UrbanSim at the EPFL

The Brussels Experience

Zachary Patterson

Transport and Mobility Laboratory, EPFL

European UrbanSim Users Workshop, ETH Zurich

17 March 2008

Starting Point

- Integrated transportation land-use modeling strategic area for research
- Many models to choose from

Why we were interested in UrbanSim:

- disaggregate (Discrete Choice Models), open-source, dynamic equilibrium
- used more and more

But:

- disaggregate approach → steep data requirements
- → **Just how hard is it to use?**

Our Approach

- Start simple
- Limited resources
- Limited data
- How far can you get?
- Evaluate for more in depth application

Test Project: Brussels

- Research partnership in Brussels
- Previous TRANUS model for the Brussels region
- Idea: use TRANUS data to develop UrbanSim model
- Master's project

TRANUS Data:

- For each of 152 zones for 2001 and 2015:
  - Employment by industrial sector
  - Households (coarse demographic information)
  - Land-prices
  - Road infrastructure
  - Travel impedance between zones
- Historical data on job and population changes

Lessons Learned - Initial Brussels Model

Possible in a few months:

- Get a working understanding UrbanSim
- Develop a model that ‘works’ (even with aggregate data)

Main challenges faced:

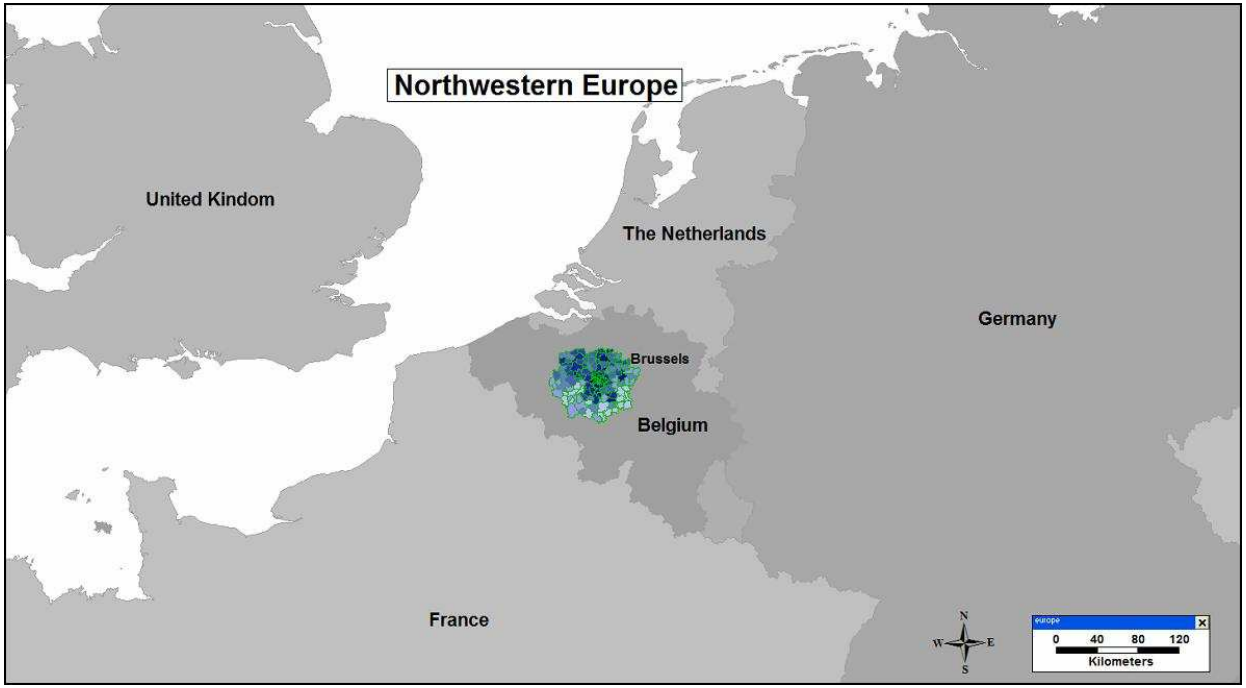
- Obtaining data
- Preparing data
- Lack of data effects quality of submodels
  - Lack in terms of spatial precision
  - Lack of historical data particularly problematic

Initial Brussels Project

Approach:

- Use Eugene example as model
- Use it to learn and evaluate model requirements
- Adapt it to be used with Brussels data
- Adapt Brussels data to be compatible with it

Study Region - Brussels, Belgium



- 1.3 million households
- 4,300 km^2
- 192,000 gridcells

Baseyear Data Preparation - 1

- Development type randomly assigned to gridcells
- One building for each developable gridcell
  - building size accommodated number of jobs or households
  - accounted for vacancy rate (function of distance from CBD)
  - jobs and households assigned to buildings (and gridcells)
- Historical construction data:
  - Based on population and employment variation by zone (1991-2001)
  - Buildings with appropriate number of jobs and households randomly selected as having been built over period
- Synthetic household characteristics assigned to households

