CENTRE FOR ADVANCED SPATIAL ANALYSIS

Working Paper Series

Paper 62

PEDESTRIAN DEMAND MODELLING OF LARGE CITIES: AN APPLIED EXAMPLE FROM LONDON

Jake Desyllas
Elspeth Duxbury
John Ward
Andrew Smith
Pedestrian Demand Modelling of Large Cities: An Applied Example from London

Dr. Jake DESYLLAS*, Elspeth DUXBURY*, John Ward, Andrew Smith**

Centre for Advanced Spatial Analysis, University College London, 1-19 Torrington Place, London WC1E 6BT, UK

*Intelligent Space Partnership, 68 Great Eastern Street, London EC2A 3JT

**Transport for London Street Management, Windsor House, 50 Victoria Street London SW1H 0TL
Abstract
This paper introduces a methodology for the development of city wide pedestrian demand models and shows its application to London. The approach used for modelling is Multiple Regression Analysis of independent variables against the dependent variable of observed pedestrian flows. The test samples were from manual observation studies of average total pedestrian flow per hour on 237 sample sites. The model will provide predicted flow values for all 7,526 street segments in the 25 square kilometres of Central London. It has been independently validated by Transport for London and is being tested against further observation data. The longer term aim is to extend the model to the entire greater London area and to incorporate additional policy levers for use as a transport planning and evaluation tool.

Keywords
pedestrian model, land use transportation planning, Visibility Graph Analysis, movement, dynamics, GIS, London
1. INTRODUCTION

The aim of transport modelling is to predict patterns of movement and the functioning of movement systems, yet research in this field until now has been almost exclusively focussed on motorised transport to the exclusion of other modes. A keyword search of the engineering INSPEC database shows 53 times more articles published since 1969 with ‘vehicle’ in the title than ‘pedestrian’ (counts of 10,211 and 192 respectively).

Perhaps the neglect of pedestrians in the research arose because modelling started at the same time as automobile dependence became a key feature of transport, so attention was focussed on understanding vehicular traffic. With the continuing rise of the car in the last 25 years, walking fell in importance from 35% of all trips in England to 26%. With this rising automobile dependence, the disparity in research between motorised and non motorised modes became ingrained in policy. Gemzøe points out that “there is no city that has a ‘pedestrian department’, recording the numbers, flow and behaviour of people on foot on the same regular basis as traffic departments record the vehicular traffic, so the pedestrians tend to be invisible in the planning process - because there are no data about them” (2001).

Despite its decline, walking still accounts for over 80% of all journeys made under a mile in length (National Statistics, 2001). In big cities, walking is a key component of transport: 40% of all inner London trips are by foot (DETR 1999), yet walking has been almost completely ignored by modellers until recently. As Brög noted; “Walking is neglected in transport policy and planning because it is often not considered in traditional transport models. But even if it is included in behavioural transport surveys, the methods applied are very often inadequate and insufficient to show its relevance for everyday mobility. And from this neglect … walking is underestimated for transport and town planning.” (2001). The transport planners who must regulate and manage our urban movement networks do not currently have modelling tools to help them understand and therefore plan for pedestrian flows.

However, there has been growing political pressure to develop more sustainable transport polices in response to automobile dependence and this is beginning to change the agenda for transport modelling. The International Council for Local Environmental Initiatives (ICLEI) formed in 1990, has had some influence on local government planning policy through the Local Agenda 21 campaign ‘dedicated to sustainability through participatory multi-stakeholder sustainable development planning’ (ICLEI 2000). The ‘visions of transport in 2030’ section of the OECD guidelines on Environmentally Sustainable Transport (OECD 2000) includes ‘a focus on reducing the number of long-distance trips and on much greater use of non-motorised means for short distance trips, with a large increase in the provision of supporting infrastructure for non motorized travel’ To achieve this goal, forms of travel such as walking, cycling and the use of public transport must be encouraged, in order to reduce motor traffic and its adverse effects and reduce dependency on fossil fuels.

