BADANIA FIZJOGRAFICZNE
R. I – SERIA D – GOSPODARKA PRZESTRZENNA str. 7–19

SCENARIO – BASED SIMULATION MODEL OF LAND USE DEVELOPMENT

JAKUB VOREL[1], KAREL MAIER[1], STANISLAV GRILL[2]

[1]Czech Technical University in Prague, Faculty of Architecture, Department of Spatial Planning, Thákurova 7, Praha 6, 166 34, Praha
[2]Charles University in Prague, Faculty of Science, Department of Geography, Albertov 6, Praha 2, 128 43, Praha

Abstract: The paper presents the application of theoretical concepts and methodology of land use modelling on prediction of future development of territory. The chosen approach is to replicate the urban growth dynamics of medium-size city region in the form of a simulation model and to predict alternative scenarios of future land use pattern. The concept, implementation and validation of the simulation model are briefly discussed. The Kappa, Fuzzy Kappa and Fractal dimension measures indicate the correspondence of the simulation model with the alternative random model and with the observation of real land use patterns. While the simulation model outperforms the alternative random model in the location of family houses and recreation houses, its predictions still have weak correspondence for multi-family houses and region-scale retail facilities with the observed land uses. The paper presents presumable causes of the inaccuracy of location predictions and proposes several possibilities for improving the model.

Keywords: land use, urban growth, simulation model, methodology

INTRODUCTION

Land use planning has a significant long-term impact on the social, environmental and economic aspects of the future development of the living environment. Only recently the impact assessment of planning proposals have become an obligatory and integral part of the planning process and associated documentation. New national legislation in the Czech Republic in the field of spatial planning and management adopted European Directive 2001/42/EC on the effects of certain plans and programmes on the environment that prescribes the obligatory impact assessment of policies, concepts and plans on the sustainable development of the environment. It is assumed that new legal requirements will stimulate a demand for predictions of land use development and for assessment of its impact on various aspects of the living environment. The urban simulation models are one of several ways of meeting these requirements.

The theoretical background for urban simulation modelling has been developed over the last five decades on the basis of Alonso's (Alonso 1964) and Mill's Urban Land Market simulation models (Wilson 1974, Barra 1989), Lowry's interaction based models, Wingo's entropy-based models (Wilson 1974), Forrester's Urban Dynamics model (Forrester 1969) and Stocks and Flows urban models (Wilson 1974). The latest modelling approaches focus on the decision-making of individual agents such as households, companies or developers. Environmental change is considered as the emergent outcome of many individual choices made in constrained circumstances (Waddell 2002; Allen 2004; Barra 1989; Briassoulis 2000).

The paper presents the application of theoretical concepts and methodology of land use modelling in the prediction of future development of terrain. The approach chosen aims to replicate the urban growth dynamics of a medium-size city region in the form of a simulation model and to predict alternative scenarios for future land use pattern. The concept, implementation, outcomes and validation of the simulation model are briefly discussed in the following sections.

BASIC CONCEPTS OF THE SIMULATION MODEL

The simulation model presented was inspired by the concept of Lowry's well-known model that considers population concentration as the major attractor for the localization of secondary economy activities (Wilson 1974; Barra 1989).

The proposed model first distributes the demand for specific land uses by locating and relocating individual households. Factors such as access to public infrastructure and services, proximity to natural elements and adjacency of other land uses influences the population growth potential of particular localities. The attractivity of the living environment, together with the vacancy rate of housing stock and the housing type preferences of the population determine the demand for residential land uses. The distribution of the population in space successively creates opportunities for the location of small-scale mixed retail activities and services as well as region-scale retail centres. The basic modelling cycle runs through the following steps (Vorel 2007; Maier 2007; Grill 2008):

1. Initiation of the model by base year data and parameters of simulated scenario.
2. Determination of cells attractivities for location of residential land uses.
3. Households and housing stock allocation based on cell attractivities.
4. Allocation of retail and mixed land-uses with regards to spatial distribution of population.
5. Actualization of variables and starting next cycle at step 2.
6. If end of simulation period à assessment of scenario.

