Modelling Railway Station Choice

Marcus Young

UNIVERSITY OF Southampton
Transportation Research Group

Improving railway station demand forecasting models using probability-based catchments.

The problem: demand models for new stations use unrealistic catchments

- Trip end models are primarily used to forecast demand for new stations in the UK. Total entries/exits per year are a function of a range of variables, such as population, jobs and service frequency.
- Need to define a station catchment – the unit of aggregation for model variables. Typically by assigning zones to their nearest station (Fig. 1).
- Assumes choice of station is deterministic – all residents in a zone will always use the zone's nearest station. Catchments are discrete.
- In reality catchments are much more complex – they overlap and stations compete (Fig. 2).

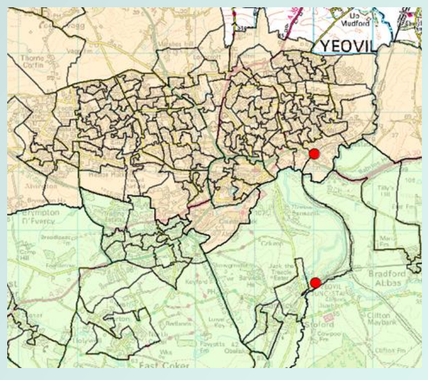

Fig.1 – theoretical catchments - zones assigned to nearest station

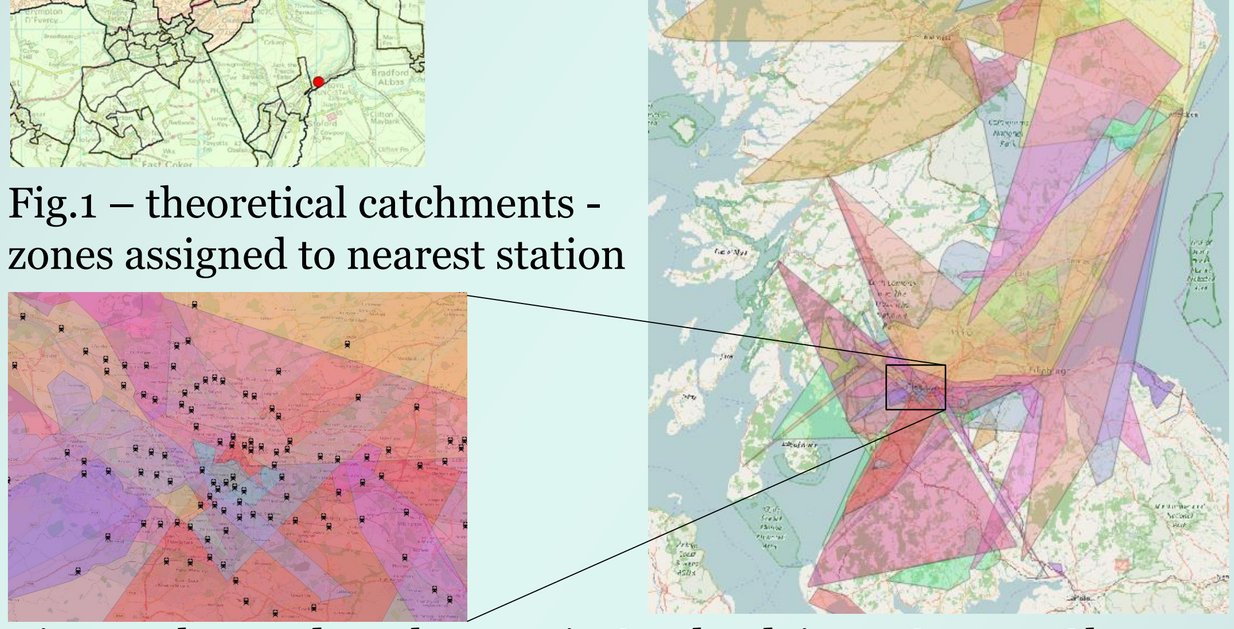

Fig.2 – observed catchments in Scotland, inset Greater Glasgow

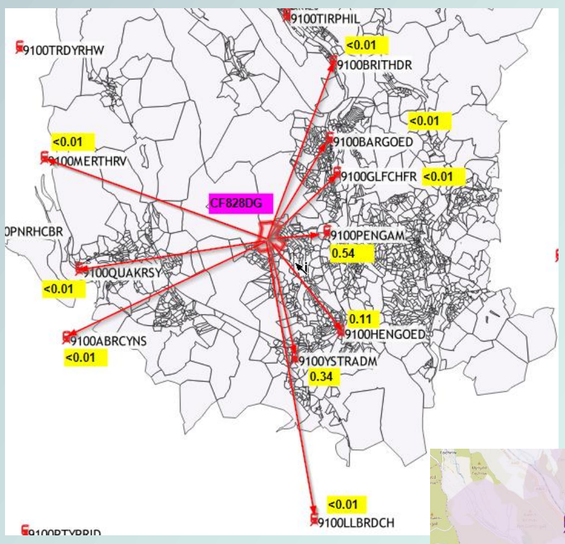

Fig 3 – station probabilities calculated for a single unit postcode

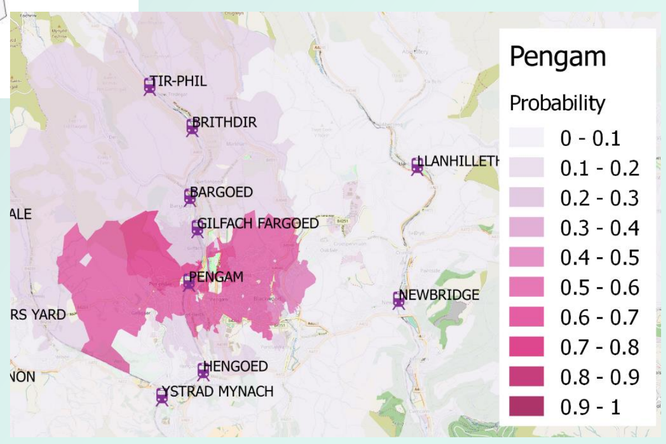

Fig 4 – example of a probability-based catchment. Each postcode has a probability of the station being chosen

An alternative approach: deriving probabilistic catchments

- Discrete choice models were calibrated using trip data from passenger surveys carried out in Scotland and Wales. The ten nearest stations to each origin postcode, plus the nearest major station if not present, formed the choice set.
- A data processing framework, consisting of a PostgreSQL spatial database, the R software environment and an instance of OpenTripPlanner, was developed to automate generating explanatory variables from a range of open transport data.