The move to more sustainable transport has led to specific policy changes in London. A key aspect of the Mayor's Transport Strategy (Livingstone 2001) is to "make London one of the most walking friendly cities for pedestrians by 2015". The Central London Partnership (CLP) which represents businesses and Stakeholders in London, considers it essential to promote an increase in walking for short to medium length trips (2001). Transport for London (TfL), which is the strategic transport authority for the city, is tasked with developing a methodology to measure progress against these objectives.

In order to meet these objectives, Transport for London and the Central London Partnership commissioned a pedestrian modelling consultancy called Intelligent Space Partnership (ISP) to develop a model of total walking volumes for every node on the street network in Central London. This paper introduces the project and the framework for its policy implementation. The main aim
is to outline the new approach that has been adopted in developing a large scale pedestrian model, but some initial findings are also presented.

2. REQUIREMENTS OF THE PEDESTRIAN MODEL

The model has been designed to meet the requirements of a strategic transport authority and avoid some of the previous problems of large scale urban modelling (Lee 1972). The requirements and approach adopted are summarised as follows:

- The kind of model required is a comprehensive pedestrian demand model. Therefore the output variable was pedestrian flows (measured as average people per hour) on every node in the road network as defined by ordnance survey road centre lines.

- The modelling has to be based on standard, best practice statistical techniques. The approach adopted is the standard Multiple Regression Analysis (MRA) technique of statistical inference for modelling the relationship between a series of independent variables on a dependent variable (Rawlings 1988).

- The model has to be empirically tested from the start and capable of ongoing tests. The required tests show how well the predicted values correlate with observed values, identifying the accuracy of the model. This has been achieved by constructing the model with MRA using observed data and testing it against pedestrian flows in further observation studies.

- The model has to be open to scrutiny and the modelling results had to be independently verified by the transport authority. Using the standard technique of MRA makes this relatively straightforward, as all modelling data was provided for independent verification by TfL engineers.

- The model composition has to be extendible, to provide a flexible framework for testing the influence of different factors on pedestrian flows in the future. If new policies need to be tested in the future, these can be accommodated in the modelling process as new variables to be tested in the MRA.

- The model has to be applicable at an urban-wide scale in order to be useful as a policy evaluation tool. The modelled area is 25 square kilometres and incorporates all 7,526 street segments within this region. This is considered large for both pedestrian models and some road network models such as Saturn (Van Vliet 1982) designed for 100 to 150 intersections. This ruled out some of the microsimulation approaches (Hoogendorn et al. 2000) as these are not designed for large scale spatial systems. Scalability has been a key issue in the choice of all model components.

In summary, the approach adopted was to focus on the development of a flexible and testable pedestrian modelling framework using MRA. The plan for progressive improvement of the model is to implement continuous retrials for new areas or new policy levers, and to use independent validation by TfL. Modelling components developed in this project are also being used by TfL’s internal research programme on pedestrian modelling, which is in development phase. With every
extension of the model, the importance of all factors will be retested to ensure that the best balance between accuracy and applicability of the model design. The modelling process is summarised in figure 1 below.

![Figure 1: The Modelling Process](image1)

At time of writing, the pilot study of Phase 1 has been completed. Phase 2 is underway, which involves the extension of the model to a larger area of Central London, shown in figure 2 below.

![Figure 2: The Study Area](image2)
3. OBSERVATION STUDIES

The first phase of model testing was undertaken using data from three observation studies within the Central London study area. Two of these came from Intelligent Space’s consultancy projects (Duxbury and Desyllas, 2000 and 2001) and the other from an research project undertaken by UCL. Counts were made between 08:00 and 19:00 in the first two studies and between 12:00 to 18:00 in the UCL study. The dates of each study were March 2000, July 2001 and August 1999. According to Hillier (1992); ‘experience shows that weather has relatively little effect on natural movement’. However Gehl and Gemzøe (1996) have shown that seasonal variance of Copenhagen pedestrian traffic can be as great as 40% between summer and winter. Hence further investigation of the seasonal effect is needed and the results incorporated in development of the model.

