A GIS-based method for the selection of the location of residence

Xavier ALBACETE*, Kari PASANEN and Mikko KOLEHMAINEN

Department of Environmental Science, Research Group of Environmental Informatics, University of Eastern Finland, Yliopistoranta 1 E, P.O. Box 1627, 70211 Kuopio, Finland

(Received 18 February 2011; final version received 2 January 2012)

Homebuyers are usually interested in both the accessibility of services and the quality of the local environment, and real estate agents frequently offer some web-based systems for home searches. There is however hardly any information about the quality of local living environment in those web-based systems. The purpose of this study was to develop a method for homebuyers, adaptable to the environmental variables of interest to homebuyers when selecting a home location. In this paper, a multicriteria spatial analysis method is proposed and demonstrated for the homebuyers’ selection process, using data from the City of Kuopio, Finland. Several spatial variables are applied, including environmental and service factors in the home searching process. A geographical information system (GIS) is used for creating maps for decision variables and mapping suitable areas. The method for ranking alternative dwellings is based on the difference between levels of the decision variables for each dwelling and the target levels given by the user. The method presented in this paper is adaptable to other geographical and social contexts. This decision analysis tool will be useful for both customers and real estate agents, and can also be used for city planning as a participatory-GIS (P-GIS) tool. It introduces new possibilities in the home selection process. The availability of spatial data on the living environment in the web-based services for homebuyers is likely to have effects on customers’ requirements and house markets, and also promote better spatial city organization in the long run.

Keywords: home selection; real estate market; living environment; GIS

1. Introduction

Several studies have focused on the spatial analysis of urban areas for homebuyers. Some of these studies focus on the spatial evaluation for the homebuyer’s selection (1, 2), whereas others relate the living environment to the quality of life (3).

However, the relation between home locations and individual perceptions in the analysis and evaluation of urban areas proved to be a key issue (3), since consideration of amenities such as climate, environmental, and urban conditions is critical when analyzing subjective well-being (e.g. proximity to landfill, cost, hazardous waste facility, rail station, and airport). Additionally, in some studies, local opinion concerning specific urban aspects, such as air quality and the amount of green space, has been considered when proposing development and policies (4–7). Nevertheless, none of the studies examined the relation between the location of residence and the perceptions of the living environment.

Geographical information systems (GIS) has been widely applied in studies on urban environment and spatial decision analysis (2, 7–11). The main focus in the scientific discussion has been on the methodology used to quantify and integrate the variables involved in the urban spatial analyses (5, 12–14). However, the proposed decision support tools have been developed mostly for city planners rather than for ordinary people.

In order to fill this gap, Natividade-Jesus et al. (1) proposed a decision support system (DSS) for housing evaluation, applying up to 210 decision variables, including also many spatial aspects of the living environment. The system is suitable for home selection analysis, where the attributes of interest, weights, and thresholds are given by the user. Additionally, the houses were ranked and visualized on maps. A second study shows how various spatial and nonspatial criteria and multicriteria evaluation tools in spatial DSS for homebuyers were also applied (2). Variables based on location, proximity, and direction were applied for spatial criteria, and the simple additive weighting, a spatially adjusted method was applied for the calculation of final scores for alternative homes that were ranked and visualized on the map. Nevertheless, in both cases the selection of the suitable areas was limited to neighborhoods. Additionally, the need for nonlinear scaling of the distance-based criteria was highlighted.

Additionally, when looking at the current possibilities within Finland, the options for the selection of a place of residence are still quite limited in home search portals. In most cases, the user can only focus on the

*Corresponding author. Email: xavier.albacete@uef.fi

ISSN 1009-5020 print/ISSN 1993-5153 online
© 2012 Wuhan University
http://dx.doi.org/10.1080/10095020.2012.708159
http://www.tandfonline.com
characteristics of the dwelling and define an area for making the search. Some portals show the closest services for selected dwellings, and some information on the local environment is provided on web-map services, which apply basic maps, aerial images, and street view photos (e.g. Jokakoti.fi, Huoneistokeskus.fi).

