reviewed paper

Urban simulations: Decoding alternative futures

Jakub VOREL, Karel MAIER, Stanislav GRILL

(Ing. arch. Jakub VOREL, PhD. Czech Technical University in Prague, Faculty of Architecture, Thákurova 7, 166 34, Praha, jakub.vorel@sklep.fa.cvut.cz)

(Prof. Ing. arch. Karel MAIER, Czech Technical University in Prague, Faculty of Architecture, Thákurova 7, 166 34, Praha, maier@fa.cvut.cz)

(Mgr. Stanislav Grill, Charles University in Prague, Faculty of Science, Albertov 6, 128 43, Praha, grillst@natur.cuni.cz)

1 ABSTRACT

Local land-use planning significantly influences environmental, social and economic aspects of future territorial development but formal procedure of impact assessment of planning regulations has been just recently integrated in local plans as their obligatory part. Information technologies are essential in this respect for effective and efficient collecting, sharing and providing the relevant information. But, however large amount of data we may have, their practical application in the assessment is often biased by limited human capacity. Here the use of urban modelling has not been fully appreciated yet: existing urban models are often perceived as a hobby for academics, not being comprehensive, precise and reliable enough to provide with guidance for real-world decisions.

The paper presents an approach to the modelling of urban change on the municipality level that would help the professionals to predict the impact of local plans by development of scenarios on the base of policy options. The model would demonstrate the outcomes so that also lay users may assess their choices.

The main focus of the paper is to describe factors that are reflected in the model and their contribution to the model validity. Both endogenous and exogenous factors are dealt with and the role of the user-centred knowledge in the interaction with the model is discussed. Various measures of the model's „Goodness of fit“ are presented. Modelling qualitative environmental aspects and effort for integration of the normative and descriptive knowledge are addressed in the paper.

2 OVERVIEW OF PREVIOUS RESEARCH BACKGROUND

Previous research set up the conceptual scheme of the modelling tool that could be used to define, predict and assess ex-ante sustainability of plans for development of an urban territory (Grill, 2007; Drda, 2007). The tool is intended to present alternative scenarios of future development under various conditions (amount of public investment, land-use control etc.) and in this way to offer a mutual comparison of scenarios.

The tool is based on the model simulating the spatial processes in the territory. The model predicts the future land-uses taking into consideration the functional compatibilities of cells, access to infrastructures and physical conditions of sites. The spatial relations are modelled using the two-dimensional grid of cells. Each cell is an autonomous decision unit (agent) with fixed position and fixed number of neighbouring agents. Each agent tends to maximize own satisfaction related to its preferences by maximizing the compatibility with the neighbouring land-use activities and by choosing the land-use with the highest possible utility from the accessibility to the infrastructure. The potential transition to more satisfying land-use must be achieved at minimum cost.

The transition of the land-use starts from the cell that shows the best satisfaction/cost ratio through the cell with the second best ratio, as long as exogenous demand for the new use is not satisfied or endogenous limits such as accessible public infrastructure, available resources or developable land is not faced.

The model distinguishes two stages of the cell development. In the first stage the agent decides among three alternatives of land-use: residential, manufacturing or undeveloped cell. In the second stage the population of the area is considered. When the number of inhabitants overpasses a certain threshold level, the land-use is completed by public or commercial services. Basic commercial services integrate in residential uses within single cell while the higher-rank commercial and public services locate in particular cell separately from other uses. The services tend to choose the locality (cell) of best accessibility from all residential cells in the area.

As in all models, also this model is based on several simplified assumptions. Firstly, it assumes that the preference structure of all agents in the territory is homogenous. Secondly, the agents are assumed do not to coordinate their present decisions, all the decision-making of individual agents being influenced only by past decisions of the other agents. We assume that the simplifications do not impact seriously on the simulated processes and that they can be relaxed in future generations of the model, however on the account of increasing complexity of the model.

The main task of the model is to localize the residences and servicing infrastructures. Land is considered mainly as consumption commodity and not as production factor. The production (industry) and residential land-uses use the same factor for their localization, even though with different sensitivity for each factor, while some important factors for industry localization are neglected as for example accessibility to labour force or proximity to natural resources and customers.

The user decides upon the exogenous factors of the model: the amount of public investment in infrastructure; flexibility or toughness of zoning regulations, overall demand for housing, office and manufacture development.