The methodology adopted was similar to that of Benham et al. (1976). Pedestrian flows were sampled on every pavement in the study zones. Observation points were located where possible on both pavements at the mid point of each street segment. The pavement midpoints were grouped into circuits for sampling. Each observer walked a circuit, taking pedestrian flow counts at each pavement midpoint for 5 minutes in such a way that each mid-point was covered once for each hour of the survey. If the observer finished a circuit within one hour they were instructed to return to the circuit starting point and wait until the end of the hour before commencing the next count. All counting was done using mechanical counters and timed with a stopwatch. Flow rates per hour were derived by multiplying the 5 minute counts by 12.

Sampling flows in this way provided a scalable framework for future testing, as new sample locations can be added to the model easily. Counts can either be undertaken by an onsite team, or using automated counting systems with technologies such as passive infrared or CCTV (UTMC 2000), depending on the budget of the transport authority and the scale of the study. Furthermore using flow as an output variable rather than density reduces the statistical effects of localised pedestrian platooning identified by Pushkarev et al. (1975).

A key characteristic of sample data that the model has to capture is the marked variance with respect to direction: there are junctions in the study area that have flows averaging over 3,000 people per hour in one street and 0 people per hour in the perpendicular street. This fine scale anisotropic character of the movement network can be masked by aggregating the data, leading to the statistical error in modelling known as the ‘ecological fallacy’ (Robinson 1950). For these reasons, the modelling has been undertaken on a street segment level, meaning that the unit of the model is the flow down any street, including both pavements if the street is not pedestrianised. There was no attempt to average flows in any other way, such as along a line of sight (Hillier 1993).

Another key characteristic of observed movement the model has to replicate is the uneven distribution of flows, with a small number of streets exhibiting far higher flows than all others. Pedestrian flows are logarithmically distributed as can be seen in figure 3 below. A model therefore has to be able to identify the primacy of these key streets in order to represent the pattern successfully.
The total sample size for the first stage of model testing was 231 street segments. Previous models have used various levels of aggregation, leading to sample sizes of 20 Blocks (Benham 1976) 50 Lines of Sight Routes for Central London (Hillier 1993) and 600 Block Sectors (Pushkarev 1975). The framework for model development allows for continual retesting with any new observation data.

4. COMPONENTS OF THE MODEL

The components of the model that were tested break down into four key groups. The groups and the measures tested are shown below.

- **Capacity**: Pavement Width, Street Width, Carriageway Width.
- **Land Use**: Adjacent Retail, Adjacent Commercial, Adjacent Office, Adjacent Residential, Adjacent Public Buildings, Adjacent Vacant Buildings and Adjacent Parking. Percentage and count values were used for all.
- **Street Grid Configuration**: Ln Visibility, N2 Accessibility, Maximum Radial line of sight.
- **Transport Accessibility**: Tube Station Entrance, Accessibility (graph depth)

5. CAPACITY

A measure of footway capacity was included in the modelling to test the importance of this physical constraint on the distribution of flows. The influence of footway capacity has been well established since Fruin formalised the Level Of Service (LOS) concept for walkways (Fruin 1971),
showing that reduced capacity leads to congestion and thereby reduced flows. Pushkarev et al. (1975) found walkway area to be a significant independent variable in their Manhattan Study although it was not included in the tests by Hillier or Benham.

The capacity measure used for the London Pedestrian Model was the average effective footway width in metres per street segment. This was sampled from high resolution Ordnance Survey mapping data using an automated GIS algorithm. Figure 4 below shows a spectral range colour scale representation of the results, with red as the highest average pavement width and blue as the lowest.

![Figure 4: Pavement Width Analysis for Central London](image)

As well as testing for pedestrian capacity, vehicular capacity was also measured. Both the width of the carriageway and the total street width were tested to check for any influence between the pedestrian space and other spaces.