Baseyear Data Preparation - 2

Transport data:

- Proximity to infrastructure with TransCAD
- Transport Model:
  - No stand-alone model
  - Developing transport model outside scope of initial project
  - Interzonal travel impedances available from TRANUS
  - Same impedances used for all simulation years

Vacancy, job and household relocation rates:

- hypotheses from Stratec

Exogenous data:

- Total jobs...
- Total households...
- Extrapolated from TRANUS data

Submodel Results

- ‘Skeletal’ structure, based on Eugene example
- Coarse household location choice model
- 3 employment location choice models
- 3 real estate development choice models
- Coarse land-price model
- Residential models shown for brevity

Household Location Choice Model

Variable	Coefficient	Std. Error	t-value
Cost to Income Ratio	-0.064	0.025	-2.599
% high income if high income	0.032	0.001	32.800
% low income if low income	0.059	0.001	60.167
% mid income if mid income	0.027	0.001	26.517
Travel Time to CBD	4.106E-04	0.000	2.796
Likelihood Ratio Test:	0.000		
Number of observations :	129269		

- Meet *a priori* expectations except CBD distance
- Pleasantly surprising

Residential Real-estate Development Model

Variable	Coefficient	Std. Error	T-value
Log commercial surface w.w.d.	-0.313	0.048	-6.569
Log Access to population	0.247	0.142	1.745
Log industrial surface w.w.d.	-0.118	0.019	-6.193
Log total population w.w.d.	0.447	0.054	2.765
Travel Time to CBD	0.006	0.002	3.111
Average Income w.w.d.	-0.217	0.045	-4.837
Likelihood Ratio Test :	0.000		
Number of observations :	1332		

- Problem variables: CBD distance, income, price
- Pleasantly surprising

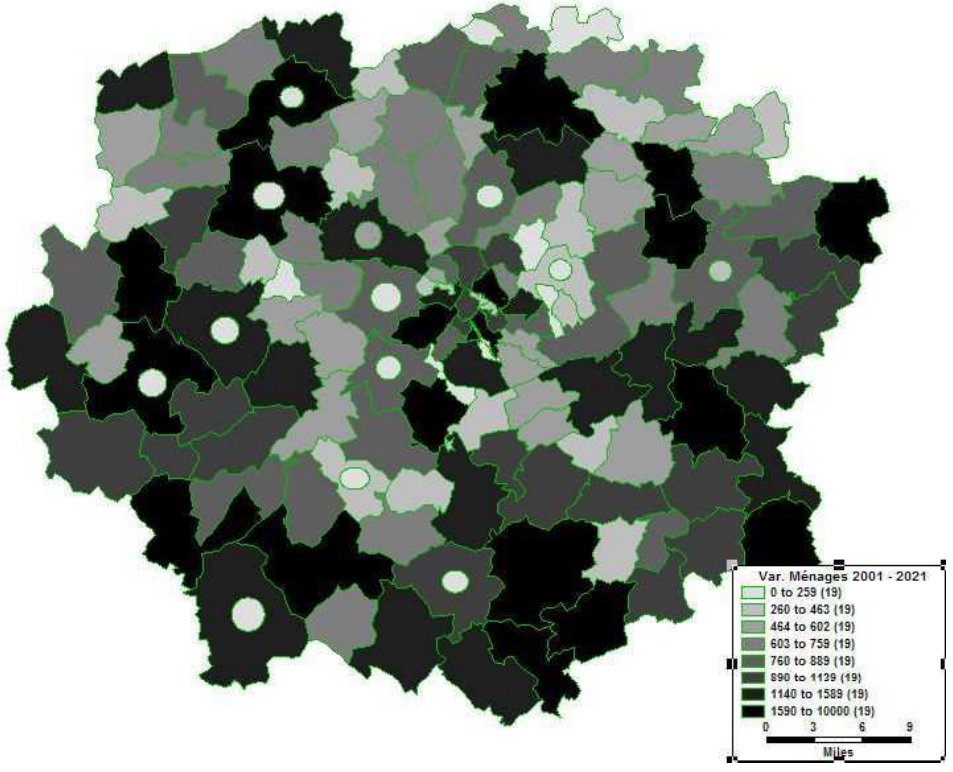
Land-price model

Variable	Coefficient	Std. Error	T-value
Log total empl. basic sect. w.w.d.	-0.239	0.002	-110.714
Log Access to empl.	0.770	0.001	660.032
Log residential units	0.116	0.001	113.787
Log total empl. w.w.d.	0.461	0.003	170.515
Log Access to population	0.014	0.002	6.678
% High income w.w.d.	0.001	0.000	5.168
Travel Time to CBD	-0.001	0.000	-11.291
Number of observations :	165780		
Adjusted R-Squared :	0.473		