The mutual compatibility of land-uses was introduced normatively. It is based on the following "common sense" assumptions:

- the concentration of mixed uses in central district is caused by the agglomeration economy and other effects of spatial concentration;
- basic services and shops are located so that the inhabitants should be provided with reasonably comfortable access;
- residential uses benefit from neighbourhood of green areas and open landscape;
- mixed uses of central districts are ambivalent to green areas. On one hand the green areas make the public space attractive, on the other hand green areas occupy valuable space;
- various types of green areas in the open landscape are neutral each to other;
- monofunctional areals diminish the attractiveness of neighbouring residential zones;
- monofunctional areals in central districts are pushed out by the zones of mixed uses that benefit from the mutual compatibility and attraction for the customers;
- hotel services and offices concentrate a large number of people who in turn generate demand for higher-order services located in central district;
- mixed and residential use is not compatible with technical utility areas;
- manufacturing has neutral or positive relations with regards to technical infrastructure;

Presented assumptions were operationalized in the form of matrix that presents the compatibility of all combinations of neighbouring land-uses on the scale from fully compatible (+10) to fully incompatible (-10). Similarly the utility of accessibility to services are assessed for each land-use. Even though the arbitrarily stated relations among the land-uses seemed to be intuitively right they did not produce any reasonable land-use pattern predictions.

3 METHOD OF MODEL CALIBRATION

The previous research established the basic principles of the model but it did not create the operational form of model. The ongoing research focuses on the model calibration and validity aspects. First task is to empirically justify the relations among particular land-uses using the data on real land-use changes.

To decide on the best possible way of the model calibration, an overview of similar models was made. Out of several of them (UrbanSim, CUF I, CURBA, SLEUTH), the California Urban Futures II model (CUF II) was identified as the most relevant response to our problem (LANDIS, 2001; CHENG, 2003; EPA, 2003).

The CUF II is the land-use allocation model which recognizes four types of land-use: a) undeveloped land, b) residential (single family or multifamily), c) commercial (retail and offices) and d) industrial. The land uses are assigned to regular cells of 1 hectare area (in our model it is 0,56 ha respectively 0,12 ha). The CUF II allocates population according to probabilities of land-use change. The probabilities derive from the land-uses changes realized during certain period (e.g. 1985 to 1995). Each particular probability of land-use change depends on probabilities of other land-uses; therefore the probabilities can be used to simulate bidding mechanism of one land-use the others.

By using the multinomial logit model, observed dependent variables (land-use changes) are related to the series of independent variables (factors) that are assumed to influence the land-use change. CUF II identifies following land-use changes as dependent variable (LANDIS, 2001):

- undeveloped → single family residential
- undeveloped → multifamily residential
- undeveloped → commercial (retail and office)
- undeveloped → industrial
- no change in undeveloped status
- non-residential → residential
- non-commercial → commercial
- non-industrial → industrial
- no change in previously developed status

The CUF II model includes several independent variables that represent the factors influencing the land-use change: a) initial land uses, b) demand factors that are measured on city level by the rate of household and job growth in the previous five years, c) regional accessibility measured by the distance of nearest city centre from the site of land-use change, d) the distance to the nearest freeway or transit station, e) the slope of terrain representing the most significant physical constraints, f) the accessibility of infrastructure measured as distance of the site to the border of urbanized area that is served by infrastructure, g) the impact of adjacent land-uses and h) inter-use externalities and proximity effects measured by Euclidean distance from each site to the nearest commercial, industrial or public-use site (LANDIS, 2001).

For our model calibration, the Multinominal Logistic Regression (MNLR) was used that establishes the relationship between single dependent variable *y* with several categories *j* and several independent variables *x* that can be categorical or continuous. The regression model measures the probability that the dependent variable *y* (land-use change) belong to particular category *j* (type of land-use change) with regards to certain combination of independent variables (LIAO, 1994).

$$P(y=j)=\frac{e^{\sum_{k=1}^{K}\beta_{jk}x_k}}{1+\sum_{j=1}^{J-1}e^{\sum_{k=1}^{K}\beta_{jk}x_k}}$$

There can be several categories of dependent variable (thereby multinominal logit) $j = 1, 2, 3, \dots, J-1$. The last category J is the reference category. The reference category is calculated by the following formula (LIAO, 1994):

$$P(y=J)=\frac{1}{1+\sum_{j=1}^{J-1}e^{\sum_{k=1}^{K}\beta_{jk}x_k}}$$

To calculate the probabilities, first the β_j parameters must be estimated by using the maximum-likelihood estimation technique. It is necessary to estimate the β_j parameters for each k-th independent variable x and j-th dependent (response) variable j.

