CHALLENGES OF DATA REQUIREMENTS FOR MODELLING RESIDENTIAL LOCATION CHOICE: THE CASE OF BERLIN, GERMANY

Benjamin Heldt
Kay Gade
Dr.-Ing. Dirk Heinrichs
German Aerospace Center, Institute of Transport Research

The locations of households and firms play a significant role in daily mobility patterns. Location decisions as long-term mobility decisions considerably affect short-term mobility decisions such as destination choice or mode choice (Wegener & Fürst, 1999). Applying land use models helps to better understand the corresponding relations of land use and residential mobility to transport, especially in the urban context. While land-use models in general, and specifically location models, are very useful tools to analyse and plan urban development, they imply considerable data requirements (Wegener, 2011). Although data availability is improving, selecting suitable data is still a challenging process and often requires compromises.

This paper addresses a set of related questions: First, what criteria are best suited to evaluate and select available data for location-choice modelling? Second, what are the 'typical' limitations of such data and what are the available methods to overcome them? The paper explores these questions empirically by describing the setting up a residential location model for Berlin, Germany. Berlin, a city with a dense and well-functioning transport system, has recently featured a strongly growing population and thus provides an interesting case for analysing the interactions between long-term and short-term mobility. The applied model is the land use modelling framework 'Cube Land' (Martinez, 1992, Martinez and Henriquez, 2007; Martinez and Donoso, 2010).

In the paper, we define model characteristics (spatial level of detail, model components, factual level of detail, interrelations, and parameters) that serve to determine whether data suits a model's purposes. This way we introduce a framework for the description of a model's data requirements and define the needed data for estimating Cube Land's residential location model. The application of this framework to the different data sources available for the Berlin case identifies Mikrozensus as the most adequate data source that is currently accessible. The framework reveals, however, that all compared data sources have their limitations. Either they do not include all the necessary attributes for residential location modelling, or data confidentiality requires their aggregation, leading to loss of information. Methods for reducing this information loss and adding missing attributes are described. In general, this paper demonstrates that surveys suited to residential location modelling are rare, in particular with reference to the linkages between longer term residential mobility and daily mobility decisions. This shows the need for the inclusion of corresponding questions in surveying and closer integration of land use and transport domains.

The application of the criteria developed in this article helps to critically evaluate data sources and assess their advantages and disadvantages regarding spatial and factual depth of information. In this way, the described method may be of value to increase awareness of a model's quality depending on the data at base.

BIBLIOGRAPHY

Martínez, F. (1992) The Bid-choice land use model: an integrated economic framework. *Environment and Planning A,* 24 871-885.

Martínez, F., Henríquez, R. (2007) The RB&SM: a random bidding and supply land use equilibrium model, *Transport Research B,* 41 631-651.

Martínez, F., Donoso, P. (2010) The MUSSA II land use auction equilibrium model, in F. Pagliara, J. Preston & D. Simmonds, eds, *'Residential Location Choice', Advances in Spatial Science*, Springer Berlin Heidelberg, pp. 99–113.

Wegener, M. (2011) From Macro to Micro – How Much Micro Is too Much?, *Transport Reviews*, 31 (2) 161-177.

Wegener, M., Fürst, F. (1999) Land use Transport Interaction: State of the Art, *Berichte aus dem Insititut für Raumplanung*, 46.