

Time-specific population modelling using new forms of data

David Martin SmartPop Seminar/Brussels/11 October 2017



Overview

- Background to Population 24/7 projects: the problem
- Introduction to population24/7 approach and surfacebuilder247
- Integrating near real-time (NRT) population sources and estimates



Background to Population24/7 projects: the problem

The problem ... A typical small area population map



SOUTHAMP.KS0010001



Problems ...

- People enumerated at residential location only
 - Some limited workplace and education data
 - 2011 census improved workplace statistics
- UK censuses only enumerated every 10 years (+ annual estimates at lower spatial resolution)
- 'Fixed' spatial units modifiable areal unit problem (MAUP)
- 'Fixed' temporal units modifiable spatio-temporal unit problem?

Inadequate spatio-temporal resolution



Southampton

Required spatio-temporal resolution



Southampton

Population24/7 research projects

- ESRC Population24/7 project, 2009-2011
 - Spatiotemporal population estimates by time of day, week, term, season, etc.
 - General approach/framework
 - Software: SurfaceBuilder247 (.NET)
 - Sample outputs
- Various PhD projects population exposure to hazards
- ESRC Population24/7 Near Real Time project (current)
 - Extending from conventional data to new forms of data

Developing a Flexible Framework for Spatiotemporal Population Modeling

David Martin,* Samantha Cockings,* and Samuel Leung[†]

*Geography and Environment, University of Southampton [†]School of Civil Engineering and Surveying, University of Portsmouth

This article proposes a general framework for modeling population distributions in space and time. This is particularly pertinent to a growing range of applications that require spatiotemporal specificity; for example, to inform planning of emergency response to hazards. Following a review of attempts to construct time-specific representations of population, we identify the importance of assembling an underlying data model at the highest resolution in each of the spatial, temporal, and attribute domains. This model can then be interrogated at any required intersection of these domains. We argue that such an approach is necessary to moderate the effects of what we term the *modifiable spatiotemporal unit problem* in which even detailed spatial data might be inadequate to support time-sensitive analyses. We present an initial implementation of the framework for a case study of Southampton, United Kingdom, using bespoke software (SurfaceBuilder247). We demonstrate the generation of spatial population distributions for multiple reference times using currently available data sources. The article concludes by setting out key research areas including the enhancement and validation of spatiotemporal population methods and models. *Key Words: GIS, modifiable spatiotemporal unit problem, population, spatiotemporal.*

Martin, D., Cockings, S., & Leung, S. (2015). *Developing a flexible framework for spatiotemporal population modeling*. *Annals of the Association of American Geographers*, *105*(4), 754-772 doi: 10.1080/00045608.2015.1022089



Introduction to population24/7 approach and surfacebuilder247



Southampton Residential locations (houses, postcodes, census output areas)



Point locations with residential populations, split into sub-groups

These populations sum to total population in model



could be mapped onto grid





which would generate this

r			



Southampton Also contains non-residential locations (buildings, postcodes, workplace zones)



Point locations with non-residential populations, split into sub-groups

Associated spatial extents, catchment areas ("wide area dispersion") and time profiles



could also be mapped onto grid





Which would generate this





Add a transportation network



to generate a background layer

	Ν	N	N	
	N	N		

Traffic densities could be expressed as persons per cell – expected relative density of people in transit – layer is time-specific

Nodata cells represent locations which cannot contain population (sea)

A universe of potential population locations



All of which change continuously

?	?					?
?					?	?
?			?			
	?	?				
?	?	?		?		
?						
	?				?	?







Night-time model: residential only



Day time: mixed activities, but which ones?





Lazy Sunday morning..?



Wednesday morning, school holidays...?





Population24/7 in a slide...