The β_j parameters (logits) are not directly interpretable. For beter interpretation they have to be transformed to power of natural logarithm e^{β} that represents the change in the odds of the modeled category of dependent variable that is caused by one-unit change in the independent variable x_k. When $\beta_j > 0$ ($e^{\beta_j} > 1$) then the dependent category will more probably belong to the category j, when $\beta_j < 0$ ($e^{\beta_j} < 1$) than the opposite is true. When $\beta_j = 0$ ($e^{\beta_j} = 1$) then the independent variable does not have any impact on the behaviour of the dependent variable.

4 DATA MINING

The objective of the data collection was to get the data that would document the urban change in certain historical period. The most difficult part was to find the right data for each factor that is considered significant from the researchers' point of view. Data on historical land-use development were unavailable in the form that would enable direct computer processing. Therefore the register of buildings (RB) maintained by the Czech Statistical Office (CZSO, 2008) was used to derive land-uses. Digital Topographical Maps in the scale 1:200 000 (DMÚ 200) were used to derive the location of road and railway network, the polygons of built-up areas and of basic categories of land-use. Digital Elevation Model (DEM) that is part of the ZABAGED geographic database was used to derive the slope of terrain.

4.1.1 Data on buildings

The RB provides with location, use, age and number of floors of buildings. The attribute of "building age" sorts the time of building constructions into several time periods: before 1899, 1900 – 1919, 1920 – 1945, 1946 – 1960, 1961 – 1970, 1971 – 1980, 1981 – 1990, 1991 – 1995, 1996 and later. About 30% of data are missing: for example the data on construction period are indicated as missing, unknown or not defined in more than 50% of cases. To complete the missing values on housing age, the house numbers in each part of municipality were used to derive the time of construction assuming that house numbers form the time series.

There are no data the age of buildings for individual recreation at all. Therefore the sequence of buildings numbers for temporary buildings is used to attribute age to buildings equally between 1961 and 1990. We assume that this was the period when the buildings for individual recreation typically emerged.

For the use of building two classifications were used. First the standard classification ISKN 76 having 16 categories was used. The ISKN categories were reclassified into the categories that are used by model:

category used in model	description of land-use
BI	family houses (single or two dwelling units per house)
BH	multifamily housing
RI	individual recreation
VV	public services
KV	commercial use (administration, shops, hotels and other services)
D	transportation
T	technical utilities
P	industry and warehousing
N	non-developed land (vacant land, agriculture and forest)
OTHER	non-specified use

After processing the reclassification, still about 20% of all buildings were identified as having "non-specified use". Therefore the non-classified buildings were sorted according to types of construction using the robust classification KSD 5522 that use 408 types of buildings. The cross tabulation of classifications ISKN 76 and KSD 5522 proved their mutual consistency. By using both classifications the number of non-specified cases was reduced from 20% to 10%. Even though it is clear that most remaining "non-specified uses" are temporary buildings for individual recreation, they are further kept under category "non-specified".

Having the temporal dimension of building location it was possible to establish several frames documenting the location of buildings existing in the year 1961, 1981, 1991, 1996 and 2008.

4.1.2 Deriving land-uses

The cell land-use is 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. Some exceptions are made in the case of buildings used for manufacturing industry (P) and public services (VV). Industry sites and public services such as schools, hospitals and public administration tend to form a monofunctional areals, therefore every single instance of such building automatically determines the use of the cell without consideration of other buildings in the same cell.

The cells are artificial construct. They reflect neither physical nor economical nor legal reality (LANDIS, 2001). Their main function is to abstract from complex reality and to reflect only selected characteristics, in our case the land-uses. It was expected that the size of cells influence the translation of building uses into the cell land-use. Two cell sizes: the 35x35 m and 75x75 m were tested. The reduction of the cell size reduces the ability of the model to capture mixed land-uses (S). Number of other cell land-uses is growing proportionally to the reduction of the cells' area. The ability of each size of cells to capture various types of land-uses can be seen in the table that presents total number of cells according to specific land-uses and cell sizes as tested on the pilot case of a small Czech town.

cell's land-use 2008	number of cells 75x75 m	number of cells 35x35 m
BH	112	281
BI	5439	11765
RI	2450	5402
S (mixed land-uses)	226	81
KV	11	30
VV	83	237
P	1295	4548
D	4	5
T	8	15
N	105239	508210
OTHER	1053	2226

Some aspects of the regular cells application are not still resolved. The expectation that the 35m cells would more realistically represent the average size of plots in residential areas was not proved. The 35m cells covered only the building plot and not the gardens that are in reality attached to the building plot. In this case the cell size proved to be too small for capturing the relations between the buildings and their plots.

