Workshop on Modelling Population 24/7

26 January 2010
Welcome, introductions and context
Programme for day

• 10.30 Welcome, introductions and context
• 11.30 Guest presentations
• 12.30 Pop24/7 project overview
• 13.00 Lunch
• 13.45 Algorithms, data and demonstration
• 15.30 Applications, challenges and future work
• 16.30 Close
About ourselves

- David Martin
- Samantha Cockings
- Samuel Leung
- Population surface models
- Geo-refer: geographical referencing for social scientists
- Census output areas and workplace zones
Housekeeping

- Refreshments
- Toilets
- Emergencies
- Presentations and questions
Population24/7 - context

• ESRC-funded project 1 March 2009-28 Feb 2011
• Space-time population modelling “24/7/52”
• Drawing on census, mid-year population estimates
• Taking advantage of new data sources
• Based on existing population surface modelling tools
• Aiming to develop a generalizable methodology
Introduction

- Importance of small area population mapping
- Deficiencies of current “night-time” approaches
- Advantages of gridded population models: esp. stability over time and reconstruction of settlement geography
Importance of small area population mapping

- Resource allocation large areas > small areas
- Targeting services/marketing
- Site location decisions/transportation demand
- Emergency planning
- Appropriate denominator population counts
Deficiencies of current “night-time” approaches

• Limited data on workplaces, travel to work, student residences
• No data on other activity patterns
• Frozen in time - census year/reference date
• Frozen in space – residential base
Photos:

David Martin
Advantages of gridded population models

• Stability over time
  – So very important if many different times represented

• Reconstruction of underlying settlement and neighbourhood geography/model is not space-filling
  – e.g. Distinction between commercial and residential neighbourhoods

• Integration with other modelled sources – esp. environmental

• Computational efficiency
OA populations: 2001 census
“One of the most important and difficult problems now facing city planners is the development of accurate, usable techniques for estimating the current daytime population of census tracts in urban areas”

Why? Reasons included modelling the location and size of bomb shelters and the potential casualties resulting from a nuclear attack

Challenges to be addressed

• Night-time grid on its own is not the solution, but can grid be harnessed to this problem?

• Variety of new data sources becoming available but no integrating modelling framework – data review and integration

• Review contemporary studies of the issues

• Develop extensible methodology for constructing 24/7 population models
Aims of today’s workshop

• Share common interests in this problem space
  – Find out more about what others have been doing
• Demonstrate project work in progress
• Seek feedback/suggestions/extensions/uses
• What can we most usefully do from here/are there opportunities for working together?
• Networking opportunity for diverse communities with a common interest
Guest Presentations
Guest presentations

- 24-hour Edgware Road
- National Population Database – Staffordshire/HSE
Population 24/7
project overview
Observations...

• This entire sub-branch of geographical and cartographic enquiry is essentially concerned with mapping the population in bed

• We have seen enormous advances in geovisualization techniques, computing power and dynamic modelling sophistication

• We have not adequately tackled the entire area of time-specific population modelling
Space-time population modelling

- ~99% of work based on night-time; ~0.5% daytime?

- Numerous motivations for time-specific models: emergency planning, transportation, business location, etc.

- General approach is to start with night-time population map and transfer population subgroups to specific daytime locations, e.g. schools, workplaces

- Longstanding difficulty of obtaining data with sufficient space/time resolution

- In reality, many different timescales to be modelled
Examples

- Landscan USA - Bhadhuri et al. (2007)
- Helsinki - Ahola et al. (2007)
- UK Health and Safety Executive – Smith and Fairburn (2008)
LandScan USA

Harris County & the city of Houston

http://www.ornl.gov/sci/gist/landscanUSA/landscanUSA_factsheet_ORNL.pdf
Modelling framework
Data sources - residential

- Census- or register-based, using “usual place of residence”
- Residence definition – equivalent to night-time population locations, students counted at term-time residence
- Decennial update interval – long-term population change
- UK census at output area level (OA, pop ~300) and official mid-year estimates (MYEs) at Lower Super OA level (LSOA, pop ~1500)
- Some workplace data from census, but not all people have workplaces and not all workplaces are daytime locations
Data sources – non-residential

- New administrative sources esp. from government, Neighbourhood Statistics Service (NeSS)
- Huge growth in availability and frequency since 2001 census
- Annual Business Inquiry dataset (employers, employees)
- Schools, hospitals, visitor attractions, long-distance visitor numbers, transportation flows
- Indirect measures of retail, leisure activity, points of interest
- Can all be related to NeSS geography hierarchy, NSPD
Mapping population to activities/places...