ESTIMATION OF LOCATION ATTRACTIVITY FOR RESIDENTIAL USE

The relation between the land use drivers and resulting land use changes is established by the Multi-Nominal Logistic model:

$$\mathrm{P_j} = \frac{e^{U_j}}{\sum_{j}^{J} e^{U_j}} \tag{1}$$

Where utility function has linear form:

$$U_j = \sum_{k=1}^{K} \beta_{jk} x_k \tag{2}$$

The Multi-Nominal Logistic (MNL) model measures the probabilities that the calculated cell will realize one of the predefined category *j* that belongs to the set of land use transformations $j \in \langle 1..J \rangle$. The probability of a particular transformation *j* depends on utility improvement *Uj,* obtained by land use transformation. The utility function *Uj* has linear form and consists of $k \in \langle 1..K \rangle$. environmental characteristics *xk* that are specific to each cell but invariant for all categories of land use transformation and *ßjk* parameters that are constant for all cells but specific to each land use transformation category.

The model represents land uses in the form of orthogonal cellular space with the size of cell being 75 x 75 m. As the data directly describing the land uses are not available, the land uses are derived from data on uses of buildings. The Register of Buildings, maintained by the Czech Statistical Office (CZSO, 2010) contains data on use, age and number of floors of buildings in several time periods. The land-use of a particular cell is therefore determined by the use of the buildings that are located within the cell. Generally the most frequent building use determines the use of the cell. The land use changes are derived for the periods: 1961, 1981, 1991, 1996 and 2008.

The overview of several land use model applications and the vast amount of theoretical publications on land use change have resulted in establishing a preliminary set of land use change drivers:

a. externalities stemming from adjacency of different land uses,
b. intrinsic physical properties of each cell,
c. accessibility of transport infrastructure (Henderson 2004; Nijkamp 1986; Briassoulis 2000; EPA 2000; Beran 2005; Vorel 2007).

The concept of urban centres serves as a compound variable representing the concentration of cultural, educational, health care and sport facilities. The single compound variable is preferred to several individual variables due to strong spatial correlation among individual variables. Municipalities hosting between 14 and 49 public facilities were identified as the urban centres of local importance and municipalities with more than 49 public facilities were considered as urban centres of micro-regional importance.

To calculate the probabilities, first the *ßj* parameters of utility function *Uj* must be estimated by using the maximum-likelihood estimation technique. It is necessary to estimate the *ßj* parameters for each *k*-th independent variable *xk* and each land use transformation category $j \in \langle 1..J \rangle$. The $e^{ß}$ parameters (logits) of the Multi-Nominal Logistic model represents the change in odds of the probability of a particular land use change type *j* that is caused by a one-unit change in the independent variable *xk*, where $j = 0$ stands for change from un-built land to apartment houses and $j = 1$ stands for change from un-built land to family houses. The change of land use from non-built land to individual recreation houses $j = 2$ is the reference category. When $ßj > 0$ ($e^{ßj} > 1$) then the dependent variable *j* will be more probable, when $ßj < 0$ ($e^{ßj} < 1$) than the opposite is true. When $ßj = 0$ ($e^{ßj} = 1$) then the independent variable does not have any impact on the behaviour of the dependent variable.

Table 1. The $e^{ß}$ parameters (logits) of the Multi-Nominal Logistic model

index *k*	independent variables *xk*	$e^{ß}$ for $j = 0$	$e^{ß}$ for $j = 1$
0	distance to forest	3,255	1,707
1	distance to motorway slip-roads	1,083	1,239
2	road distance to urban centres of I. category	1,269	1,322
3	distance to road class I, II and III	0,895	0,881
4	distance to river	1,379	1,460
5	number of public facilities land-use cells in neighbourhood	2,849	0,961
6	number of apartment houses land-use cells in neighbourhood	5,179	2,617
7	number of family houses land-use cells in neighbourhood	1,178	2,170
8	number of recreation houses land-use cells in neighbourhood	0,850	0,306
9	number of mixed land-use cells in neighbourhood	2,919	1,801

The MNL model calculates the probability of land use transition from non-urban use to family residences, apartment houses and individual recreation. The resulting probabilities make it possible to predict the distribution of households in particular localities.