- Assemble multi-source data library of locations, populations and time profiles
- Transfer total population from origins locations to cells in output grid
- Principal loop is of destinations, identifying population required at specific target time
- Some population remains at origins; some transfers to destinations; some allocated into background (in transit), local dispersion applied
- All processing repeated per population subgroup



SurfaceBuilder247

Southampton

<u>₽</u> 24/7	Ballinia: Marriel Press Party
	Data selection
	Select bkgmd
	0,0,3500,350
FAT Y _ DU	Analysis area
	425000,7500
	Rea
	Select time se SICCluster Time Series Time Series Time Series Time Series
	Select dest file
	Dests_EN Dests_Sol NFDests_ VFHighfie NFHighfie NFSolentE
	Select origin fil CEN01TE V NFSolentE Origins_EV Origins_Sc Origins_Sc
	Validation
	V
	Show validati
	Read data Data session
	TestDAve
	Read data
	Show data log
	Dests read
🤼 🔲 🔁 🔁 🖏 🕕 Pop247 👘 📜 24-7 👘 Microsoft Power	oj

ata selection		- Run parameters
Locate directory C:\Users\Dave	\Documents\F	Run name
elect bkomd file pop247bkolave	rtxt 👻	Default
		Choose dest p
,0,3500,3500,200		18_74
nalysis area	Buffer	Origin pop ban
25000,75000,350,350,200	10000	POP06_1874
Read and display backgrou	und	
elect time series file		Model Date 0
SICClusters Time Profiles xls		Model time
TimeSeries1 xls		Bun
Time Series 2 xls		
TimeSeries4.xls		Show run
1 . 1 . M		Save resu
elect dest file		Due detette
Desis_EN_Secondary2003x0.cs	sv 🌔	Run details
NFDests_Solent_Workplace200	1.csv	ElapsedTime:
NFHighfieldTestDest.csv	-	
NFHighfieldTestDestJustUni.csv	=	
NFSolentExtraTestDest.csv	Ŧ	Chookaum
		Checksum
CENUTTERM_Orgins_EW_Popt	MMXU.CSV	
NFSolentExtraTestOrigin.csv	rev E	
Origins EW CEN01TERM MOV	/.csv	Display results
Origins_Solent_PopIMM.csv		
Origins_Solent_PopIMMx0.csv	-	
/alidation		
Validate single file		Show
Show validation log		Show de
Read data		Sho
ata session name		Sho
estDAve 🔻 Save	data session	Show orig

TestDAveDataLog.txt

-

Nun name												
Default												
Choose dest p	oop band											
18_74 🗸												
Drigin pop bar	nd used											
POP06_1874ST_PC -												
Model Date	03/02/2011											
Model time	12:00	2:00										
Run												
Show run log												
Save resu	ults file											
Run details												
ElapsedTime	: 37.3464											
Checksum												
	_											
splay results												
	Show back	ground										
	Show de	sts										
Show	w dest In Trav	vel at 12:00)									
Show de	ests On Site s	pread at 1	2:00									
Sho	ow NightTime	OriginPop										
Sho	ow origins Po	p at 12:00										
Show or	igins OnSite :	spread at 1	2:00									

_ **D** _ X



Acronyms: QLFS Quarterly Labour Force; DCSF Department for Children, Schools and Families; HESA Higher Education Statistics Agency; Survey; DCMS Department for Culture, Media and Sport; ALVA Association for Leading Visitor Attractions; DfT Department for Transport; TfL Transport for London; CAA Civil Aviation Authority



Bristol outputs, 2011 open data





Integrating near realtime population sources and estimates

Background



 Spatiotemporal population estimates – by time of day, week, term, season etc

Southampto

- General approach/framework
- Software: SurfaceBuilder247 (.NET)
- Example output layers
 - e.g. 8am weekday term-time, for all of Eng & Wales, 200m resolution, 2001





Acronyms: QLFS Quarterly Labour Force; DCSF Department for Children, Schools and Families; HESA Higher Education Statistics Agency; Survey; DCMS Department for Culture, Media and Sport; ALVA Association for Leading Visitor Attractions; DfT Department for Transport; TfL Transport for London; CAA Civil Aviation Authority Population247 Near Real Time Southampton (Pop247NRT) Project

- ESRC New and Emerging Forms of Data Policy Demonstrator Project
- 12 months (Feb 17-Feb 18)
- Aims
 - Enhance existing methods and tools for harvesting, processing, integrating and calibrating new, emerging and existing data sources
 - Deliver two case studies with project partners to show how methods can be used to inform policy and practice
 - Promote knowledge transfer between and beyond partners

New and emerging forms of data



OS POI data

The classification scheme has three levels. There are nine Groups and 52 Categories as shown here. These are broken down over 600+ Classes. Customers can select Points of Interest from the complete classification or from any combination of Group and/or Category.