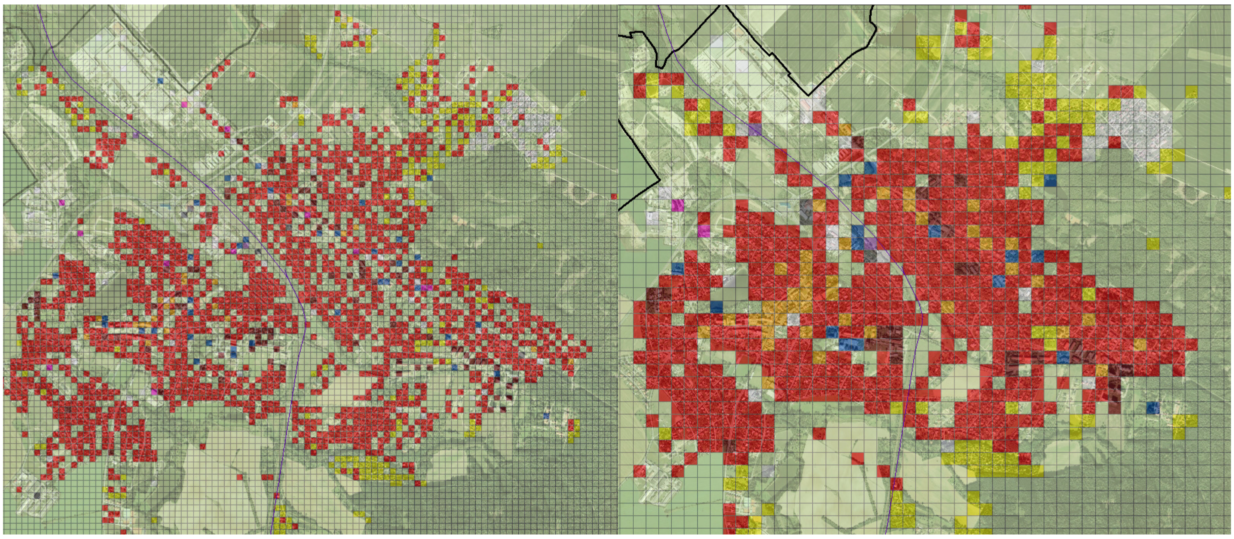

Fig. 1: The comparison of 35m and 75m cell grid

The problem can be clearly illustrated on the industrial and public buildings that are usually situated in plots of big size. In this case the grid of cells captures only immediate surrounding of building. Provisionally the aerial pictures were used to complete the information on the use of big areals.

4.1.3 Identifying land-use changes

The objective is to create a model that would describe the past land-use changes. Such model could help us to extrapolate the past trajectory into future development and to predict the future land-uses. To obtain such model it is necessary to identify the cells that were in certain period affected by land-use change. The table below shows the frequencies of particular land-use changes in the time period between 1961 and 2008.

cell's land-use change 1961 - 2008	number of cells 75x75 m	number of cells 35x35 m
N → BH	80	198
N → BI	2373	6397
N → RI	2347	5379
N → S	31	19
N → KV	15	32
N → VV	68	147
N → P	671	2571
N → D	0	0
N → T	8	14
BH → OTHER	14	5
N → OTHER	997	2136
NO CHANGE	109336	515902

The table indicates that frequencies of land-use changes differ a lot. Surprisingly, no land-use changes were identified in the built-up area. To correct the deficiency another source of data on land-use will be needed that would complete the data on use of buildings.

The positive aspect of the cell size reduction is that some less frequent land-use changes emerged as statistically significant (BH, KV). On the other hand one cannot expect that increased sample size would increase the variability of factors and eliminate possible multicollinearity.