Thus, there is a need for new spatial descriptive environmental variables and simple decision analysis tools for the home selection process, which is manageable by ordinary citizens. In this paper, we present a spatial analysis method for homebuyers with a city analysis extension and nonlinear scaling of the distance-based criteria. The method uses several spatial variables including environmental and service factors (e.g. noise level, children’s playgrounds, sports centers, public health facilities, connectivity, and access to nature) for ranking alternative dwellings and mapping suitable areas.

2. Materials and methods

2.1. Overview of the method

The method presented in this paper for home selection has the following main steps:

1. Defining the spatial extent for the analysis.
2. Creation of raster maps for the decision variables.
3. Classification of the decision variables into five levels.
4. Locating dwellings alternatives.
5. Spatial linking of the decision variables and home locations (Table 1).
6. Selection of decision variables and target levels by the user.
7. Mapping suitable areas using GIS.
8. Ranking of the dwellings.
9. Visualizing suitable areas and ranked dwellings.

2.2. Creation of decision variables

Spatial data and the proposed decision variables for the method are presented in Table 1. The data include road network, location of services (schools, children’s playgrounds, children’s daycare centers, centers of neighborhoods, libraries, sports centers, groceries, and public health centers), users’ points of interest (e.g. location of relatives and place of work), location of dwellings for sale, water areas, and forest (25 × 25 m raster for timber volume). Modeled noise propagation maps and dispersion maps of particulate matter (PM$_{10}$) concentrations are the most important environmental data often available for urban areas. All data can be obtained from the National Land Survey of Finland, the Finnish Transport Agency, the Finnish Forest Research Institute, the Finnish Meteorological Institute, and municipalities.

The nature variable was constructed as a combination of distance to dense forest areas (forest areas with wood density ≥2500 m$^2$ in 2000 m radius) and to natural water areas. In addition, the purpose of the variable centers of neighborhoods was to express the proximity to social and commercial centers located in every neighborhood of the city. These centers are potential spaces for developing social networks in Finnish cities. The proximity to special interest points was conceived as an option to be applied in web-based services, where the customers could define some important locations that they want to reach easily.

### 2.3. Mapping and classification of the decision variables

Two types of maps are created. First, variable maps are created for describing the category of each variable in the city and after that final maps are created for showing suitable areas and dwellings based on the priorities of homebuyers. In order to make the method applicable, the created maps of the decision variables should have certain characteristics: every variable have its corresponding map, so that the system is sensitive when combining maps of the variables; variables should be self-explanatory for citizens and potentially useful for further research; the final maps should show the homebuyer the resulting suitable areas and the ranking of the dwellings; and the quality levels for each variable should be classified on the basis of the maximum distance to the closest service points within the populated areas.

The quality levels should be based on objective and measurable parameters. Using an objective method for classification ensures an evaluation without discussions that subjective classification would entail. In this sense, in the case of variables which have international classification systems (e.g. air quality and noise level), the preestablished classification is used; while for other variables, the classification was based on fixed percentage values of maximum distance to the closest service point. The ranges of values should not be too wide or too restrictive for the benefit of citizens. This is why we used a mistake-proofing method to try to find the percentage values which express the different quality levels and at the same time show suitable areas for the most frequently encountered cases.

The extension of the maps must occupy the entire municipal areas, so that nonpopulated suitable areas can also be visualized. This enables the evaluation of potential developing areas, if they present good variable levels or characteristics close to the levels required by the customers.

Considering all the prerequisites, the creation of the maps for the variables follows these steps:

First, the extent of the analyses must be defined. For this, only populated areas based on social and economic data are selected.

