Transportation Research Part A 118 (2018) 68–88

Contents lists available at ScienceDirect

Transportation Research Part A

journal homepage: www.elsevier.com/locate/tra

Predicting travel flows with spatially explicit aggregate models On the benefits of including spatial dependence in travel demand modeling

Kasper Kerkman[a,*], Karel Martens[b,c], Henk Meurs[c,d]

[a] *Institute for Management Research, Radboud University, P.O. Box 9108, Nijmegen 6500 HK, The Netherlands*
[b] *Faculty of Architecture and Town Planning, Technion – Israel Institute of Technology, Israel*
[c] *Institute for Management Research, Radboud University, The Netherlands*
[d] *MuConsult, Amersfoort, The Netherlands*

ARTICLE INFO

Keywords:
Spatial autocorrelation
Travel demand
Spatial interaction modeling
Public transport travel flows
Spatial prediction

ABSTRACT

The prediction of travel demand is a key step in transport planning and is a topic of intense discussion of the literature. This paper adds to the debate about the accuracy of travel demand prediction by addressing the 'technical' problem of spatial autocorrelation. This paper aims to systematically assess the predictive performance of spatially explicit models that take spatial autocorrelation into account vis-à-vis more conventional models. We compare the performance of both types of models in predicting the transit passenger flows for alternative transit network designs in the region of Arnhem-Nijmegen, the Netherlands. We find that models taking spatial dependence into account outperform the conventional models in nearly all respects: model fit, parameters of variables, and the quality and stability of the predictions. Results show that taking spatial autocorrelation into account is not only important for the analysis of spatial interactions, but also result in different and more accurate predictions of the impact of interventions. We conclude that travel demand models should account for spatial dependence in order to avoid overprediction of the impact of transport system changes. We end with a discussion about the relevance of our findings for the debate about the causes for the observed systematic over-estimation of travel demand in the practice of transport planning.

1. Introduction

Travel demand modeling is an important tool to support transport planning in general and decision making about major transport supply interventions in particular. Obviously, reliable predictions of the effects of transport interventions are essential for making adequate decisions on possible interventions. Although there is much literature on different methods and models to predict future travel demand, there has been much less research into the actual (real world) quality of travel demand forecasts (van Wee, 2007).

The literature distinguishes two types of long-term travel demand forecasts: reference forecasts and policy forecasts (Andersson et al., 2017). Reference forecasts concern the future transport volumes in a do-nothing scenario, whereas policy forecasts predict the travel effects of policy interventions. Recently, Andersson et al. (2017) analyzed the accuracy of reference forecasts for passenger transport. Comparing historic Swedish national passenger transport forecast with the actual outcomes, they found substantial differences between the forecasts and the actual outcomes.

* Corresponding author.
E-mail address: k.kerkman@fm.ru.nl (K. Kerkman).

https://doi.org/10.1016/j.tra.2018.08.029
Received 20 October 2017; Received in revised form 24 May 2018; Accepted 23 August 2018
0965-8564/

An increasing number of studies has analyzed the accuracy of policy forecasts, which are of key importance for a proper assessment of the expected effects and benefits of policy interventions. The available literature suggests that the accuracy of these forecasts is often rather poor, and seem to systematically overestimate future travel demand, especially for public transport investments (Flyvbjerg et al., 2005; Perry, 2017; van Wee, 2007).

Various authors have suggested a range of reasons for the poor accuracy of travel demand models. The reasons include technical explanations (imperfect techniques, inadequate data), psychological explanations (planning fallacy, optimism bias), and political-economic explanations (deliberate and strategic manipulation of model results to generate desired outcomes) (Flyvbjerg, 2007; Flyvbjerg et al., 2005; Flyvbjerg and Holm, 2002; van Wee, 2007). Because public transport demand is seemingly structurally overestimated, technical explanations are often considered to be the least likely explanation of the poor prediction accuracy, as the assumption is that technical faults in the model would somehow lead to a balance between over- and underestimations of future travel demand (van Wee, 2007).