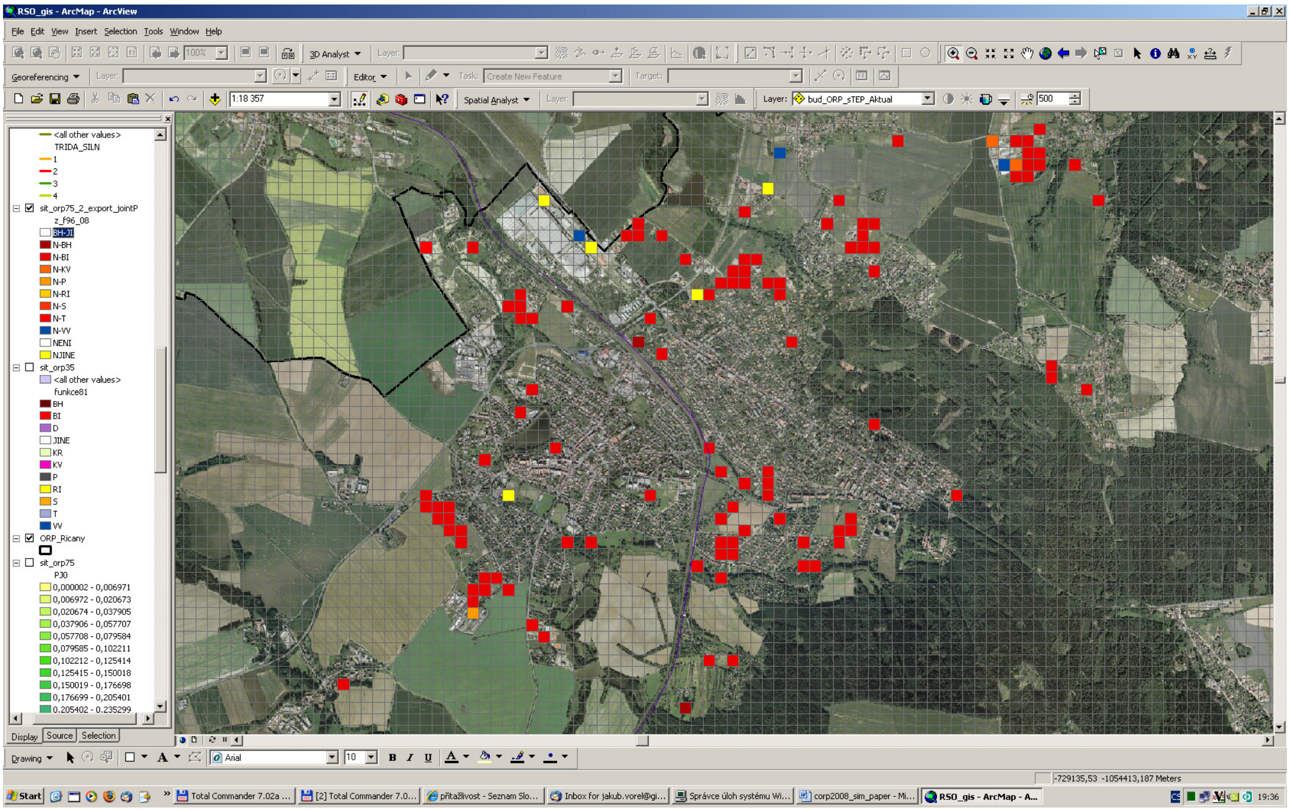

Fig. 2: Land-use changes realized between 1996 and 2007:
red: N → BI, brown: N → BH, orange: N → S, blue N → VV, yellow N→ OTHER

4.1.4 Data on land-use adjacency

The mutual effects of neighbouring land-uses are expected to play important role in the land-use localization. For the inclusion of the effects into the model it is necessary for each cell to identify its neighbours.

Series of statistical tests proved that the most relevant size of neighbourhood is 9 x 9 cells. That means that not only directly adjacent land-uses, but also the next but direct land-uses influence the land-use of cell. The strength of the effect depends not only of the presence of absence of specific land-use category, but also on the proportion of land-uses in the neighbourhood (VOREL, 2006). In our case the proportion of land-uses proved not being significant, therefore only binominal indicators were used to indicate whether at least one cell of specific land-use category is present in the neighbourhood. This corresponds to the approach of CUF II model that test whether there is a specific land-use in distance of 200 meters (LANDIS, 2001).

4.2 Data on infrastructure accessibility

The decision of each cell on its land-use change is influenced also by its location with regards to public infrastructure. The proximity to public services, important public spaces and transport and technical infrastructure has different impact on different land-uses.