6. LAND USE

A measure of land use was included in the model in order to test the influence that different building uses have as 'trip generators' for pedestrian flows. Testing land use was central to the early pedestrian models of the 1970s following the tradition of land use transportation models by Lowry and McLughlin (Willumsen 1994) and large scale urban models of the 1960s (Lee 1973). Both the Benham (1976) and Pushkarev (1975) studies showed a strong correlation between a series of independent land use variables and pedestrian movement.

It was intended that official land use statistics would be used in the model from the first stage; these were used for both the American studies of Benham and Pushkarev studies. However, government statistics on use data are not easily available at disaggregated level in the UK and for phase 1 of the model development, it was necessary to resort to manual land use surveys for the
test areas. Entrances to buildings were marked on each street and the ground level land use was classified as follows: office; retail; commercial; education; vacant; public buildings and residential. This allowed for an initial testing of land uses in the model to evaluate how significant they might be, but was not a scalable solution for large scale modelling. For the next phase of the model, official land use statistics will be used and there are likely to be at least 2 national databases of land use more easily available in the next 2 years (NLUD 2002).

7. STREET GRID CONFIGURATION

Pedestrian models based on land use alone do not encompass differences in the layout of streets as movement networks. The concept of land use data employed to assign values of attractiveness to origins and destinations is clear, but it begs the question as to which streets pedestrians use to get between these different land uses. Street systems are complex and irregular, so any model that aims to predict pedestrian flows within them must be linked to street grid configuration in some way. Pushkarev attempted to account for street layout by including dummy variables for "street" versus "avenue" in his Manhattan study. In a city such as London, this simplification is too crude and a better measure is required.

Whereas the Benham and Pushkarev studies focussed on trip generation by attractors in the tradition of land use-transportation modelling, Hillier's study came from architectural research and started from a very different focus: the morphology of the street grid. Hillier suggested that "the city is a structure in which origins and destinations tend to be diffused everywhere, though with obvious biases towards higher density areas and major traffic interchanges. So movement tends to be broadly from everywhere to everywhere else. To the extent that this is the case in most cities, the structure of the grid accounts for much of the variation in movement densities" (Hillier 1996). This is a useful simplifying assumption: to the extent that land uses are evenly distributed, the layout of the street grid itself must influence the pattern of pedestrian movement and can therefore be used to predict flows on individual streets. Hillier suggested using measures of the street grid's spatial configuration as independent variables in the model of pedestrian flows in order to quantify this influence.

The aim for the London Pedestrian Model was to incorporate both a traditional trip generation component and also a component for street grid configuration. Although neither the earlier land use studies nor the Hillier configuration approach statistically tested the influence of both land use and configuration together in the regression analysis, the two approaches are not incompatible and have been combined in this model.

The two main methods for representing and analysing street grid configuration are Visibility Graph Analysis (VGA) first developed by Braaksma (1980) and the Axial Map developed by Hillier (1984). In a previous study, both representation techniques were tested against observation data and the VGA representation showed a better correlation (Desyllas and Duxbury 2001). VGA also offered a number of methodological advantages for use in large scale modelling, principally relating scalability and reproducibility as VGA can be fully automated whereas axial map representation requires a user to manually draw lines along each street.

The Fathom software application (Intelligent Space Partnership 2000) was used to provide an analysis of the visibility relations between sample locations in the street network. On the basis of high resolution mapping data (OS MasterMap), sample points were generated in areas classified as pedestrian space (all pavements and pedestrianised areas). The size of the sample was 302,000 points within the Central London area. A further 250,000 points had to be processed in a buffer area around the study zone, in order to remove the 'edge effect' from network accessibility measures (Desyllas and Duxbury 2001), giving a total sample of over half a million points.
The software produces an analysis of both the local and global characteristics of each sample point's location within the street grid configuration. A local configurational measure is the visibility of a sample point, sometimes referred to as visual field, viewshed or isovist area (Benedikt 1979). This measure identifies the area of pedestrian movement space in square metres that is visible from any sample location by testing for direct lines of sight from each point to all other points, using buildings and private spaces as occlusions. The visible area is calculated using the graph measure of modal degree (Wasserman 1994): the number of vertices that a node has in a visibility graph represents the number of other points that are visible (i.e. connected by line of sight) from a specific sample point. For the purposes of pedestrian movement, this measure highlights the difference between 'desire lines' or important visual links through the street grid and the more secluded back streets. A spectral range colour representation of this analysis is shown in figure 5 overleaf. Another local measure calculated by the software is the maximum radius of the visual field from each sample point. This identifies the maximum distance that a pedestrian can see from that location.