- Total population
  - Residential
    - Private dwellings
    - Communal ests.
  - Non-residential
    - Education
    - Employment
    - Temp accomm.
    - Healthcare
    - Family/social
    - Retail
    - Leisure
    - Tourism
    - Generalized local
  - Transport
    - Road
    - Rail
    - Metro/subway
    - Air
    - Water

...and further subdivisions
What do current maps cover?

- Total population
  - Residential
    - Private dwellings
    - Communal ests.
  - Non-residential
    - Education
    - Employment
    - Temp accomm.
    - Healthcare
    - Family/social
    - Retail
    - Leisure
    - Tourism
    - Generalized local
  - Transport
    - Road
    - Rail
    - Metro/subway
    - Air
    - Water
Data sources for all non-residential

Total population

+/- external visitors

Residential

Private dwellings
Communal ests.

Census, MYE
Census, MYE

Non-residential

Education 24%
Employment 49%
Temp accomm.
Healthcare
Family/social
Retail
Leisure
Tourism
Generalized local

NeSS, EduBase
Census, ABI, QLFS
VisitBritain, ABI
HES
VisitBritain
ABI, commercial
DCMS, ALVA, etc.
DCMS, ALVA, etc.
ID2007, NeSS

Transport

Road
Rail
Metro/subway
Air
Water

DfT, AADF
National Rail
TfL, etc
CAA
TfL
Modelling framework

• Builds on Martin (1989, 2006), Martin et al. (2000) currently implemented in SurfaceBuilder program

• One of a variety of methods for reallocation of population counts onto a series of geographical features

• Akin to dasymetric models where known population counts are allocated to most likely set of spatial locations

• Adaptive kernel estimation, treating each centroid as a high information point.

• Weight each cell to receive population reallocated from local centroids, hence volume preserving
Extension to the spatio-temporal problem

<table>
<thead>
<tr>
<th>Spatial centroid:</th>
<th>e.g. census output area</th>
</tr>
</thead>
<tbody>
<tr>
<td>– Population count</td>
<td>– Census population</td>
</tr>
<tr>
<td>– Spatial extent</td>
<td>– Modelled</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Time-space centroid:</th>
<th>e.g. primary school</th>
</tr>
</thead>
<tbody>
<tr>
<td>– Population capacity</td>
<td>– Pupil numbers</td>
</tr>
<tr>
<td>– Spatial extent</td>
<td>– Small (one cell)</td>
</tr>
<tr>
<td>– Time profile</td>
<td>– Term dates, school day</td>
</tr>
<tr>
<td>– Area of influence</td>
<td>– Catchment area (modelled time/space)</td>
</tr>
</tbody>
</table>
Time profile example – school