The previous research the general public infrastructure was listed. For the use in the model only the public infrastructure that is present in the pilot case and that proved to have significant impact on the localization of land-uses was selected.

public infrastructure	applicability in the model
train station of high importance (international train links)	not in the area
train stations of regional importance	not significant
bus stations	not significant
underground lines	not in the area
tram, trolleybus	not in the area
bus stops with interval less than 15 minutes	not significant
bus stops with interval less than 60 minutes	not significant
motorway slip-roads	significant
important urban street	not significant
I., II. and III. class roads	significant
police station	not significant
fire department	not significant
zones of infrastructure provision	significant
railway lines	significant
basic schools and kinder gardens	not significant

There can be several reasons why some public infrastructure has insignificant impact on localization of land-uses. First reason results from the method of statistical regression. It is important that the independent variables that enter the regression are not correlated. The usual co-location of several public infrastructures can cause that at least some of them would become insignificant. For example the co-location of bus stops and the road network can cause that the bus stops are evaluated as insignificant for the localization of housing land-uses.

Another reason of insignificance is that the spatial relation between specific types of public infrastructure and land-uses is weaker than it was expected. The pilot area is intensively urbanized and the public infrastructure is generally well accessible. The dependence of habitants on public transportation and pedestrian accessibility to public services is diminished by the lifestyle that is car oriented. Therefore at present the factors of accessibility play less important role for the localization of housing land-uses than for example legal limits or land-use plan regulations.

The accessibility of technical infrastructure for newly developed sides was included into the model by demarcation of zones of infrastructure provision (in CUF II model this is called „sphere of influence“). The zones cover the built-up areas that are served by technical infrastructure (water conducts, sewage system, electricity and gas distribution). We assume that newly located land-uses should take advantage of already installed public utilities and therefore the land development near the borders of built-up areas should be preferred to the land development that is taking part further from the outside borders of built-up areas.

4.2.1 Data on physical conditions

Several limiting factors for specific land-use exists: the morphology of terrain, type of soil and underground water conditions. With regards to the accessibility of data, only slope of terrain was included into the model. Significance of other factors will be tested in later phases. The CUF II model uses five intervals of slope: 1-2%, 3-5%, 6-9%, 10-15%, 16% and more. In our case only three intervals are used: 1-6%, 6-12%, 12% and more.

ISBN: 978-39502139-4-2 (CD-ROM); ISBN: 978-39502139-5-9 (Print)
Editors: Manfred SCHRENK, Vasily V. POPOVICH, Dirk ENGELKE, Pietro ELISEI

5 FACTORS KEPT EXTERNAL TO THE MODEL

Some factors are not included as independent variables into the regression model and remain only as external to the model. The main reason is that the data on these factors may not be available in this moment. For example the time series of the demand for land-uses is missing. Also historical data on jobs is difficult to obtain.

We can compare the CUF II model with our model in this respect. In addition to our model, the CUF II model includes the demand for space into regression model as three independent variables: The rate of household and job growth during the period 1980-1985 and job/household ratio (LANDIS, 2001). In our model demand does not enter the regression model and therefore it does not influence the probability of land-use changes. Demand enters the allocation process as an amount of each function distributed according to already counted probabilities.

Another factor that is remaining external to the regression model is legal limits to land-use change. While legal limits to the land-use development were included into the regression model in the CUF II, in our case they are external to the regression model. As the cells affected by legal limits are not developable at all (probability of development equals zero), the regression model would not provide us with any new information.

The regression model does not take into consideration the regulations of local land-use plans. Even though it is decisive factor in the determining of land-uses, it is too complex to be traced through history. Similar problem was identified by the authors of the CUF II model. The land-use plans were changing very often and it is difficult to trace the history of the changes of regulations and their influences on realized land-use change (LANDIS, 2001). Our model does not include the land-use regulations and spatial limits in the land-use change probabilities, but it use them only in land-use allocation process.

6 THE MODEL CALIBRATION

The Multinominal Logit Regression (MNLR) was used to establish the probabilistic relation among the dependent and independent variables. Both the continuous and categorical variables are entering the model. There are 12 independent variables entering the regression model, as we can see in the table below.

name of independent variable	type of variable, units
distance to motorway slip-roads	continuous, distance unit = 100 meters
distance to road class I, II and III	continuous, distance unit = 100 meters
distance to zones of infrastructure provision	continuous, distance unit = 100 meters, inside of zone distance = 0 meters
distance to railway lines	continuous, distance unit = 100 meters
slope of terrain	categorical: 0%–6%, 6%–12%, > 12%,
BI, KR, KV, P, RI, S, VV in neighbourhood	categorical, is or is not in the neighbourhood

The purpose of the model is to predict the localization of residential and industry land-uses that correspond with the most frequent land-use changes from non-developed land to the individual housing, buildings for individual recreation and industry land-uses.

the category of land-use change	description
N → BI	non-developed land to individual housing
N → RI	non-developed land to individual recreation
N → P	non-developed land to industry

Following tables present the e^{β} parameters for each value of independent variable x_k. Each independent variable has two coefficients that are related to two categories of dependent variable y_1 (N → BI) or y_2 (N → P). Third category of dependent variable y_3 (N → RI) serves as reference category.