Global measures of the network accessibility of each sample location are also calculated. These identify how much of the movement network is reachable to a pedestrian within a given complexity of trip. The complexity of the trip is defined as the number of changes of direction that a pedestrian makes. If a pedestrian walks off from a given point in any direction and allows themselves only one further change of direction, the area reached can be defined in graph terms as the 'neighbourhood' accessible within two 'steps' of the graph (Wasserman 1994). In a visibility graph, each step of graph depth represents a move from one visible area to the area around the corner of an occlusion. For the London Pedestrian model, neighbourhood has been calculated to step 2.

Figure 5: Visibility Analysis of Central London
8. TRANSPORT ACCESSIBILITY

Other modes of transport are relevant to the modelling of walking volumes within the city because of the physical limits to the distances that individuals will walk on any trip. Although these limits vary greatly in different kinds of urban environment, most walking in London is within 0.96km (National Statistics 2001) and the practical limit for walking trips is 3.2km (Fruin 1971). In London, the underground is a key transport system distributing pedestrians around the street grid for longer distance trips, with over 1.9 million trips per day in the city as a whole and 1.1 million within the Central London area (LU 2001). Pedestrian flows are also important in the integration of different transport modes. Approximately 73% of rail trips, 79% of trips to the London underground and 50% of all bus trips involve at least one walk of 50m or more (National Statistics Update 2000).

Accessibility of public transport was shown as an important independent variable by Pushkarev (1975) who used the measured distance from the transit entrance. Pushkarev’s method for obtaining this distance is not clear, but it appears to be the Euclidean distance from a terminal entrance. For the London Pedestrian Model, a measure of the accessibility of underground stations within the street grid configuration has been developed using the visibility graph analysis. The Fathom software calculates the path complexity from each sample point to the nearest tube station entrance. In graph theory terms, this measure is the minimum ‘depth’ from any origin or set of origins, with each step of depth representing a visual connection around the occlusions of street blocks. A graphical representation of the analysis of tube accessibility is shown in figure 7 below. A spectral colour scale is used with red represents the most accessible (or lowest depth) spaces and blue represents the least accessible.
9. FUTURE ADDITIONS TO THE MODEL

The framework for pedestrian modelling has been designed to allow for additional factors to be included and a variety of possible policy levers are already being considered for inclusion. The most promising component for extension of the model at present is to include the variable of urban density. Previous research has shown that there is a clear, log-linear negative relationship between urban density and private transport energy use per capita, both comparing cities as a whole and comparing within urban regions (Newman and Kenworthy 1999). This suggests that higher densities result in higher pedestrian flows and this could itself be a significant variable within the pedestrian model. The real distinction brought out by previous research is shown between inner areas and suburban areas, and at present, the Pedestrian model only covers the dense inner urban area. However, extension of the model to the suburbs will necessitate testing a density measure.

Other components being considered for inclusion are:

- Traffic flows as used in Sketch Plan methods developed by Matlick (1996).
- Additional transport modes (such as buses, river transit, coach stations, light rail, trams and car parks).
10. INITIAL RESULTS

Standard quality control procedures have been used in the statistical modelling, such as checking the frequency distributions of each of the variables to ensure that all logarithmic distributions were transformed, checking that all coefficient signs were as expected and testing for autocorrelation of independents. Stepwise regression was used to identify the variables that were found to have a statistically significant influence on observed pedestrian movement. The stepwise regression resulted in the following model composition:

\[ \log(\text{flow}) = A \times \log(\text{average visibility within the street network}) + B \times (\text{accessibility to a London Underground station}) + C \times (\text{pavement width}) + D \times (\% \text{ of frontage that is retail}) + \text{constant} \]

where A, B, C, D are calibrated constants.