- The calibrated models have a very good fit and predict station choice substantially better than the base model (which assumes the nearest station has a probability of one).
- Using the station choice model it is possible to calculate station probabilities for each postcode (Fig. 3) and generate a probability-based catchment for a station (Fig. 4).

Application: integrating station choice into trip end models

- Trip end models - using both deterministic and probabilistic catchments - were calibrated for all category E and F stations in GB, using estimated entries and exits in 2011/12 as the dependent variable.
- When defining a deterministic catchment for a station, the population of each postcode which has that station as its nearest is weighted by a distance decay function and this is summed for all applicable postcodes.
- When defining a probabilistic catchment, the population of each postcode which has the station in its choice set is weighted by the probability of the station being chosen and then by the distance decay function.
- The models were used to forecast demand for three recently opened stations (Table 1). The probabilistic catchments resulted in substantive 'corrections' to model predictions, that were consistent with observed station demand.

Station	Forecast (deterministic)	Forecast (probabilistic)	% diff.	Actual		
				2013/14	2014/15	2015/16
Conon Bridge	23,840	20,722	- 13%	18,114	15,510	15,276
Fishguard & Goodwick	14,064	14,205	+ 1%	17,062	19,874	19,946
Energlyn & Churchill	46,554	59,394	+ 28%		69,390	74,206

Table 1 – demand forecasts (total entries/exits) for new stations (2011/12 base year)

EPSRC
Engineering and Physical Sciences Research Council

Marcus Young
Transportation Research Group, Engineering and the Environment,
Boldrewood Innovation Campus, University of Southampton,
SO16 7QF
Email: m.a.young@soton.ac.uk