The independent variables are sorted into three tables. The first table presents continuous variables. The interpretation of e^{β} parameters is straightforward: e^{β} parameters represent the contribution of independent variables to the probabilities of specific category of dependent variable y_j (N → BI or N → P). For example moving 100 meters from a road means 0,995 times smaller probability of locating the BI on N then locating the reference category RI on N. Moving the cell to the distance of 200 meters would equal to the decrease of relative probability to $0{,}995^2 = 0{,}990$, etc.

$x_{1..}\ x_4$ independent variable values	$e^{ß}$ parameters for y_1 (N→BI)	$e^{ß}$ parameters for y_2 (N→P)
distance to motorway slip-roads	1,006*	0,977
distance to road class I, II and III	0,995	0,954*
distance to zones of infrastructure provision	1,025*	1,101
distance to railway lines	0,995*	1,001

The second table presents multinominal variable x_5, in which three categories represent slope of terrain. The table presents coefficients only for two categories, because the third category (slope of terrain more than 12%) is reference value, with $e^{ß} = 1$. Parameters $e^{ß}$ describe how many times one category of independent variable x_k increase or decrease the probability that the dependent variable y_j will change to a specific category (N → BI or N → P) with regards to the reference category. For example choice of the cell with the slope of terrain smaller than 6% increase the relative probability of an industry development 1,969 times in comparison to the choice of the cell having the slope bigger then 12%

x_5 independent variable values	$e^{ß}$ parameters for y_1 (N→BI)	$e^{ß}$ parameters for y_2 (N→P)
slope of terrain between 0% and 6%	0,883	1,969
slope of terrain between 6% and 12%	0,974	1,342

The third table presents series of binominal independent variables $x_{6...}x_{12}$. They describe the instance of specific land-uses in the neighbourhood of the cell. Each variable has one reference category representing absence of cell with specific land-use in the neighbourhood. The reference category has $e^{ß} = 1$. The interpretation of coefficients is similar to the previous variable. For example a cell with individual housing (BI) in the neighbourhood increases 17 times the probability that another cell in the neighbourhood changes from non-developed land to residential (BI). On the other hand having the BI in neighbourhood would decrease 0,063 times the probability that the same cell would change into industrial (P).

$x_{6..}x_{12}$ independent variable values	$e^{ß}$ parameters for y_1 (N→BI)	$e^{ß}$ parameters for y_2 (N→P)
BI in neighbourhood	16,855*	0,063*
N in neighbourhood	1	1
KV in neighbourhood	1,256*	0,895
P in neighbourhood	1,212	1,806
RI in neighbourhood	0,028*	0,00013*
S in neighbourhood	4,457*	2,685*
VV in neighbourhood	2,739*	5,5E-9*

The $e^{ß}$ parameters marked by * are statistically significant, other parameters being insignificant. It was decided not to leave the insignificant parameters out of model for this moment. We are aware that the inclusion of insignificant parameters into model can decrease the performance of the model when applied on other data, therefore further research will be aimed at improvement of operationalization of factors and inclusion of new factors into model.

The authors of the CUF II model faced similar problems with factor significance. Not all factors proved to be significant in all eight counties in which the CUF II model was applied. Among the most significant factors in all the counties was distance to city limits (in six cases), slope (in five cases), distance to major highway (four cases).

The model is calibrated to describe the urban development of the pilot area in the time period between 1961 and 2008. The values of parameters therefore do not describe all historically developed land-uses. Therefore it may seem that some parameters as for example „distance to road class I, II and III" should be much more distinctive (eccentric).

The ultimate goal of the modelling effort is to obtain relative probabilities of dependent variables for every cell taking into account all independent variables. The result is presented in figure 3 bellow.

