DETERMINATION OF FACTORS INFLUENCING LAND USE POTENTIALS IN HONG KONG

C.K. Wong
Research Student
Department of Civil and Structural Engineering
The University of Hong Kong
Pokfulam Road, Hong Kong
Fax: (852) 25595337
E-mail: wongck@hkusua.hku.hk

C.O. Tong
Associate Professor
Department of Civil and Structural Engineering
The University of Hong Kong
Pokfulam Road, Hong Kong
Fax: (852) 25595337
E-mail: cotong@hkucc.hku.hk

S.C. Wong
Lecturer
Department of Civil and Structural Engineering
The University of Hong Kong
Pokfulam Road, Hong Kong
Fax: (852) 25595337
E-mail: hhecwsc@hkucc.hku.hk

abstract: This paper describes the second stage of the development of a calibration procedure for the Lowry-type land use model, which was formulated and calibrated for the city of Hong Kong. The model contains, for each zone, a population and an employment potential which indicated, respectively, the attractiveness of the zone to residents and workers. A Genetic Algorithm was developed to calibrate these potentials. The aim of this paper is to investigate the relationship between these potentials and various land use variables, using a multivariate regression analysis. A case study in Hong Kong is presented to illustrate the potential application of this methodology.

1. INTRODUCTION

This paper is the second stage of the development of a calibration procedure for the Lowry-type land use model, which was formulated and calibrated for the city of Hong Kong. Lowry model has been commonly used worldwide to model the population and employment distributions in a city region (Wilson, 1974) A comparison between Lowry model and conventional regression analysis in modelling the population and employment distributions in relation to the land use parameters was made (Foot, 1974), in which it was found that Lowry-type models were generally stable and reliable, while the regression analysis may result in absurd predictions as negative population and employment may be estimated with the future values of the explanatory variables. The regression analysis, however, is able to identify the significant variables through statistical testing.

The calibration of Lowry model was studied by Putman and Ducca (1978a) for population distribution using a maximum likelihood approach. A product form of the population potential indicating the attractiveness of a zone to residents with respect to the land use variables was established. The unbiased estimators of these potentials were determined by maximizing a likelihood function. However, statistical testing on the significance of the chosen land use variables was not conducted. The selection of important land use variables was largely based on the experience and judgement of planners. The method was applied to several cities (Putman and Ducca, 1978b, Putman, 1980). It was found that the results varied from one city to another, while some cities gave a high R^2 value between observed and modelled population distributions, some did not. In the model, the land use factors used in the attractiveness measure were pre-specified no matter they were significant or not. For some cities, the attractiveness measures do have very close relationships with those preset land use variables. For others, however, the attraction may be constituted by some omitted land use variables.

The proposed model contains, for each zone, a population and an employment potential which indicated, respectively, the attractiveness of the zone to residents and workers. In the first stage of model development, by using the 1992 base year data of zonal population and employment and travel time between zones, the model was successively calibrated by a Genetic Algorithm and stable values of land use potentials were obtained for each zone (Wong et al, 1996). The aims of the present research are to investigate various factors that may be influencing the values of these potentials and to establish models. Once calibrated, these models could be used for future predictions of zone land use potentials and also as policy models to evaluate the effectiveness of various land use policies on population and employment distributions.

The first part of the work is to identify various factors that may influence land use potentials. For population potential, these could be residential floor area, number of school places, level of facilities, etc. The same kind of land use variables could also influence the employment potential, however, some modifications may be appropriate. For instance, the residential floor area could be replaced by commercial, industrial, storage, and shopping floor areas. The second part of the work is to measure the values of various factors for the year 1992 for Hong Kong. The following data sources are used in the analysis. Different types of land use area are measured from outline zoning plans produced by the Town Planning Department in Hong Kong. The permitted plot ratios of different sites are governed by the Building (Planning) Regulations. These two components together determine the floor area. Other data are mostly extracted from the 1992 Hong Kong Travel Characteristics Survey (Transport Department, 1993). In the third part, multivariate linear regression analysis is employed to establish relationships between these factors and the potentials. The objective is to generate two equations, one for determining population potential and the other for determining employment potential. By using stepwise regression, only those factors which significantly influence land use potentials are retained in the equations.

