-421.
- Batty M., Mackie S., 1972, *The calibration of gravity, entropy, and related models of spatial interaction*, Environment and Planning, 4, 2, pp. 205-233.
- Bell M., Blake M., Boyle P., Duke-Williams O., Rees P., Stillwell J., Hugo G., 2002, *Cross-national comparison of internal migration: issues and measures*, Journal of the Royal Statistical Society, Series A, 165, 3, pp. 435-464.

- Bezák A., 2000a, *Funkčné mestské regióny na Slovensku (Functional urban regions in Slovakia)*, Geographia Slovaca, 15, Slovak Academy of Sciences, Institute of Geography, Bratislava.
- Bezák A., 2000b, *Interregional migration in Slovakia: some tests of spatial interaction models*, Geografický časopis, 52, 1, pp. 15-32.
- Fotheringham A.S., O'Kelly M.E., 1989, *Spatial Interaction Models: Formulations and Applications*, Kluwer, Dordrecht.
- Garrett A.J.M., 1991, *Macroirreversibility and microreversibility reconciled: the second law*, [in:] B. Buck, V.A. Macaulay (eds.), *Maximum Entropy in Action*, Clarendon Press, Oxford, pp. 139-170.
- Jaynes E.T., 1957, *Information theory and statistical mechanics*, Physical Review, 106, 4, pp. 620-630.
- Jaynes E.T., 1979, *Where do we stand on maximum entropy?* [in:] R.D. Levine, M. Tribus (eds.), *The Maximum Entropy Formalism*, MIT Press, Cambridge (Mass.), pp. 15-118.
- Masser I., Scheurwater J., 1978, *The specification of multi-level systems for spatial analysis*, [in:] I. Masser, P.J.B. Brown (eds.), *Spatial Representation and Spatial Interaction*, Martinus Nijhoff, Leiden, pp. 151-172.
- Pohyb obyvateľstva v Českej a Slovenskej Federatívnej Republike v roce 1991, 1993*, Český statistický úřad, Praha.
- Pohyb obyvateľstva v Československé socialistické republike v roce ... [1981, 1982, 1983, 1984, 1985], 1982-1986*, Federální statistický úřad, Praha.
- Pohyb obyvateľstva v Slovenskej republike v roku ... [1992, 1993, 1994, 1995], 1993-1996*, Štatistický úrad Slovenskej republiky, Bratislava.
- Rogerson P.A., 1990, *Buffon's needle and the estimation of migration distances*, Mathematical Population Studies, 2, 3, pp. 229-238.
- Roy J.R., Thill J.-C., 2004, *Spatial interaction modelling*, Papers in Regional Science, 83, 1, pp. 339-361.
- Stillwell J.C.H., 1983, *SPAINTE: A Computer Program for Spatial Interaction Model Calibration and Analysis*, Computer Manual, 14, University of Leeds, School of Geography, Leeds.
- Wilson A.G., 1967, *A statistical theory of spatial distribution models*, Transportation Research, 1, 3, pp. 253-269.
- Wilson A.G., 1970, *Entropy in Urban and Regional Modelling*, Pion, London.
- Wilson A.G., 1974, *Urban and Regional Models in Geography and Planning*, Wiley, Chichester.
- Yaglom A.M., Yaglom I.M., 1983, *Probability and Information* [Translated from the Russian (1973) by V. K. Jain], D. Reidel, Dordrecht.