This gave an intermediate value F. A linear regression of \(\exp(F)\) was made against observed flows, and this gave a very impressive correlation of \(r^2 = .82\), shown in figure 8 below. This is a far higher correlation than any of the previous published urban pedestrian models; Benham's correlation of .764 was the closest, but with a much smaller sample of 20 cases (1976). The published correlations of the Pushkarev (1975) and Hillier (1993) studies were .61 and .55 respectively.

![Figure 8: Correlation between forecast and observed pedestrian flows in the London Pedestrian Model.](image)

The model has been independently verified by TfL engineers and the programme of testing against additional observation data has begun. An initial test has been completed with the existing 3 areas used in the pilot study. This was undertaken by recalibrating the model using the values for Regent Street and Shoreditch Triangle to predict values for Tottenham Court Road (which was not used in model calibration). All the same factors were shown to be significant, with some trivial changes to the coefficients. The result of this model was a correlation of \(r\)-squared of .74, as shown in figure 9 below. For the next stage of testing, pedestrian movement data for 3 further study areas within the central London area have been collected by TfL, containing a total of 180 additional sample locations.
As the model based on two areas and validated against the third area shows, the general shape of the model holds up well to scrutiny against data which has not been used to calibrate the model. However, the scale of the model did not match as well, so although we obtained a very strong r-squared value, the absolute values differed significantly. We expect that a large part of this difference can be attributed to differences in urban density between the areas (both overall density and the density of particular land uses). We plan to incorporate density in the next stage of the modelling in order to capture these differences.

These initial findings are raising some interesting questions for theory. For example, why is visibility the most important variable, and not a more global accessibility measure of the street grid configuration? Given that pedestrian trips are generally short, it may be that visibility is the most appropriate spatial measure of the street network for capturing the aggregation of short journeys. Another assumption would be that simpler routes tend to use the more visible spaces and pedestrians are economising on route complexity, not distance. Visibility may be also considered the most direct aid to wayfinding, especially as pedestrians have imperfect knowledge of the street network and therefore pedestrians are picking out the most obvious routes, which are also the most visible.

11. THE FUTURE OF TRANSPORT MODELLING

It is hoped that by presenting the development of the London Pedestrian Model, we have demonstrated that it is both possible and beneficial for planners and engineers to move towards a more balanced view of transport modelling that encompasses both motorised and non-motorised modes. This means doing the necessary data gathering and modelling work for non-motorised modes, as well as that which has been long established as a requirement for vehicles. The London Pedestrian Model is an ongoing project and the longer term aim is to extend the coverage to Greater London and to incorporate additional policy levers. This will facilitate its use as a transport planning and evaluation tool. When transport planners have the research and analysis tools to monitor and predict pedestrian flows as an integral part of urban transport, decision makers will be able to act upon a far more balanced understanding of the functioning of streets in our cities and develop policies that are beneficial to all users.
ACKNOWLEDGEMENTS

This paper describes research that was part-funded by Transport for London and Central London Partnership. We are grateful to the EPSRC VR Centre for the Built Environment Grant at University College London which provided the support which produced pedestrian flow data from the Regent Street research project in this study. Many thanks, (in alphabetical order) to Keith Harris, Sam Hatoum, Frank Hebbert and Julie-Ann Jankó at Intelligent Space Partnership for their work on the model components and graphics used in this paper. Includes map data licensed from Ordnance Survey with the permission of the Controller of Her Majesty’s Stationary Office. © Crown copyright 2003. All rights reserved.

REFERENCES


Hillier, B., Penn, A., Hanson, J., Grajewski, T. and Xu, J. (1993) Natural Movement: or, configuration and attraction in urban pedestrian movement, Environment and