- No vacancy rate variable
- Pleasantly surprising

Simulation Results - 1

Change in Households 2001-2021 with Population Growth



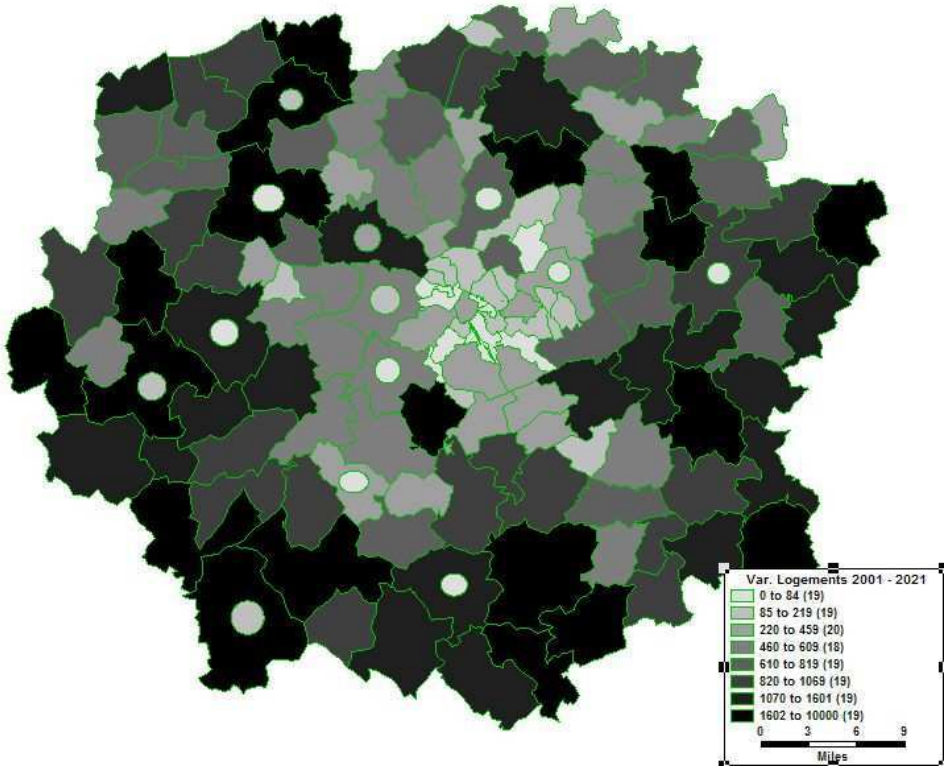
- Population increasing primarily away from center

TRANSP-OR

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

Simulation Results - 2

Change in Residential Units 2001-2021



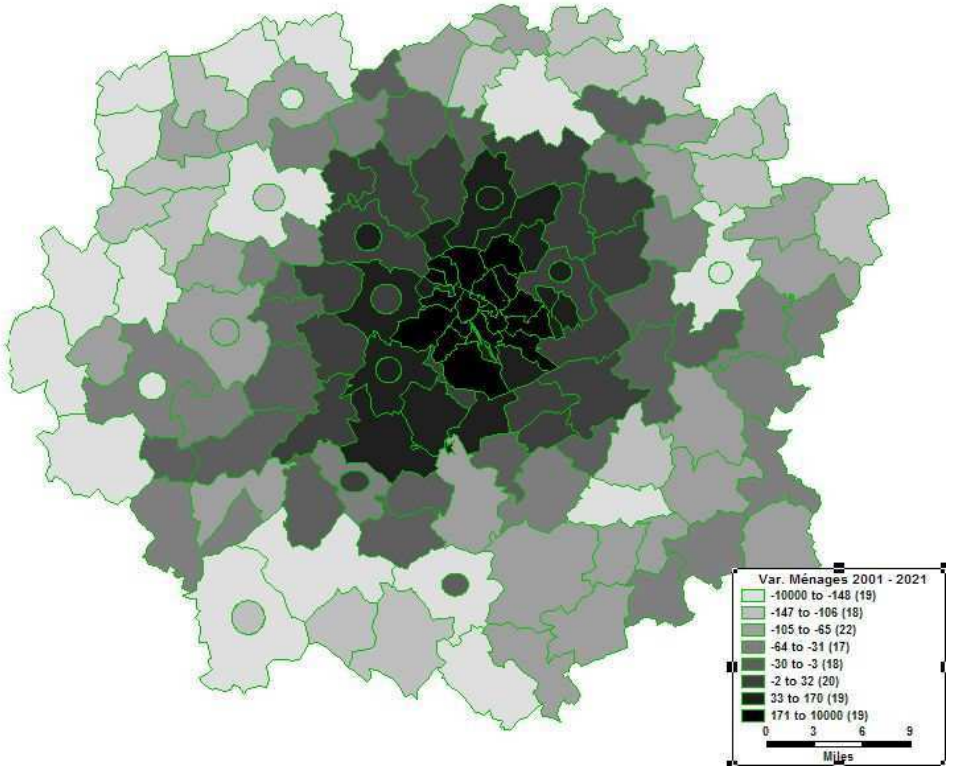
- Residences built outside of center
- Vacancy rates low outside center

TRANSP-OR

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

Simulation Results - 3

Change in Households 2001-2021 without Population Growth



- With no construction...
- ...households choose downtown.