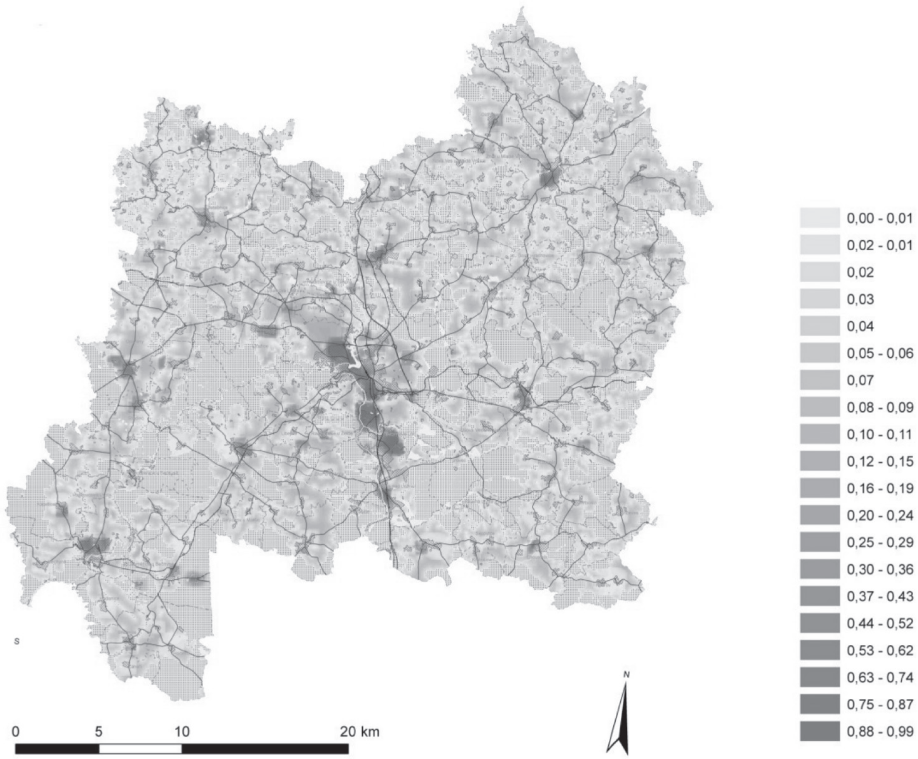

Fig. 1

POPULATION AND HOUSING STOCK ALLOCATION

The changes in the size of existing households induce the relocation of the inner-region population.

New population coming from outside the region and the population relocating inside the region form new households of an a priori specified size. In the case of overall decline in number of households, the households are de-allocated from the least attractive cells.

New households are successively allocated to cells with the highest attractivity with regards to their housing type preferences. Actual local vacancy rate in a neighbourhood is decisive for the decision whether to allocate new households into existing housing stock or to annex un-built land for construction of new houses.

LOCATION OF MIXED LAND USES

If the local population increase is big enough to economically sustain an extension of basic services, then the cells with residential use which at the same time have the highest accessibility to the local population are transformed to mixed land use. The number of residential cells to be transformed into mixed land uses is defined by the number of habitants, the per capita expenditure and the productivity of retail units. The origin-based interaction model is used for the derivation of attractivities of cells:

$$P_i^L = \sum_{j \in B_i} N_j * e^{-\beta d_{ij}} \tag{3}$$

$P^L i$ – attractivity of the cell *i* for location of mixed land use,
i – index of calculated cell,
j – index of cells of residential land use in basic statistical unit (neighbourhood),
Ni – number of habitants living in cell *j* of residential land use,
ß – distance decay parameter,
dij – distance between cells *i* and *j*,
Bi – basic statistical unit (neighbourhood) of the cell *i*.

Fig. 1. Example of one of three probability maps showing the cell attractivity for households that prefer to live in family houses

LOCALIZATION OF REGIONAL SCALE RETAIL SERVICES

Large scale retail facilities have a specific localization behaviour. Cell attractivity is calculated using the origin-based interaction model, taking into consideration the population of the whole modelled region.

$$P_i^R = \sum_{j \in R} N_j * e^{-\beta d_{ij}} \tag{4}$$

Pi^R – attractivity of the cell i for location of regional scale retail services,
i – index of calculated cell,
j – index of cells of residential land use,
Ni – number of habitants living in the cell *j* of residential land use,
ß – distance decay parameter,
dij – distance between the cells *i* and *j*,