Section 2 summarizes the calibration procedure of the Lowry-type model to determine the population and employment potentials. In Section 3, the second stage of establishing the relationship between these potentials and land use variables is presented. Section 4 demonstrates a case study in Hong Kong, and Section 5 discusses the application of the proposed methodology. Suggestions for further investigations are provided in Section 6.

2. CALIBRATION OF LOWRY-TYPE MODEL

The Lowry-type model comprises two distribution steps. One is the residence location and the other is the employment location. The distribution process is based on a gravity distribution procedure as used in the trip distribution models. The population in a zone is estimated by the following,

$$P_i = \alpha \sum_{j=1}^{n} E_j \frac{W_{Pi}\, e^{-\beta c_{ji}}}{\sum_{i=1}^{n} W_{Pi}\, e^{-\beta c_{ji}}} \tag{1}$$

Similarly, the distribution of employment places can be determined using the gravity- type model shown below:

$$E_j = \frac{1}{\alpha} \sum_{i=1}^{n} P_i \frac{W_{Ej}\, e^{-\beta c_{ij}}}{\sum_{j=1}^{n} W_{Ej}\, e^{-\beta c_{ij}}} \tag{2}$$

where

P_i is the estimated population at zone i,
E_j is the employment at zone j,
W_{Pi} is the population potential at zone i,
W_{Ej} is the employment potential at zone j,
c_{ji} is the travel cost from zone j to zone i,
n is the number of zones,
α is the regional population to employment rate, and
β is an impedance coefficient of the travel cost.

Equations (1) and (2) are applied alternately until the population and employment patterns stablise, given some initial values assumed for the population and employment potentials and the coefficient β. By minimising the discrepancy between observed population and employment distributions in the base year and the modelled results, a set of population potentials, employment potentials and a sensitivity coefficient of the travel cost β are then calibrated.

The following objective function derived from the likelihood function is employed to measure the discrepancy between observed and modelled results in the base year:

$$F(\mathbf{W}_P,\ \mathbf{W}_E,\ \beta) = \left(\sum_{i=1}^{n} \frac{(\overline{P}_i - P_i)^2}{\sigma_{Pi}^2} + \sum_{i=1}^{n} \frac{(\overline{E}_i - E_i)^2}{\sigma_{Ei}^2} \right) \tag{3}$$

where

F is the objective function (i.e. the sum of error between the observed and modelled values),
$\overline{E}_i$ is the number of observed employment at zone i,
E_i is the modelled number of employment at zone i,

$\overline{P}_i$ is the observed population at zone i,
P_i is the modelled number of population at zone i,
$\mathbf{W}_P$ is the vector of the population potentials,
$\mathbf{W}_E$ is the vector of the employment potentials,
σ^2_{Pi} is the variances of the observed population, and
σ^2_{Ei} is the variances of the observed employment.

The least-square form of the objective function is derived from a maximum likelihood approach. In general, the number of total population of a modelled area exceeds the number of total employment, but this scaling effect can be accommodated by the variances of the observed population σ^2_{Pi} and employment σ^2_{Ei}.

3. RELATIONSHIP BETWEEN POTENTIALS AND LAND USE VARIABLES

From the first stage of model calibration, a set of population potentials $\mathbf{W}_P$ and a set of employment potentials $\mathbf{W}_E$ were determined. This section attempts to establish the relationships between these potentials and the various land use variables. To avoid giving negative potentials, a product form of relationship is specified (Putman, 1980).

The population potential is given by

$$W_P = a_0 X_1^{a_1} X_2^{a_2} X_3^{a_3} \ldots\ldots X_m^{a_m}, \tag{4}$$

and the employment potential is

$$W_E = b_0 Y_1^{b_1} Y_2^{b_2} Y_3^{b_3} \ldots\ldots Y_n^{b_n}, \tag{5}$$

where
$a_0, a_1, a_2, \ldots, a_m$; $b_0, b_1, b_2, \ldots, b_n$ are coefficients to be calibrated,
$X_1, X_2, \ldots, X_m$ (≥ 1); $Y_1, Y_2, \ldots, Y_n$ (≥ 1) are land use variables,
m is the number of land use variables for population potential, and
n is the number of land use variables for employment potential.

Since zone characteristics vary within a modelled area, some land use variables may be missing in particular zones. In such circumstances, a minimum value of '1' instead of '0' is assigned to those zero zonal land use variables to handle the numerical difficulty.