01 Accommodation, eating and drinking

- 01 Accommodation
- 02 Eating and drinking

02 Commercial services

- 03 Construction services
- 04 Consultancies
- 07 Contract services
- 05 Employment and career agencies
- 06 Engineering services
- 60 Hire services
- 08 IT, advertising, marketing and media services
- 09 Legal and financial
- 10 Personal, consumer and other services
- 11 Property and development services
- 12 Recycling services
- 13 Repair and servicing
- 14 Research and design
- 15 Transport, storage and delivery

03 Attractions

- 58 Bodies of water
- 16 Botanical and zoological
- 17 Historical and cultural
- 19 Landscape features
- 18 Recreational
- 20 Tourism

04 Sport and entertainment

- 22 Gambling
- 23 Outdoor pursuits
- 21 Sport and entertainment support services
- 24 Sports complex
- 25 Venues, stage and screen

05 Education and health

- 26 Animal welfare
- 27 Education support services
- 28 Health practitioners and establishments
- 29 Health support services
- 31 Primary, secondary and tertiary education
- 32 Recreational and vocational education

06 Public infrastructure

- 33 Central and local Government
- 34 Infrastructure and facilities
- 35 Organisations

07 Manufacturing and production

- 37 Consumer products
- 38 Extractive industries
- 39 Farming
- 40 Foodstuffs
- 41 Industrial features
- 42 Industrial products

09 Retail

- 46 Clothing and accessories
- 47 Food, drink and multi item retail
- 48 Household, office, leisure and garden
- 49 Motoring

10 Transport

- 53 Air
- 59 Bus transport
- 57 Public transport, stations and infrastructure
- 54 Road and rail
- 55 Walking
- 56 Water



ORR Train Station Usage data Southampton

X 🚽	17 - (21 -	Ŧ													estimates-o	of-station-u	usage-2015-16.xlsx -	Microsoft Excel		
File	Home	Inse	rt Pa	age Layou	ut Forr	mulas [Data	Review	Vie	ew	Developer Learning	Guide	Acrob	at Tea	m					
Ê	K Cut B Copy ▼		Calibri		* 11	· A A	= =	= _	\$%.≁	▶ ¶ ~	🖥 Wrap Text	Gene	ral	٣			Normal 4 2	Normal 4 3	Normal 4 4	No
Paste	Format Pa	inter	BI	<u>u</u> .	<u> </u>	🗞 - <u>A</u> -	E 3		< >	5	📑 Merge & Center 🔻	9 -	%,	00. 0.≁ 0.∢ 00.	Conditional Formatting *	Format as Table ▼	Normal_l;Stat	Normal_l;Stat	Normale_Fog	No
C	lipboard	- Fai		F	ont	Fai			Ali	gnmen	t G		Number	E.				Sty	les	
	A2088	•	. (=	f_x	5922															