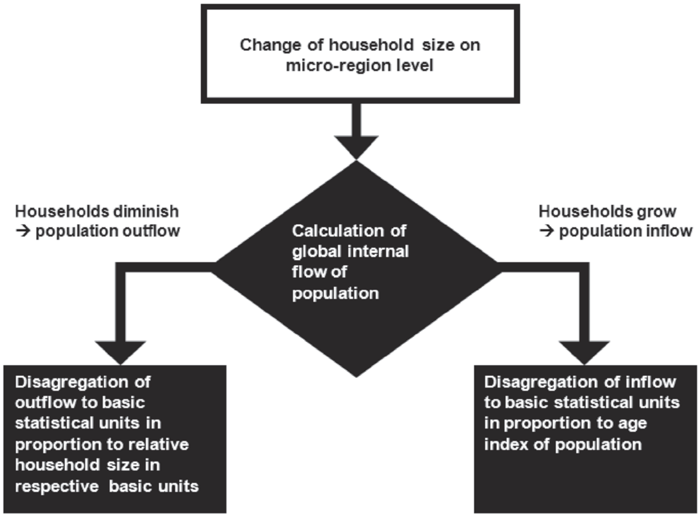

Fig. 2. Mechanism of inner population relocation based on household change

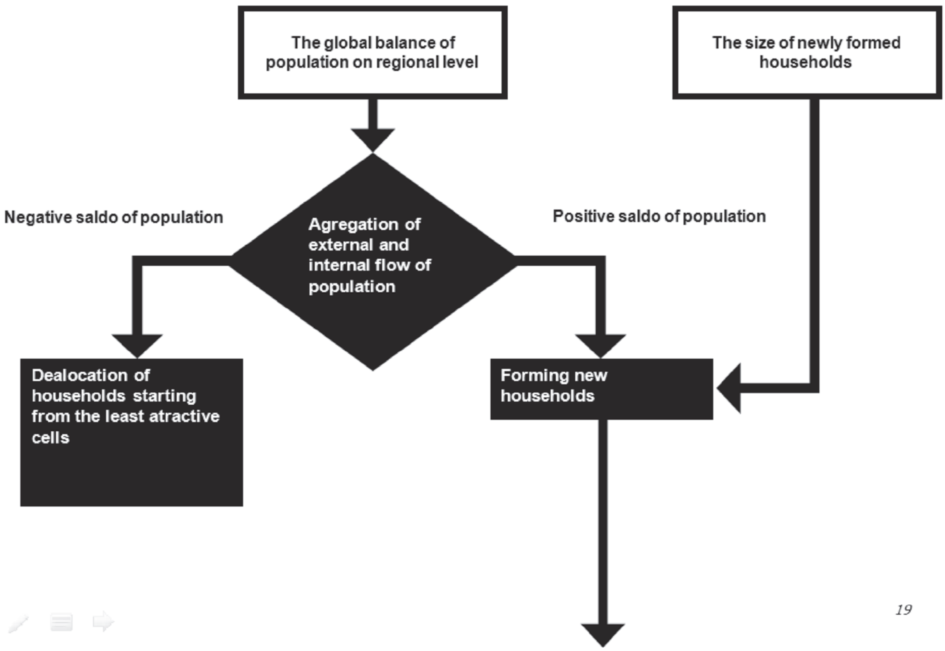

Fig. 3. Mechanism of de-allocation of existing households and formation of new households.

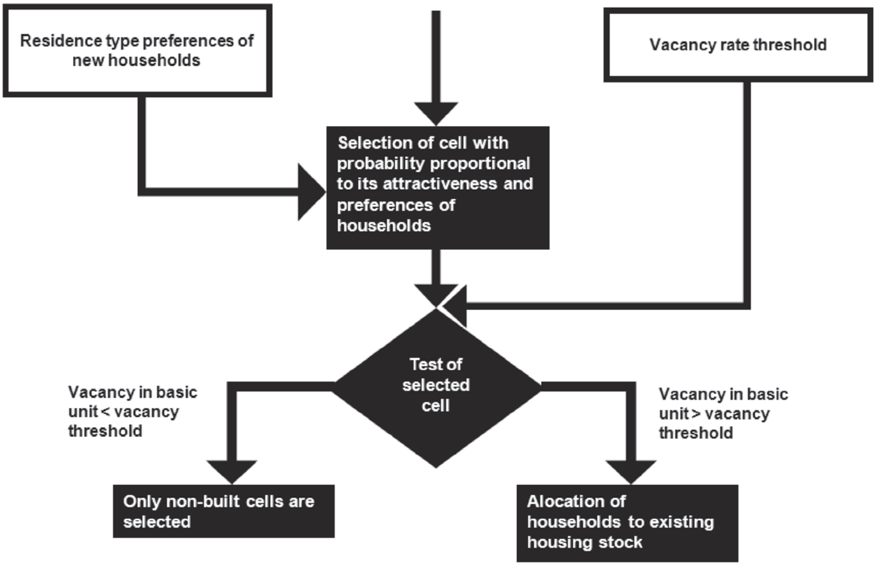

Fig. 4. Mechanism of allocation of new households

R – the area of micro-region (whole modelled territory).