[Received: June 2017; accepted: October 2017]

Appendix

Derivation of the elementary interaction models I to IV (see Table 1)

Notation: W is the number of microstates (the number of distinct arrangements of individual movements which give rise to the distribution $\{T_{ij}\}$); S is entropy (the objective function to be maximized); T_{ij} is the size of the interaction between origin i and destination j (i.e. the number of individual movements from area i to area j)

Model I

$$W = \frac{T!}{\prod_i \prod_j T_{ij}!}$$

$$S = \ln W = \ln T! - \sum_i \sum_j \ln T_{ij}!$$

Stirling's approximation:

$$\ln T! = T \ln T - T$$

$$\ln T_{ij}! = T_{ij} \ln T_{ij} - T_{ij}$$

$$\text{Constraint: } \sum_i \sum_j T_{ij} = T$$

Lagrangian:

$$L = T \ln T - T - \sum_i \sum_j T_{ij} \ln T_{ij} - T_{ij} + \omega \left(T - \sum_i \sum_j T_{ij} \right)$$

$$\frac{\partial L}{\partial T_{ij}} = -\ln T_{ij} - \omega$$

$$\frac{\partial L}{\partial T_{ij}} = 0 \Rightarrow -\ln T_{ij} - \omega = 0 \Rightarrow T_{ij} = \exp(-\omega)$$

$$\sum_i \sum_j T_{ij} = T \Rightarrow \exp(-\omega) mn = T$$

$$\text{Since } T_{ij} = \exp(-\omega) \Rightarrow T_{ij} = \frac{T}{mn}$$

Model II

$$W = \frac{O_i!}{\prod_j T_{ij}!}$$

$$S = \ln W = \ln O_i! - \sum_j \ln T_{ij}!$$

Stirling's approximation:

$$\ln O_i! = O_i \ln O_i - O_i$$

$$\ln T_{ij}! = T_{ij} \ln T_{ij} - T_{ij}$$

Constraint: $\sum_j T_{ij} = O_i$

Lagrangian:

$$L = O_i \ln O_i - O_i - \sum_j T_{ij} \ln T_{ij} - T_{ij} + \sum_i \lambda_i \left(O_i - \sum_j T_{ij} \right)$$

$$\frac{\partial L}{\partial T_{ij}} = -\ln T_{ij} - \lambda_i$$

$$\frac{\partial L}{\partial T_{ij}} = 0 \Rightarrow -\ln T_{ij} - \lambda_i = 0 \Rightarrow T_{ij} = \exp(-\lambda_i)$$

$$\sum_j T_{ij} = O_i \Rightarrow \exp(-\lambda_i) n = O_i$$

Since $T_{ij} = \exp(-\lambda_i) \Rightarrow T_{ij} = \frac{O_i}{n}$

Model III

$$W = \frac{D_j!}{\prod_i T_{ij}!}$$

$$S = \ln W = \ln D_j! - \sum_i \ln T_{ij}!$$

Stirling's approximation:

$$\ln D_j! = D_j \ln D_j - D_j$$

$$\ln T_{ij}! = T_{ij} \ln T_{ij} - T_{ij}$$

Constraint: $\sum_i T_{ij} = D_j$

Lagrangian:

$$L = D_j \ln D_j - D_j - \sum_i T_{ij} \ln T_{ij} - T_{ij} + \sum_j \delta_j \left(D_j - \sum_i T_{ij} \right)$$

$$\frac{\partial L}{\partial T_{ij}} = -\ln T_{ij} - \delta_j$$

$$\frac{\partial L}{\partial T_{ij}} = 0 \Rightarrow -\ln T_{ij} - \delta_j = 0 \Rightarrow T_{ij} = \exp(-\delta_j)$$

$$\sum_j T_{ij} = D_j \Rightarrow \exp(-\delta_j) m = D_j$$

$$\text{Since } T_{ij} = \exp(-\delta_j) \Rightarrow T_{ij} = \frac{D_j}{m}$$

Model IV

$$W = \frac{T!}{\prod_i \prod_j T_{ij}!}$$

$$S = \ln W = \ln T! - \sum_i \sum_j \ln T_{ij}!$$

Stirling's approximation:

$$\ln T! = T \ln T - T$$

$$\ln T_{ij}! = T_{ij} \ln T_{ij} - T_{ij}$$

$$\text{Constraints: } \sum_j T_{ij} = O_i, \sum_i T_{ij} = D_j$$

Lagrangian:

$$L = T \ln T - T - \sum_i \sum_j T_{ij} \ln T_{ij} - T_{ij} + \sum_i \lambda_i \left(O_i - \sum_j T_{ij} \right) + \sum_j \delta_j \left(D_j - \sum_i T_{ij} \right)$$

$$\frac{\partial L}{\partial T_{ij}} = -\ln T_{ij} - \lambda_i - \delta_j$$

$$\frac{\partial L}{\partial T_{ij}} = 0 \Rightarrow -\ln T_{ij} - \lambda_i - \delta_j = 0 \Rightarrow T_{ij} = \exp(-\lambda_i) \exp(-\delta_j)$$

$$\sum_j T_{ij} = O_i \Rightarrow \exp(-\lambda_i) \sum_j \exp(-\delta_j) = O_i \Rightarrow \exp(-\lambda_i) = \frac{O_i}{\sum_j \exp(-\delta_j)}$$

$$\sum_i T_{ij} = O_i \Rightarrow \sum_i \exp(-\lambda_i) \exp(-\delta_j) = D_j \Rightarrow \exp(-\delta_j) = \frac{D_j}{\sum_i \exp(-\lambda_i)}$$

$$\begin{aligned} T_{ij} = \exp(-\lambda_i) \exp(-\delta_j) &\Rightarrow T_{ij} = \frac{O_i}{\sum_j \exp(-\delta_j)} \frac{D_j}{\sum_i \exp(-\lambda_i)} = \\ &= \frac{O_i D_j}{\sum_i \sum_j \exp(-\lambda_i) \exp(-\delta_j)} = \frac{O_i D_j}{T} \end{aligned}$$

JÁN PAULOV | ANTON BEZÁK

OGRANICZENIA JAKO CZYNNIKI REDUKUJĄCE ENTROPIĘ ROZKŁADÓW: PRZYKŁAD MODELU MAKSYMALIZACJI ENTROPII PRZESTRZENNEJ INTERAKCJI

Celem artykułu jest dostarczenie empirycznego dowodu twierdzenia, że ograniczenia nałożone na funkcję celu mogą zredukować entropię odpowiednich rozkładów, stworzoną przez modele maksymalizacji entropii. Tę ideę ocenia się na podstawie zastosowania do modelu maksymalizacji entropii przestrzennej interakcji jako typowego reprezentanta rodziny modeli maksymalizacji entropii. Jednocześnie różnych wersji modelu przestrzennej interakcji dopasowuje się kolejno do sześciu zbiorów danych dotyczących międzyregionalnych migracji na Słowacji. Dla każdego modelu wpierw prognozuje się rozkłady przepływów, później oblicza się odpowiednią przewidywaną entropię, a na końcu porównuje się wartości entropii odnoszące się do wszystkich modeli.

Otrzymane wyniki jednoznacznie wskazują, że ograniczenia nałożone na funkcję celu sukcesywnie redukują początkową maksymalną entropię, i że ta redukcja zależy od liczby i charakteru nałożonych ograniczeń. Warto podkreślić, że wspomniana redukcja jest stopniowa. Co więcej, w artykule wyprowadzamy, przynajmniej jeśli chodzi o pewne wnioski na temat rodzajów ograniczeń te, które są najlepsze w redukcji początkowej maksymalnej entropii. Wreszcie, wykazujemy, że w zbiorze rozważanych modeli podwójnie ograniczony model przestrzennej interakcji z funkcją odległości Tannera można uważać za najlepszy, przynajmniej jeśli chodzi o zastosowaną metodologię.

Główne twierdzenie niniejszego artykułu uzyskało zatem dowód empiryczny, trzeba jednak pamiętać, że wspomniane uogólnione wnioski oparte są na ograniczonej liczbie przypadków. Dlatego wskazane są dalsze badania idące w podobnym kierunku.