Before applying multiple regression analysis, a simple transformation is applied by taking logarithm on both sides of equation (4) and (5), we have respectively

$$\ln \mathbf{W}_P = \ln a_0 + a_1 \ln X_1 + a_2 \ln X_2 + a_3 \ln X_3 \ldots\ldots + a_m \ln X_m \tag{6}$$

and

$$\ln \mathbf{W}_E = \ln b_0 + b_1 \ln Y_1 + b_2 \ln Y_2 + b_3 \ln Y_3 \ldots\ldots + b_n \ln Y_n \tag{7}$$

By the method of least squares, all the coefficients a_1, a_2, a_3, ... a_m and b_1, b_2, b_3, ... b_n as well as the constant terms ln a_0 and ln b_0 can be determined.

Besides, all significant land use factors will be retained in the equations and those factors of less importance will be discarded through the stepwise linear regression analysis. Since factors attracting residence are not identical to those attracting workers, some of the variables retained in equation (4) may or may not be the same as those used in equation (5), which depends largely on the level of significance. Moreover, the choice of land use variables is more flexible.

4. CASE STUDY AND RESULTS

4.1 Case study

In this section, a case study of Hong Kong is presented. In Hong Kong, government's policy dominates future land uses. These policies are implemented through Outline Zoning Plans undated by the Hong Kong Planning Department at regular intervals. Every Outline Zoning Plan contained detailed descriptions on the future land uses of a district. Various types of land uses may be specified, such as commercial, residential and industrial uses. Floor areas are calculated from the plans based on plot ratio and site area. Nearly all land use data contained in Outline Zoning Plans form part of the database in the study. In addition, various school places, number of cinema seats, market stalls, hotel rooms and hospital beds are also made used. All these data are extracted based on the year 1992 which is compatible with the results obtained in stage 1. A list of them is given on Table 1.

Furthermore, all the variances of the observed population σ^2_{Pi} and the variances of the observed employment σ^2_{Ei} are set to '1' in this case study.

4.2 Results

4.2.1 Equation for population potential

After implementing the stepwise regression analysis by entering all the land use information, the regression results for population potential are shown in Tables 2 and 3.

Although the number of zones in the data sets is 274, the number of cases used in the analysis of population potential is only 213. This is because some zones have little development. The R^2 value of 0.65 indicates that there is a reasonably close relationships between population potential and land use variables. However, some coefficients estimated have a high standard error.

All significant land use variables related to population potential are listed in Table 3. In Hong Kong, almost all primary school places are allocated to students by the Education Department according to their home locations. Most kindergartens are private owned and they are mostly located within residential districts. Thus, they are entered into the model with high significance.

Table 1 List of land use data

No.	Land use type	Remarks
1	Office floor area	measured in meter square
2	Commercial floor area	measured in meter square
3	Flatted factory floor area	measured in meter square
4	Specialised factory floor area	measured in meter square
5	Warehousing floor area	measured in meter square
6	Building materials storage area	measured in meter square
7	Container storage area	measured in meter square
8	Comprehensive Development Area	area under intensive development
9	Commercial and Residential floor area	
10	Residential floor area of density type - I	residential area of highest plot ratio, i.e. highest density
11	Residential floor area of density type - II	residential area of middle density
12	Residential floor area of density type - III	residential area of lowest density
13	Residential floor area of density type - IV	residential area located in rural area or new town
14	MTR catchment area	area covered by radius of 600m from station exits
15	KCR catchment area	area covered by radius of 600m from station exits
16	Village	
17	Government, Institute and Community	facilities including schools, churches and government offices
18	Open space	recreation uses
19	Other specified uses	government uses (mostly provide services)
20	Undetermined	not commonly seen on plans
21	Green Belt	
22	Country Park	
23	Kindergarten places	
24	Primary school places	
25	Secondary school places	
26	Tertiary school places	including universities, and other post secondary institutes
27	Cinema seats	
28	Market stalls	
29	Hotel rooms	
30	Hospital beds	