The principle of economy of scale causes that regional retail premises usually occupy clusters of several adjacent cells. The model uses the adjacent cells' attractivities as positive feedback amplifying the attractivity of the calculated cell. The resulting attractivity (clustered attractivity) equals the product of the attractivities of the cells in immediate contact:

$$P_i^C = P_i^R * \prod_{j \in S_i} P_j \tag{5}$$

Pi^C – clustered attractivity of cell *i*,
i – index of calculated cell,
j – index of cells of residential land use,
Si – the neighbourhood of directly adjacent cells to cell *i*,
Pj – attractivity of the cell *j*,
Pi – attractivity of the cell *i*.

The total amount of allocated regional scale retail services is proportional to the changes in the size of the population in the modelled region.

MODEL VALIDITY

The validity of the model is measured by a comparison of the model outcome with the outcome of an alternative (reference) model and with observed land uses. The accuracy of fit indicates to what extent the output of the model corresponds to the observations. A random model was chosen as the reference model. It was generated using the Random Constrained Match of Map Comparison Kit software tool (RIKS, 2005). The random model randomly allocates the amount of land use categories that is identical with the tested model. The validity of the simulation model is considered as confirmed in the case where the accuracy of fit of the random model is less than that of the simulation model.

Three accuracy of fit measures were used:

- Kappa statistics,
- Fuzzy Kappa statistics,
- Fractal dimension.

The Kappa statistics assess the pixel by pixel similarity of two maps depicting observed or simulated land uses. The Fuzzy Kappa statistics express the degree of match between particular cell land uses. In the case of a perfect location match of the cells' land use the Fuzzy Kappa statistics equal one, in the case of the displacement of cell land use the statistics are calculated by the distance decay function (in this case the exponential decay, halving distance by 2 and the radius of a neighbourhood equal to 10 were used).

The Fractal dimension is an indicator of the shape complexity of land use patches. Values close to one represent a simple and smooth shape of the patch border, whereas values close to two represent more complex and convoluted shapes of patch border (RIKS 2005; Jasper 2006).

THE RESULTS AND DISCUSSION

The *ßj* parameters of the households allocation model were estimated using the land-use data in the years of 1981 and 1991. The model was then used to simulate the residential land use allocation in the same time period, and the land uses predicted by the simulation model for 1991 and the output of the random model were compared with the observed land uses in the same year.

The following table presents a comparison of the simulation model and the random model (in brackets) with the observed land uses based on the Kappa and Fuzzy Kappa measures. The last column presents the Fractal dimension of simulated and observed land use patterns. The land uses already existing before

1981, as well as more recent land use changes in the period of 1981–1991, were included in the comparison.

Table 2. Comparison of resulting pattern utilizing the accuracy of fit measures

Land use categories	Kappa	Fuzzy Kappa	Fractal dimension (observed/simulated pattern)
Family houses	0,944	0,974	1,485 / 1,483
Multi-family houses	0,689	0,765	1,653 / 1,604
Recreation houses	0,755	0,825	1,587 / 1,616
Mixed land uses	0,306	0,429	1,704 / 1,710
Regional scale retail services	0,462	0,468	1,512 / 1,983

Both the Kappa and Fuzzy Kappa measures indicate that the model has the best fit in the case of family houses' location and medium fit in the case of multi-family houses' and recreation houses' locations.

The poor prediction in the case of mixed land use can be partly explained by the lower number of land uses distributed by the model as a result of the underestimated variables of per capita expenditure and productivity of retail unit variables.