Population

In transit

Present

00  06  12  18  00   Time of day
<table>
<thead>
<tr>
<th>Location</th>
<th>July 2009</th>
<th>July 2008</th>
<th>Change</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aberdeen</td>
<td>287 491</td>
<td>314 355</td>
<td>-9</td>
</tr>
<tr>
<td>Barra</td>
<td>1 403</td>
<td>1 328</td>
<td>6</td>
</tr>
<tr>
<td>Belfast City (George Best)</td>
<td>245 186</td>
<td>232 733</td>
<td>-4</td>
</tr>
<tr>
<td>Belfast International</td>
<td>538 311</td>
<td>633 620</td>
<td>-15</td>
</tr>
<tr>
<td>Benbecula</td>
<td>2 976</td>
<td>3 236</td>
<td>8</td>
</tr>
<tr>
<td>Birmingham</td>
<td>969 219</td>
<td>1 036 104</td>
<td>-17</td>
</tr>
<tr>
<td>Blackpool</td>
<td>39 638</td>
<td>53 888</td>
<td>-26</td>
</tr>
<tr>
<td>Bournemouth</td>
<td>88 749</td>
<td>110 037</td>
<td>-17</td>
</tr>
<tr>
<td>Bristol</td>
<td>632 337</td>
<td>655 538</td>
<td>-17</td>
</tr>
<tr>
<td>Cambridge</td>
<td>268</td>
<td>17</td>
<td>1457</td>
</tr>
<tr>
<td>Campbeltown</td>
<td>1 018</td>
<td>826</td>
<td>23</td>
</tr>
<tr>
<td>Cardiff Wales</td>
<td>189 896</td>
<td>234 444</td>
<td>-19</td>
</tr>
<tr>
<td>City of Derry (Eglinton)</td>
<td>40 134</td>
<td>41 288</td>
<td>-29</td>
</tr>
<tr>
<td>Coventry</td>
<td>-</td>
<td>43 508</td>
<td></td>
</tr>
<tr>
<td>Doncaster Sheffield</td>
<td>92 841</td>
<td>103 326</td>
<td>-10</td>
</tr>
<tr>
<td>Dundee</td>
<td>6 898</td>
<td>7 097</td>
<td>-1</td>
</tr>
<tr>
<td>Durham Tees Valley</td>
<td>28 355</td>
<td>71 903</td>
<td>-61</td>
</tr>
<tr>
<td>East Midlands International</td>
<td>503 889</td>
<td>583 853</td>
<td>-14</td>
</tr>
<tr>
<td>Edinburgh</td>
<td>956 199</td>
<td>905 309</td>
<td>-5</td>
</tr>
<tr>
<td>Exeter</td>
<td>90 736</td>
<td>105 677</td>
<td>-15</td>
</tr>
<tr>
<td>Glasgow</td>
<td>819 309</td>
<td>942 906</td>
<td>-14</td>
</tr>
<tr>
<td>Gloucestershire</td>
<td>2 050</td>
<td>1 855</td>
<td>10</td>
</tr>
<tr>
<td>Humberside</td>
<td>34 160</td>
<td>48 992</td>
<td>-30</td>
</tr>
<tr>
<td>Inverness</td>
<td>61 507</td>
<td>73 918</td>
<td>-17</td>
</tr>
<tr>
<td>Islay</td>
<td>2 765</td>
<td>2 821</td>
<td>-22</td>
</tr>
<tr>
<td>Isle of Scilly (St. Mary's)</td>
<td>16 608</td>
<td>16 006</td>
<td>4</td>
</tr>
<tr>
<td>Aberdeen</td>
<td>287 491</td>
<td>314 355</td>
<td>-9</td>
</tr>
<tr>
<td>Barra</td>
<td>1 403</td>
<td>1 328</td>
<td>6</td>
</tr>
<tr>
<td>Belfast City (George Best)</td>
<td>245 186</td>
<td>232 733</td>
<td>-4</td>
</tr>
<tr>
<td>Belfast International</td>
<td>538 311</td>
<td>633 620</td>
<td>-15</td>
</tr>
<tr>
<td>Benbecula</td>
<td>2 976</td>
<td>3 236</td>
<td>8</td>
</tr>
<tr>
<td>Birmingham</td>
<td>969 219</td>
<td>1 036 104</td>
<td>-17</td>
</tr>
<tr>
<td>Blackpool</td>
<td>39 638</td>
<td>53 888</td>
<td>-26</td>
</tr>
<tr>
<td>Bournemouth</td>
<td>88 749</td>
<td>110 037</td>
<td>-17</td>
</tr>
<tr>
<td>Bristol</td>
<td>632 337</td>
<td>655 538</td>
<td>-17</td>
</tr>
<tr>
<td>Cambridge</td>
<td>268</td>
<td>17</td>
<td>1457</td>
</tr>
<tr>
<td>Campbeltown</td>
<td>1 018</td>
<td>826</td>
<td>23</td>
</tr>
<tr>
<td>Cardiff Wales</td>
<td>189 896</td>
<td>234 444</td>
<td>-19</td>
</tr>
<tr>
<td>City of Derry (Eglinton)</td>
<td>40 134</td>
<td>41 288</td>
<td>-29</td>
</tr>
<tr>
<td>Coventry</td>
<td>-</td>
<td>43 508</td>
<td></td>
</tr>
<tr>
<td>Doncaster Sheffield</td>
<td>92 841</td>
<td>103 326</td>
<td>-10</td>
</tr>
<tr>
<td>Dundee</td>
<td>6 898</td>
<td>7 097</td>
<td>-1</td>
</tr>
<tr>
<td>Durham Tees Valley</td>
<td>28 355</td>
<td>71 903</td>
<td>-61</td>
</tr>
<tr>
<td>East Midlands International</td>
<td>503 889</td>
<td>583 853</td>
<td>-14</td>
</tr>
<tr>
<td>Edinburgh</td>
<td>956 199</td>
<td>905 309</td>
<td>-5</td>
</tr>
<tr>
<td>Exeter</td>
<td>90 736</td>
<td>105 677</td>
<td>-15</td>
</tr>
<tr>
<td>Glasgow</td>
<td>819 309</td>
<td>942 906</td>
<td>-14</td>
</tr>
<tr>
<td>Gloucestershire</td>
<td>2 050</td>
<td>1 855</td>
<td>10</td>
</tr>
<tr>
<td>Humberside</td>
<td>34 160</td>
<td>48 992</td>
<td>-30</td>
</tr>
<tr>
<td>Inverness</td>
<td>61 507</td>
<td>73 918</td>
<td>-17</td>
</tr>
<tr>
<td>Islay</td>
<td>2 765</td>
<td>2 821</td>
<td>-22</td>
</tr>
<tr>
<td>Isle of Scilly (St. Mary's)</td>
<td>16 608</td>
<td>16 006</td>
<td>4</td>
</tr>
</tbody>
</table>
Treatment of centroid $i$ at time $t$
Implementation, data and demonstration
Centroids, boundaries and grids