Second, raster layers for proximity to service points are created using raster tools in ArcGIS Spatial Analyst. For this, firstly, distance to the closest service point is first calculated via the road network (including roads, pedestrian, and bike lines). Then, for the pixels outside...
Table 1. Raster maps for decision variables calculated using GIS.

| Variables | Data used | Purpose of the variable |
|-----------|-----------|-------------------------|
| (1) Nature | Road network data (streets, pedestrian, and bike lanes). Forest inventory raster data. Water from topographic maps | Measure the proximity to a natural environment |
| (2) Water | Water from topographic maps | Measure the proximity to water areas |
| (3) Noise level from road traffic: day and night | Noise maps | Measure spatial variation of road traffic noise in the city, distinguishing night hours from day hours |
| (4) Noise level from rail traffic: day and night | Noise maps | Measure spatial variation of rail noise in the city, distinguishing night hours from day hours |
| (5) PM$_{10}$ concentration based on dispersion model | Maps of annual average concentration in $\mu g/m^3$ | Measure spatial variation of air quality |
| (6) Children’s playgrounds locations | Road network data, children’s playground locations | Measure the proximity to public children’s playgrounds |
| (7) Children’s daycare centers. | Road network data, kindergarten locations | Measure the proximity to children’s daycare centers |
| (8) Primary schools | Road network data, primary schools locations | Measure the proximity to public educational centers for children and young people |
| (9) Secondary schools | Road network data, secondary school locations | Measure the proximity to public educational centers for children and young people |
| (10) Libraries | Road network data, library locations | Measure the distance to libraries |
| (11) Sports centers | Road network data, sports centers locations | Measure the distance to the main sports centers. Nearby sports spaces are especially important in Finnish societies during winter time |
| (12) Markets | Road network data, food market locations | Measure the distance to food markets |
| (13) Public health centers | Road network data, public health center locations | Measure the distance to the public health centers |
| (14) Beaches | Road network data, official beach locations | Measure the proximity to public and official beaches |
| (15) Centers of neighborhoods | Road network data, neighborhood center locations | Measure the distance to the social center of neighborhoods |
| (16) Proximity to special interest points | Road network data, interest point locations | Measure the distance to the points of interest chosen by the customer and not included in other variables |

the road network, the closest road distance values are allocated and summed with the Euclidean distance to the closest road.

Finally, distance raster are classified into five categories based on 5, 15, 30, 45, and 70% of the maximum distance to the closest service point within the populated areas. Only the populated areas are used for defining the maximum distance, to avoid unrealistic class breaks. The five categories rank values from 1 to 5 (where 5 is the best) and are used at the end of the process as decision variables for defining suitable areas and ranking of the dwellings. This classification is considered to be discriminative enough and realistic for different geographical areas and users.

In the case of noise levels, the classification of the variables was done according to the Caltrans Transportation Laboratory Noise Manual (1982) (15), and the modification by the Environmental Science Associates is adapted to more demanding levels for the night cases. Maximum noise levels were fixed at 65 dBA for night and 75 dBA for day time, following the advice of local experts. The minimum noise level was 35 dBA, which is considered to be the natural background noise level.

3. Calculation
3.1. Ranking of dwellings

GIS tools are applied to link the levels of decision variable maps to the locations of alternative dwellings. For the ranking of alternative dwellings, two metrics ($D_1$ and $D_2$) based on differences between target levels and real levels of decision variables for each dwelling are calculated as follows:

$$D_1 = \sum_{i=1}^{n} |(V_i - P_i)| - \sum_{i=1}^{n} (V_i - P_i)$$  \hspace{1cm} (1)

$$D_2 = \sum_{i=1}^{n} (V_i - P_i)$$  \hspace{1cm} (2)

where $V_i$ is the real level of variable $i$ for dwelling, $P_i$ is the target level of variable $i$ given by the user, and $n$ is the number of decision variables considered. The variables that the user does not mark as relevant are excluded from the ranking procedure.

The minimum value of $D_1$ is zero, and will be reached for dwellings that fulfill all targets given by the user. $D_2$ is the sum of the differences between real and
target levels, and is needed to rank dwellings having the same value for $D_1$.