The emergence of the planned and deliberately concentrated development of large-scale retail services as well as of multi-family houses is very difficult to assess by means of statistics. The main reason is the small number of observed land use changes during the calibration period and the low variation of factors causing land use changes. One possible solution would be to overcome

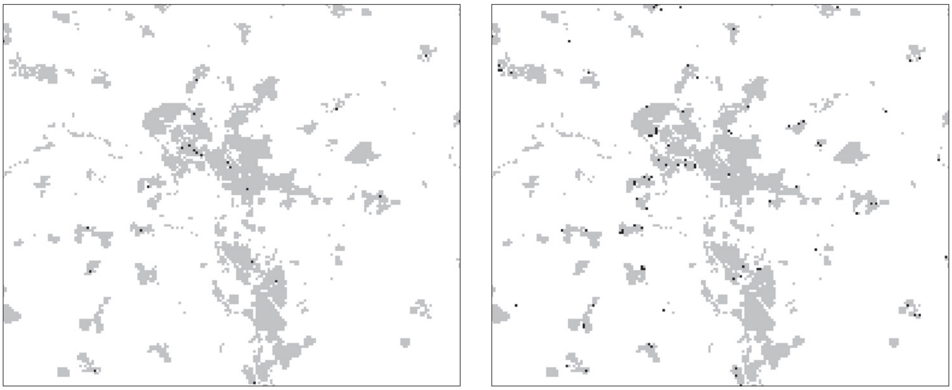

Fig. 5. Visual comparison of simulated (picture on left) and observed (picture on right) conversions from un-built land to residential land use of family houses from 1981 to 1991

the lack of observations by an extension of the study area, as the localization preferences of retail services and multi-family houses are expected to be similar across regions. But generally, it is becoming more and more obvious that the employment of the Multi-Nominal Logistic model is not appropriate in the situation of arbitrary decision making on the placement of large-scale, concentrated development, be it housing estates, industrial or public facilities. Even much more complex simulation models such as UrbanSim, leave the arbitrary location of big development project as exogenous to the model.

The Fractal dimension is an indicator of the shape complexity of patches of the same category of land uses. The relatively higher Fractal dimension of the predicted land use pattern compared to the observed land use pattern indicates the tendency of the simulation to produce fragmented patterns. The strips of individual recreation land uses stretched alongside a river or the ribbon like structures of big retail shopping centres alongside main roads are attached to their linear attractors and lack any strong mutual integration forces. In the case of family houses the behaviour of the simulation model is opposite. The land uses are allocated adjacently to existing clusters, resulting in a pattern of more compact patches.

The performance of the simulation model is usually compared with alternative models, in this case with the random model.

Table 3. Fit between simulated and observed land uses in the case of the simulation and random models

Land use transformation	Simulation model	Random model
j = 0(from non-built to family housing)	0,137	0,012
j = 1(from non-built to apartment housing)	0,069	0,002
j = 2(from non-built to individual recreation)	0,047	0,012
From non-built to mixed use	0,043	0,013

The Fuzzy Kappa measure shows that the simulation model outperforms the random model with regards to all land use transformation types.

CONCLUSION

The presented accuracy of fit measures proved useful in validating the land use simulation models.

The validation test failed in the case of multi-family houses prediction and region-scale retail services. The approach based on a statistical analysis seems inappropriate due to the low frequency of observed land use changes. The extension of the analysed territory could possibly bring a sufficient amount of the required data. A further improvement could be attained by building an individual MNL model for each of the above-mentioned land use categories with individual sets of independent variables that prove significant with regards to that specific land use category. On the other hand, the prediction of land use changes for family houses and recreation houses was verified with respect to land use match and shape complexity.

The simulation model was validated with regards to the randomized model, but the fit to observed land uses is still rather weak. One possible improvement in prediction could be made by using the presented accuracy-of-fit measures not only for the finalized model's validation, but for model calibration as well.

REFERENCES

Allen P.M. 2004: *Cities and regions as self-organizing systems*. Taylor & Francis, UK.

Alonso W. 1964: *Location and Land Use: Towards a General Theory of Land Rent*. Harvard University Press, Cambridge, US.

Barra T. 1989: *Integrated Land Use and Transport Modelling*. Cambridge University Press, Oxford, UK.

Beran V., Dlask P. 2005: *Management udržitelného rozvoje regionů, sídel a obcí*. Acad. Praha, Praha, ČR.