Left: centroid locations and boundaries; Right: centroid populations redistributed onto grid.
Distance decay function

![Diagram showing the distance decay function with weighting and distance axes.](image)
Persons per cell (200m x 200m)

gbpop91
VALUE
- 0 - 33
- 33-46
- 46-79
- 79-160
- 160-359
- 359-848
- 848-2,051
Overview of Southampton study area
Centroid set

- 1696 census OAs
- 3329 workplaces
- 211 schools and colleges
- 2 universities
Background layer: transportation network
Background layer

- Population “capacity”
  - Land use
  - Land cover
  - Pop density
  - Transport
Basic time-space interpolation algorithm

- Specify study area $a$ and time $t$
- Identify background layer $b$ (cells that can contain population) for time $t$
- Adjust for external visitors in/out of $a$ at time $t$
- Sum all residential centroids to obtain population $P$
- Examine each centroid $i$ to obtain populations $p$ in local extent $d$ and area of influence $j$ at time $t$
- Redistribute $P$ across $d$ and $j$, constrained by $b$
Early results

• Southampton, UK as test area

• Using existing model with pre-prepared data extracts for specific time slices (SurfaceBuilder program)
  – Does not require full time profiles for each centroid
  – Does not take background layer into account, hence no population is allocated into transport layer

• Time-space interpolation program currently being written in .Net using existing and new code components

• Preliminary validation against known patterns
Southampton, 200m cells

02:00

Residential “night-time” model
Southampton, 200m cells

08:00

Early workplaces, docks, industrial estates; rest as residential
Southampton, 200m cells

09:00

Workplaces, educational institutions, “daytime” model; low residential; very high central densities
Southampton, 200m cells
18:00
Late workplaces remain, education closed; return to residential; high central densities
Initial visualization of output

Output to kml

Multiple layers overlaid in Google Earth

3D navigation and exploration

Time slider allows time sequence to be “played”
(Schmitt, 1956, p. 83).

- “One of the most important and difficult problems now facing city planners is the development of accurate, usable techniques for estimating the current daytime population of census tracts in urban areas”

- Why? Reasons included modelling the location and size of bomb shelters and the potential casualties resulting from a nuclear attack

Conventional population data sources

- Census- or address register-based
- Residence definition – equivalent to nighttime population locations
- Long update intervals
Software tool SurfaceBuilder24/7
Applications, challenges and future work
Questions, discussion.