However, authors may have been too quick to accept this conclusion, as (recent) literature suggests in related domains (LeSage, 2014). This literature underscores the importance of taking into account the phenomenon of spatial dependence or spatial autocorrelation when analyzing spatial flows using Spatial Interaction Models (SIMs). An increasing number of studies shows that including spatial dependence through unobserved spatial factors in models of spatial flows may fundamentally alter the structure of the models and their predictions (LeSage and Llano-Verduras, 2014; LeSage and Pace, 2008). Since travel demand models are a specific kind of SIM, these recent insights underscore the importance of taking spatial autocorrelation into account in transport models. Yet, to the best of our knowledge, virtually no (applied) travel demand model accounts for this phenomenon. Failure to include spatial dependence fully in travel demand models might result in a structural error in the forecasts they produce, which may in part explain the structural over-estimation of future travel demand. The problem applies to both the SIM-models presented in this paper, as to travel demand models of the logit type (LeSage and Pace, 2008; Bhat and Guo, 2004). Even so-called marginal models, predicting changes in mobility, may be flawed, since they only take direct effects of interventions into account and not the indirect effects.

Against this background, the paper aims to assess whether spatially explicit aggregate models indeed perform better in predicting the impact of interventions in a transport network on travel flows. We focus on long-term predictions of the impacts of major interventions in the transport network, with a time horizon of 10–20 years. With the term "spatially explicit" we refer to models that systematically take into account the spatial dependence among locations and interactions. More specifically, we will compare spatially-explicit models to "conventional" models that do not account for spatial dependence. We employ our models to an analysis of hypothetical interventions in the public transport system in the Arnhem-Nijmegen region in the Netherlands.

The paper is organized as follows. Following this introduction, we describe the methodology employed for estimating the model and producing predictions for future travel flows, with a specific focus on prediction accounting for spatial dependence (Section 2). In Section 3, we describe our model specifications, including a definition of the data sets. Furthermore, we describe the performance indicators that will be used to assess and compare the performance of the different models. In Section 4, we turn to the study region and the alternative transit network designs which will be used to test the predictive performance of the different models. In Section 5, the results of the simulations are described and analyzed. We end with a discussion and reflection on the results (Section 6). In that final section, we also discuss the potential contribution of spatially explicit aggregate models for improved accuracy of predictions. More specifically, we discuss whether accounting for spatial dependence could indeed address the seemingly structural over-estimation of travel demand by conventional demand models.

Note that the paper builds on an earlier paper (Kerkman et al., 2017), but extends this work in two ways. First, where Kerkman et al. (2017) demonstrate the existence of spatial dependence in travel flows in the estimation of models, the current paper explicitly considers the problem of prediction in case of spatial dependence in travel flows. Since the latter requires an alternative estimation procedure than the former, this is the first contribution of this paper. Second, Kerkman et al. have not reported on the impacts of the inclusion of spatial dependence on the models' performance in terms of the quality of predictions. Exactly these two elements, which are vitally important for travel demand modelling, are the main focus and contribution of the current paper.

2. Taking spatial autocorrelation into account in travel demand modeling

2.1. Introduction

Recent studies in a range of research domains consistently stress the importance of taking spatial dependence – usually referred to as spatial autocorrelation – into account when modelling spatial flows (Fischer and Griffith, 2008; LeSage and Pace, 2008; Chun et al., 2012; Kerkman et al., 2017). Not accounting for spatial autocorrelation results in loss of valuable information (Schabenberger and Gotway, 2017) and poorer and biased model estimates and predictions.

The literature generally distinguishes between two forms of spatial dependence in data: spatial lag dependence and spatial error dependence (Anselin and Bera, 1998; Anselin, 1988). The spatial lag dependence encompasses two phenomena that are difficult to distinguish in practice:

- True contagion: Spatial lag dependence can be caused by a direct influence of characteristics or behavior from one location to other close locations, due to spill-overs, copy-catting, or diffusion. In transport, this may be related to network characteristics, e.g. the number of boardings at a bus stop may be affected by the number of boardings at previous stops due to crowding. In this case the inclusion of lagged variables has a substantive, theoretical meaning.
- Spatial scale mismatch: A difference between the spatial scale of phenomenon under study and the spatial scale at which it is