Thus, for ranking dwellings that fulfill all target levels ($D_1 = 0$), ranking is based on the value of $D_2$, and the best solution is the one that has the highest value for $D_2$. If there are still dwellings that have the same value for $D_2$, they are ranked equal.

If there are no dwellings that fulfill all target levels (the lowest value for $D_1 \lor 0$), the ranking takes place as follows. The lower the value of $D_1$, the higher the ranking. The alternatives that have the same value for $D_1$ are ranked on the basis of the value of $D_2$, and dwellings that have the same value for $D_2$ are ranked equal.

4. Case study

The method was developed and tested using the City of Kuopio, which is situated in the North Savo region in Eastern Finland, as case study area. Fifty dwellings in Kuopio with an area bigger than 70 square meters and more than three rooms were selected from home search portals. This was done using the existing searching tools offered on the web site. Some natural environments and the city center were applied as proximity to a special interest point.

Target levels for two simulated cases of homebuyers are presented in Table 2. The target levels for the two case studies simulated a young family with small children (Case 1) and an elderly single person (Case 2).

Suitable areas and locations of best dwellings for two cases are shown in Figure 1. Different metrics and results for ranking are shown in Table 3. In Case 1, dwellings were found in suitable areas, whereas none was found in the second case.

All the dwellings which had a result value of 0 in Equation (1) were located inside the suitable area (Figure 1(a)). However, in Case 2, since no dwellings had a result value of 0 in Equation (1), there are none within the suitable areas (Figure 1(b), Table 3).

5. Discussion and conclusion

The current study proposes various spatial decision variables for living environment and method for the mapping of suitable areas, as well as ranking of alternative homes for home selection process. The method is easy to use and simple enough to be implemented into home search portals, but can also be used for city planning to find out worst and best areas, or to compare the quality of the living environment among selected neighborhoods by applying different user requirements. If applied in a home search portal, the data on user requirements could be collected for further studies or for practical city planning purposes. The variables and method demonstrated in the case study would be useful for many homebuyers. Additional variables or different classifications of criteria maps can be applied easily.

When comparing our study to previous ones research (1, 2), we have found differences worth mentioning.

First, spatial decision variables affect a wide diversity of variables for all the studies. However, variables related to the home characteristics were excluded in our study, focusing only on the living environment variables. This option facilitates further studies on evaluating the quality of living environment in neighborhood and city scale, because all houses are evaluated concerning the living environment based on the populated areas.

Secondly, the proximity measurements in our method were based on the distance via the road network, instead of using straight-line distance. This permitted a more realistic measure of the access to the different service points. Two more possibilities in measuring accessibility are suitable to be applied: first, differentiate the different mobility networks within the city (e.g. primary and secondary roads, pedestrian and cycling paths, and natural paths). Second, the distance could have included energy costs and slope of the terrain. In both cases, these options were not applied due to the lack of real data about the mobility and in order to simplify the entire process.

Thirdly, to make different decision variables comparable and easy to use, we applied classification of variables into five levels based on objective parameters (percentage of maximum distance to the closest service point within populated areas). The classification enabled quick visualization of decision variables and suitable areas. This method based on five quality levels of variables seemed to work well and would be easy to use by ordinary citizens for visualization of good areas and best home selection. Contrary to the previous studies by Rinmer et al. (2) and Natividade-Jesus et al. (1), our method does not enable the user to set any weights for target variables. In addition, application of only five levels of classification may reduce the potential diversity of solutions. However, our method can be easily modified.