TRANSP-OR

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

Simulation Results - What is Happening

1. Houses are being built:
   - Away from center
   - Regardless of land price (development model)
2. Households locating away from center where houses built
3. Weakness in development models:
   - Not enough observations?
   - Not enough variation in land-price?
4. It seems disaggregation not as critical as lack of:
   - Historical construction data
   - Good vacancy rate data?
   - Zoning?

Brussels - Part II

After this promising start, more challenging:

- Land-use data available, but...
- Continued transport data availability problems
  - Difficult to obtain data (partner too busy)
  - False starts: data made available but inadequate
    - Transport model provided incomplete and insufficient

Change of strategy:

- With lessons learned from first model
- Seek application environment with more easily available data (transport and land-use)
- In meantime, continue work on Brussels when resources available

Brussels - Incorporating land-use data

Two master's students (in math) for four months:

- Use land-use data to attribute actual development types to gridcells
- Use existing R-code to re-assign buildings, jobs and households
- Recalibrate models and run new simulations

Brussels - Incorporating land-use data

Challenges:

- Land-use data
  - Availability of data
  - Adequacy of data (e.g. commercial development type)
  - Consistency of data across study region
- Learning various software and languages proved challenging
  - GIS, R, MySQL, Python, UrbanSim

Result:

- Relatively successful incorporation of land-use data
- Recalibration of submodels $\rightarrow$ successful
  - but models suffer from (mostly) same problems as before
- Problems running model simulations (error in residential units)

Household Location Choice Model

Variable	Coefficient	Std. Error	T-value
Log total empl. w.w.d	0.124	0.0332	3.730
Travel Time to CBD	-0.013	0.0024	-5.170
Cost to Income Ratio	-9.08	0.010	-90.990
% low income if low income	0.238	0.0179	13.280
Final Log-likelihood is :	-4 247		
Log-likelihood ratio is :	0.99		
Number of observations :	129 269		
Convergence statistic is :	0.0009		

Residential Development Choice Model

Variable	Coefficient	Std. Error	T-value
Average Land Val. w.w.d.	-0.070	0.095	-0.740
Log commercial surface w.w.d.	0.125	0.026	4.890
Log Access to population	0.149	0.143	1.040
Log industrial surface w.w.d.	0.035	0.008	4.600
Log total employment w.w.d.	-0.539	0.047	-11.480
Log total population w.w.d.	0.711	0.039	18.240
Travel Time to CBD	0.013	0.002	7.080
Final Log-likelihood is :	-5 128		
Log-likelihood ratio is :	0.11		
Number of observations :	1 697		
Convergence statistic is :	0.00037		

Land Price Model

Variable	Coefficient	Std. Error	T-value
Constant	0.557	0.206	2.701
Dev. Type Industrial	-0.073	0.003	-21.912
Near Artery	0.060	0.003	20.071
Near Highway	0.085	0.007	12.069
Ln Basic Employment w.w.d.	-0.038	0.001	-32.254
Ln Comm Surface w.w.d.	0.230	0.002	116.820
Ln Service Sector Employment w.w.d.	0.031	0.002	14.012
Ln Total Employment w.w.d.	0.076	0.003	24.356
Ln Total Population w.w.d.	0.362	0.007	50.413
Ln Work Access to Employment	0.117	0.009	13.214
TT CBD	-0.010	0.000	-96.412
Number of observations:	165780		
Adjusted R-Squared:	0.4711068		

Lessons Learned

- There are limits to what can be achieved with aggregate data
- But remains useful, low-cost exercise
- From the user-side:
  - Using UrbanSim can seem a big investment
  - Would be good to be able to try it with easily available (e.g. aggregate) data
  - Our experience with disaggregating data:
    - can be difficult to do it in consistent manner
    - would be useful to have consistent method to disaggregate common data easily in UrbanSim
  - Remains difficult for newcomers to master all tools necessary
  - User guide more for programmers and not ‘users’
  - Efforts on user guide could prove helpful for attracting users

Next Steps

- In Lausanne:
  - Existing transport models (EMME and TransCAD)
  - Extensive land-use data
- Given experience with test case of UrbanSim for Brussels...
- ...have obtained funding for a prototype model for Lausanne
- Work has started on data collection and preparation
- Team of five