Table 2 General regression result for population potential

No. of cases	R^2	Standard error of estimate
213	0.6524	0.60138

Table 3 Calibration results for population potential

	Coefficient	Standard error of coefficient	t-value
Intercept	-3.45615	0.222019	-15.5669
Kindergarten places, X_1	0.13368	0.024296	5.5020
Primary school places, X_2	0.09338	0.023113	4.0402
Residential floor area of density type - I, X_3	0.03490	0.009682	3.6047
Residential floor area of density type - IV, X_4	0.03699	0.010900	3.3939
Residential floor area of density type - III, X_5	0.03727	0.013068	2.8520
Tertiary school places, X_6	-0.03611	0.017793	-2.0294

The influence of tertiary school places, on the other hand, is less important when compared to other land use factors. In Hong Kong, most tertiary institutes are situated in remote area, near a commercial center or inside a reserved zone. Thus, the development of residential type land uses and services are prohibited. And it reduces the population attractiveness directly. As a result, a negative figure for this land use type is calibrated that the overall attraction potential is lowered.

Three types of residential floor area are entered into the model since they are the basic variables attracting residents. Among them, the floor area of density type-I is of the highest importance as this type covers the largest number of population within the whole study area. Naturally, the more the residential floor area, the higher the population potential as more residents can be housed. In the present approach, the number of resident does not totally depend on the amount of these residential floor area since the population potential is also affected by other land use variables. Besides, the number of resident is determined by the Lowry model based on the overall population potential as well as the travel cost.

4.2.2 Equation for employment potential

Tables 4 and 5 shows the regression results for employment potential.

Table 4 General regression result for employment potential

No. of cases	R^2	Standard error of estimate
228	0.5596	0.74475

Table 5 Calibration results for employment potential

	Coefficient	Standard error of coefficient	t-value
Intercept	-4.07864	0.680307	-5.99530
Office floor area, Y_1	0.10053	0.022152	4.53812
Specialised factory floor area, Y_2	0.06648	0.015271	4.35332
Market stalls, Y_3	0.09977	0.024509	4.07052
Greenbelt land area, Y_4	-0.03289	0.009378	-3.50738
Residential floor area of density type - I, Y_5	-0.03887	0.011262	-3.45121
Residential floor area of density type - IV, Y_6	-0.03441	0.011572	-2.97405
Tertiary school places, Y_7	-0.06198	0.022886	-2.70842
Hotel rooms, Y_8	0.08268	0.031950	2.58764
Commercial floor area, Y_9	0.05227	0.023446	2.22958
Building material storage floor area, Y_{10}	0.03605	0.016370	2.20230
MTR catchment area, Y_{11}	0.02160	0.010636	2.03120

Although the R^2 value of 0.56 for the employment model is a bit lower than that of the residence model, a satisfactory relationship between employment potential and land use variables can be established. Again, the high standard error may be typical of land use models.

The office floor area, market places, specialised factory floor area, hotel rooms, commercial floor area and building material storage floor area all contribute to employment and therefore they have positive effect on the employment potentials. Among them, office and specialised factory are the most significant.

The MTR stations catchment areas are significant for the employment attractiveness. It is because the locations of large commercial centers are close to such stations where they can be easily accessed while these centers accommodate large number of job opportunities.

The residential floor areas of density type-I and type-IV and green belt area are negatively correlated to the employment potential since, by nature, they do not generate a large number of job opportunities. Moreover, number of jobs produced by tertiary institutes is small when compared to other commercial use areas so that it is less significant to the employment potential and as a result, has a negative coefficient.

5. MODEL APPLICATION

Using the equations developed in Section 4, population and employment potentials can be determined from land use variables. If the future land use variables are known, the future zonal population and employment potentials can be estimated. If the future travel time between zones is also known, these are then applied to the Lowry model to predict the future distributions of population and employment. The model application procedure is illustrated in Figure 1. Various land use development schemes proposed by the government can then be evaluated for their effectiveness.