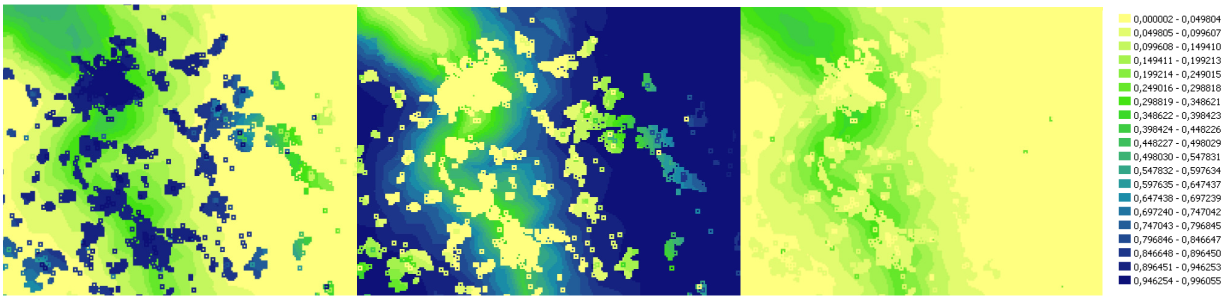

Fig. 3: Calculated cell probabilities of the new family housing (BI), industry (P) and individual recreation (RI) land-use localisation.

7 ALLOCATION OF LAND-USES ACCORDING TO LAND-USE CHANGE PROBABILITIES

The calibrated model can be used to predict the land-use changes based on the calculated probabilities. As the probabilities of land-use change are comparable across all the modelled area and all types of land-uses, they can be directly used for the allocation of demanded land-uses. The relationship between probabilities and choice of land-use is deterministic: in the auction the specific land-uses are allocated to undeveloped cells according to the highest probabilities. The land-use changes are realized as long as there is the unsatisfied demand for specific land uses.

Fig. 4: Predicted pattern of land-use changes between 1961 and 2007

The models' "Goodness of fit" measures the percentage of correctly predicted land-use changes. The following table shows the results:

	Predicted			
Observed	N→BI	N→P	N→RI	Percent Correct
N→BI	779	20	418	64,0%
N→P	37	446	1	92.1%
N→RI	56	9	884	93,2%
Overall Percentage	32,9%	17,9%	49,2%	79,6%

The table proves that the model is successful in 79,6% of all cases. It is less than in the case of CUF II model with 92% of successful prediction. Also the success of prediction differs a lot among the land-use changes. The prediction of new industrial and individual recreation land-use location is equally successful as in the case of CUF II model. Surprisingly the model is less successful in prediction of new individual housing location (N→BI). In this case the CUF II model was successful in 21-48% depending on application in each county. While the CUF II model uses „no land-use change" category for maintaining status-quo of the cell use. In the case of our model that is calibrated on the realized land-use changes only.

8 CONCLUSION

The presented approach to urban simulation proved to be reasonably successful in the predicting of future development of land-uses in suburban and exurban areas. But there is still long way to go to get the results that are persuasive not only for academics, but also for practitioners who have direct responsibility for their decisions. To improve the "Goodness of fit" of the model, the following tasks must be fulfilled:

- the validity of the model will be tested while applied to other cases that have different characteristics;
- the effort will be focused on additional data collection to include other possibly significant factors;
- other data sources and methods will be used for more precisely identification of land-uses;
- the temporal strength and significance of each factor will be tested by developing the regression model for different time periods;
- new indicators of models' "Goodness of fit" that are based on spatial statistics will be tested.

9 REFERENCES

CZSO (Czech Statistical Office), http://www.czso.cz/eng/redakce.nsf/i/home, 2008
GRILL S., VOREL J., MAIER K.: Urban development simulation and evaluation, GIS Ostrava, 2008
DRDA F., MAIER K., VOREL J., ČTYROKÝ J.: Simulation model for urban development sustainability appraisal, CORP 2007 Proceedings
EPA: Projecting Land-Use Change: A Summary of Models for Assesing the Effects of Community Growth and Change on Land-Use Patterns: Environmental Protection Agency: 2000
CHENG J.: Modelling Spatial & Temporal Urban Growth: PhD Thesis: Faculty of Geographical Sciences, Utrecht University, The Netherlands: 2003: ISBN 90-6164-212-4
LANDIS J.: CUF, CUF II, and CURBA: „A Family of Spatially Explicit Urban Growth and Land-Use Policy Simulation Models“ in Brail, Richard. K.; Klosterman, Richard E.: Planning Support Systems, ESRI, 2001
LIAO T.: Interpreting probability models, Logit, Probit and Other Generalized Linear Models: Series: Quantitative Applications in the Social Sciences: Sage Publications, 1994
VOREL J.: Learning the public preferences for habitat environment: the use of Conjoint analysis: 5th International Conference on Advanced Engineering Design proceedings ISBN number 80-86059-44-8, June 2006, Prague