| Variables                              | Case 1 | Case 2 |
|----------------------------------------|--------|--------|
| Nature                                 | 2      | 1      |
| Water                                  | 0      | 3      |
| Road traffic noise (day)               | 0      | 4      |
| Road traffic noise (night)             | 4      | 2      |
| Rail traffic noise (day)               | 0      | 3      |
| Rail traffic noise (night)             | 2      | 2      |
| PM$_{10}$ concentration               | 2      | 1      |
| Children’s playgrounds                 | 5      | 0      |
| Children’s daycare centers             | 4      | 0      |
| Primary schools                        | 4      | 0      |
| Secondary schools                      | 5      | 0      |
| Libraries                              | 3      | 4      |
| Sport centers                          | 3      | 3      |
| Markets                                | 3      | 4      |
| Public health centers                  | 0      | 2      |
| Beaches                                | 3      | 4      |
| Centers of neighborhoods               | 2      | 3      |
| Proximity to special interest points   | 0      | 3      |
to enable different classifications for variables as well as the use of weights for target variables. Because the classification of decision variables is based on the maximum distance to the closest service points within populated areas, the spatial distribution of services, as well as the spatial extent and geographical properties of the analysis area have some effects on the class breaks. Therefore, the proposed variables and classifications may be most suitable for urban areas only. The method can be applied for mapping suitable areas and ranking of homes in rural areas, but some changes for available variables and the classifications will be needed.

In some studies relating happiness levels or the quality of life with the spatial conditions, investigators emphasize the need for a tool that can facilitate communication for both citizens and researchers, and provide useful information for citizens while creating data for further research (6, 7, 16).

Participatory-GIS (P-GIS) is a field explored by different studies in order to evaluate how public participation can be included the evaluation and planning of

| Rank | Result Equation (1) | Result Equation (2) | Result Equation 1 | Result Equation 2 |
|------|---------------------|---------------------|-------------------|-------------------|
| 1    | 0                   | 16                  | 2                 | 20                |
| 2    | 0                   | 13                  | 2                 | 13                |
| 3    | 0                   | 11                  | 4                 | 14                |
| 4    | 2                   | 12                  | 4                 | 13                |
| 5    | 2                   | 10                  | 4                 | 12                |
| 6    | 2                   | 9                   |                   |                   |
| 7    | 2                   | 8                   |                   |                   |
specific areas close to them (17–19). In most cases, the experiences have brought positive results increasing the participation of teenagers and children, which are not included in the usual public queries.

Moreover, Carver et al. (17) affirms that, by his experience, the smaller the scale is to evaluate, the more the public participation increases. In this sense, we consider that our proposed method could be implemented as P-GIS for extracting positive and negative views of the current living environment and forecast quality of city areas with expectations for a high public participation. The application of the proposed method in other international contexts is a plausible option. Comparison between different urban areas based on the distances of the same variables used in the current study could provide interesting conclusions related to spatial location of the studied variables.

In conclusion, the method allows the application of new variables in mapping suitable areas and ranking alternative dwellings, which would be useful for homebuyers, real estate agents, and to some extent for city planners. In addition, it adapts to the user’s requirements, gives detailed solutions in map and ranking forms (suitable area maps and ranking of the dwellings respectively), and can provide spatial-related results when it comes to analysis of the homebuyers’ requirements.

Acknowledgement
Supported by the CIMO Fellowship Program.

Notes on contributors
Xavier Albacete, MSc in Environmental Science for the University of Barcelona (2009), is currently a PhD candidate at the University of Eastern Finland (Finland). His major research interests are in Geographic Information Science and its application into urban aspects including: (1) (visual) spatial analysis and modeling, (2) city quality evaluation, and (3) city design optimization.

Kari Pasanen holds an MSc on Forestry from the University of Joensuu in 1995. His studies have been focused on forestry and environmental epidemiology; with special interest in spatial decision analysis, environmental health, and GIS.

Mikko Kolehmainen was born in Finland, on 29 September 1962. He received an MSc degree in software engineering and a DSc degree in computational methods from the Helsinki University of Technology, Helsinki, Finland, in 1987 and 2004, respectively. He also received an MSc degree in biotechnology from the University of Kuopio, Kuopio, Finland, in 2001. He started his career with a private software company as an analyst and development manager. Since 1996, he has been with the University of Kuopio, Kuopio, Finland, first as a researcher, a research director since 2002, and a professor of environmental informatics since 2006. Under his guidance, the research group has carried out successfully both national and European Union projects. His main research and teaching interests are in applying neural networks to environmental data-sets and applying methods of bioinformatics in environmental monitoring. He was the chair of the 2009 ICANNGA conference in Kuopio and continues to work in its Advisory Committee.