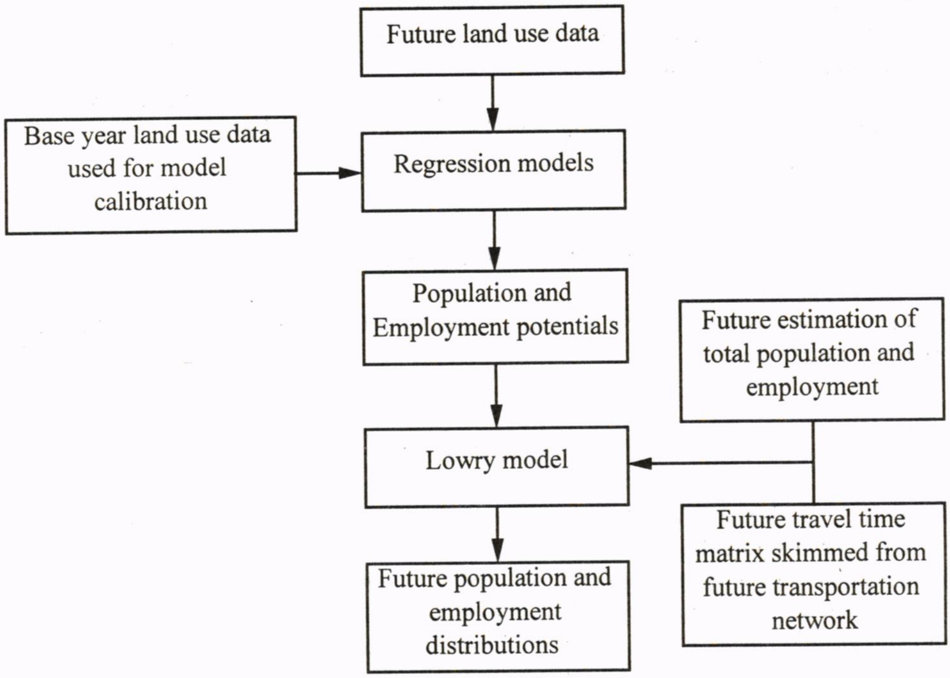

Fig. 1 Structure of model application

6. FURTHER WORKS

Relationships between population and employment attractiveness and various land use factors have been established. Using these relationships, a model for estimating aggregate population and employment can be applied. However, the applicability of the model is largely based on its level of accuracy. Thus, the next stage will be the validation of the modelling results in which the total error will be calculated. Still, the model can be improved by enlarging the database used in the regression analysis such as incorporating environmental variables. They are also important aspects in residence location, which in turn have significant correlation with the population potential.

The model can be further enhanced by stratifying the population and employment into different segments. For example, some variables may be highly correlated to one segment of the population and not the other segments, or even negatively correlated to other segments. Similarly, different employment types may be correlated to different variables. It is therefore decided that the model be further developed with population stratified by income and employment by employment type.

Even with single type population and employment, calibration of the Lowry Model by Genetic Algorithms requires much computation time. The calculation time will be increased in a geometric manner when the segments of the models are disaggregated. For instance, if both population and employment are divided into three segments, then nine sets of population and employment potentials need to be determined and therefore the problem size is nine times greater. When using Genetic Algorithms, as every individual trial genes can be treated separately, adopting parallel computation is one of the means to reduce the computation time because the total work loads are shared and taken up by several processors. The feasibility of applying parallel computation to calibrate the model will be studied during the next stage of the study (Wong et al, 1997).

ACKNOWLEDGEMENTS

This research was supported by a research grant HKU 435/94E from the Hong Kong Research Grant Council.

REFERENCES

Foot D.H.S. (1974) A Comparison of some land-use allocation / interaction models, George Over Limited.

Putman S.H. (1980) Calibrating urban residential location models 3: empirical results for non-US cities, *Environment and Planning A*, **12(2)**, 813-828.

Putman S.H. and Ducca F.W. (1978a) Calibrating urban residential models 1: procedures and strategies, *Environment and Planning A*, **10**, 633-650.

Putman S.H. and Ducca F.W. (1978b) Calibrating urban residential models 2: empirical results, *Environment and Planning A*, **10**, 1001-1014.

Transport Department with MVA (1993) *Travel Characteristic Survey : Final Report*, Hong Kong Government.

Wilson A.G. (1974) Urban & Regional Models in Geography & Planning, Wiley.

Wong C.K., Wong S.C. and Tong C.O. (1996) Calibration of Lowry Model by Genetic Algorithm, *Proceedings of The First Conference of Hong Kong Society for Transportation Studies*, Department of Civil and Structural Engineering, The Hong Kong University of Science and Technology, Hong Kong, 213-222.

Wong C.K., Wong S.C. and Tong C.O. (1997) The Parallel Implementation of the Genetic Algorithm to Calibrate a Lowry Model, *The fifth International Conference on Computers in Urban Planning and Urban Management*, Indian Institute of Technology, Bombay, India. (*accepted for publication*)