Briassoulis H. 2000: *Land Use Change: Theoretical and Modelling Approaches* [online]. Regional Research Inst., West Virginia Univ., US: [cit. 2010-04-12]. Accessible from WWW: http://www.rri.wvu.edu/WebBook/Briassoulis/contents.htm

CZSO (Czech Statistical Office), 2010: http://www.czso.cz

EPA (Environmental Protection Agency), 2000: *Projecting Land-Use Change : A Summary of Models for Assessing the Effects of Community Growth and Change on Land-Use Patterns*. Environmental Protection Agency (EPA), US.

Forrester J.W. 1969: *Urban Dynamics*. The MIT Press, Cambridge, US.

Grill S., Vorel J., Maier K. 2008: Urban *development simulation and evaluation*. [In:] *Sborník konference GIS Ostrava 2008*, VŠB-TU Ostrava, Ostrava.

Henderson J. V., Thisse J.-F. 2004: *Handbook of Regional and Urban Economics*, Vol. 4, *Cities and Geography*. Elsevier, Amsterdam, The Netherlands.

Jasper van V.; 2006: *Validation of Land Use Change Models: A Case study on the Environmental Explorer*. The Netherlands, Wageningen Univ.

Maier K., Vorel J., Ctyroky J. 2007: *Comprehensive land-use allocation model*. [In:] *Sustainable development of regions*. CTU, Faculty of Civil Engineering, Prague, Prague, 73.

Maier K., Vorel J., Ctyroky J., 2007: *Simulation model for urban development sustainability appraisal*. [In:] M. SCHRENK et al., *CORP 2007 Proceedings*. Eigenverlag des Vereins CORP – Competence Center of Urban and Regional Planning, Vienna, Austria, 536.

Nijkamp P., 1986: *Handbook of Regional and Urban Economics*. Vol. 1. Regional Economics, Elsevier.

RIKS (Research Institute for Knowledge Systems) 2005: *MCK Reader: Methods of the Map Comparison Kit*.

Vorel J., Maier K., 2007: *Learning the public preferences for living environment characteristics: the experimental approach*. [In:] PAPERS OF REAL CORP 007, Eigenverlag des Vereins CORP – Competence Center of Urban and Regional Planning, Vienna, Austria, 3. Accessible from WWW: http://programm.corp.at/cdrom2007/archiv/papers2007/corp2007_VOREL.pdf

Vorel J., Maier K., Grill S., 2007: *Urban simulation: Decoding alternative futures*. [In:] PAPERS OF REAL CORP 008, Eigenverlag des Vereins CORP – Competence Center of Urban and Regional Planning, Vienna, Austria. Accessible from WWW:http://programm.corp.at/cdrom2007/archiv/papers2007/corp2007_VOREL.pdf

Waddell P., 2002: *UrbanSim: Modelling Urban Development for Land Use, Transportation and Environmental Planning*. Univ. of Washington, US.

Wilson A.G., 1974: *Urban and Regional Models in Geography and Planning*. John Wiley, New York, US.

OPARTY NA SCENARIUSZU MODEL SYMULACJI ZAGOSPODAROWANIA PRZESTRZENNEGO

Streszczenie

W niniejszym artykule przedstawiamy zastosowanie koncepcji teoretycznych oraz metodologii modelowania zagospodarowania przestrzennego w oparciu o przewidywany przyszły rozwój danego terytorium. Postanowiliśmy powtórzyć dynamikę urbanizacji średniej wielkości regionu miejskiego w formie modelu symulacyjnego oraz przewidzieć alternatywne scenariusze przyszłych wzorców zagospodarowania przestrzennego. Pokrótce omawiamy też koncepcję, wdrożenie i zatwierdzenie modelu symulacyjnego. Wskaźniki wymiarów Kappa, Fuzzy Kappa i Fraktal wskazują na powiązanie modelu symulacyjnego z alternatywnym modelem losowym oraz obserwacją faktycznych wzorców zagospodarowania przestrzennego. O ile model symulacyjny sprawdza się lepiej od alternatywnego modelu losowego w przypadku budownictwa rodzinnego i wypoczynkowego, tworzone w jego ramach prognozy nadal mają słaby związek z budownictwem wielorodzinnym oraz regionalnymi punktami sprzedaży detalicznej w przypadku obserwowanego zagospodarowania przestrzennego. W artykule przedstawiamy przypuszczalne powody nieścisłości prognoz oraz proponujemy kilka możliwości usprawnienia modelu.