References
(1) Natividade-Jesus, E.; Coutinho-Rodrigues, J.; Antunes, C.H. A Multicriteria Decision Support System for Housing Evaluation. Decis. Support Syst. 2007, 43 (3), 779–790.
(2) Rinner, C. The Spatial Dimensions of Multi-Criteria Evaluation – Case Study of a Homebuyer’s Spatial Decision Support System. Springer 2006, 338–352.
(3) Brereton, F.; Clinic, P.; Fererre, S. Happiness, Geography and the Environment. Ecol. Econ. 2008, 68, 386–396.
(4) MacKerron, G. Life Satisfaction and Air Quality in London. Ecol. Econ. 2009, 68 (5), 1441–1453.
(5) Saaty, T.L. The Possibility of Group Welfare Functions. Int. J. Inform. Technol. Decision Making 2005, 4 (2), 167–176.
(6) Ferrer-i-Carbonell, A.; Gowdy, J.M. Environmental Degradation and Happiness. Ecol. Econ. 2007, 60, 509–516.
(7) Apparicio, P. The quality of the urban environment around public housing buildings in Montreal: An objective approach based on GIS and multivariate statistical analysis. Soc. Indic. Res. 2008, 86 (3), 355–380.
(8) Villa, F. A GIS-Based Method for Multi-Objective Evaluation of Park Vegetation. Landscape and Urban Planning 1996, 35 (4), 203–212.
(9) Vreeker, R. Evaluating Effects of Multiple Land-Use Projects: a Comparison of Methods. J. Housing Built Environ. 2006, 21 (1), 33–50.
(10) Malczewski, J. GIS-Based Multicriteria Decision Analysis: A Survey of the Literature. International Journal of Geographical Information Science 2006, 20 (7), 703–726.
(11) Rinner, C.; Malczewski, J. Web-Enabled Spatial Decision Analysis Using Ordered Weighted Averaging (OWA). J. Geogr. Syst. 2002, 12 (4), 385–403.
(12) Banai, R. Land Resource Sustainability for Urban Development: Spatial Decision Support System Prototype. Environ. Manage. 2005, 36 (2), 282–296.
(13) Karnatak, H.C. Multicriteria Spatial Decision Analysis in Web GIS Environment. GeoInformatica 2007, 11 (4), 407–429.
(14) Strager, M.P.; Rosenberger, R.S. Incorporating Stakeholder Preferences for Land Conservation: Weights and Measures in Spatial MCA. Ecol. Econ. 2006, 57 (4), 627–639.
(15) Jones & Stokes. Transportation-and Construction-Induced Vibration Guidance Manual; California Department of Transportation, Environmental Program, Environmental Engineering, Noise, Vibration, and Hazardous Waste Management Office: Sacramento, CA, 2004.
(16) Foster, A. Quality of Life: A Good Practice Guide to Communicating Quality of Life Indicators. Audit Commission, 2003.
(17) Carver, S.; Evans, A.; Kingston, R.; Turton, I. Public Participation, GIS, and Cyberdemocracy: Evaluating On-line Spatial Decision Support Systems. Environ. Plann. B: Plan. Design 2001, 28, 907–921.
(18) Cinderby, S. How to Reach the ‘Hard-to-Reach’: The Development of Participatory Geographic Information Systems (P-GIS) For Inclusive Urban Design in UK Cities. Area 2010, 42 (2), 239–251.
(19) Cinderby, S.; Forrester, J. Facilitating the Local Governance of Air Pollution Using GIS for Participation. Appl. Geogr. 2005, 25 (2), 143–158.