- 24	А	В	с	Q	R	S	т	U	V	w	х	Y	Z
	NLC	TLC	Station Name	Entries_Full	Entries_Reduced	Entries_Season	Entries_Total	Exits_Full	Exits_Reduced	Exits_Season	Exits_Total	1516 Entries & Exits	Estimated absolute change in Usage due to 2015/16 London Travelcard
1	-	–	1	~	~	-	-	-	·	-	-	*	Methodology 🚽
2080	1453	SOK	South Kenton	68,553	122,130	105,145	295,828	68,553	122,130	105,145	295,828	591,656	166,259
2081	5292	SMO	South Merton	20,121	18,245	31,542	69,908	20,121	18,245	31,542	69,908	139,816	-56,511
2082	8142	SOM	South Milford	24,559	18,787	25,883	69,229	24,559	18,787	25,883	69,229	138,458	-
2083	3057	SRU	South Ruislip	45,963	37,930	33,459	117,352	45,963	37,930	33,459	117,352	234,704	37,858
2084	7404	STO	South Tottenham	112,416	197,700	379,768	689,884	112,416	197,700	379,768	689,884	1,379,768	326,352
2085	1949	sws	South Wigston	13,599	9,475	14,355	37,429	13,599	9,475	14,355	37,429	74,858	-1
2086	6892	SOF	South Woodham Ferrers	49,021	43,850	192,715	285,586	49,021	43,850	192,715	285,586	571,172	-838
2087	3187	STL	Southall	274,015	479,966	641,600	1,395,581	274,015	479,966	641,600	1,395,581	2,791,162	451,470
2088	5922	SOA	Southampton Airport (Parkway)	288,364	335,622	285,730	909,716	288,364	335,622	285,730	909,716	1,819,432	-165
2089	5932	SOU	Southampton Central	825,380	1,655,155	699,311	3,179,846	825,380	1,655,155	699,311	3,179,846	6,359,692	-1,194
2090	5263	SOB	Southbourne	21,366	43,614	47,970	112,950	21,366	43,614	47,970	112,950	225,900	2
2091	6947	SBU	Southbury	101,116	185,126	130,752	416,994	101,116	185,126	130,752	416,994	833,988	20,779
2092	5339	SEE	Southease	2,633	5,842	508	8,983	2,633	5,842	508	8,983	17,966	9
2093	4787	SIA	Southend Airport	73,833	108,288	30,459	212,580	73,833	108,288	30,459	212,580	425,160	7,052
2094	7456	SOC	Southend Central	347,863	506,775	691,515	1,546,153	347,863	506,775	691,515	1,546,153	3,092,306	-177,586
2095	7457	SOE	Southend East	198,090	288,582	393,782	880,454	198,090	288,582	393,782	880,454	1,760,908	-101,126
2096	7420	SOV	Southend Victoria	161,931	235,906	321,903	719,740	161,931	235,906	321,903	719,740	1,439,480	-82,667
2097	6890	SMN	Southminster	13,816	18,538	43,617	75,971	13,816	18,538	43,617	75,971	151,942	-1
2098	2262	SOP	Southport	284,234	1,422,538	366,760	2,073,532	284,234	1,422,538	366,760	2,073,532	4,147,064	-
2099	5286	SWK	Southwick	42,232	92,567	46,237	181,036	42,232	92,567	46,237	181,036	362,072	-1
2100	8527	SOW	Sowerby Bridge	77,723	63,130	55,030	195,883	77,723	63,130	55,030	195,883	391,766	-
2101	6388	SPA	Spalding	25,332	54,415	12,951	92,698	25,332	54,415	12,951	92,698	185,396	-1
2102	8836	SBR	Spean Bridge	2,409	1,257	-	3,666	2,409	1,257	-	3,666	7,332	-
2103	2199	SPI	Spital	95.684	82.370	56.821	234.875	95.684	82.370	56.821	234.875	469.750	-

New and emerging forms of data





An ESRC Data Investment

Centre				Topics	Products	Geography	 All Datasets 	Maps	Tutorials	CDRC	
/ Organizations / Consum	er Data Res	search Cent	tre / Loca	al Data Co	ompany -	UCL					
Local Data Company - UCL SmartStreetSensor Footfall	Dataset	🚰 Groups	② Activity	y Stream							
Data Secure	The footfa Company	ll sensor data i (LDC) in partne	is based on ership with	passive Wi- UCL. This da	Fi signal pro ata contains:	bing from a se	ensor network de	ployed by	Local Data		
Followers 4	 Sensor locations Anonymized hashed MAC addresses of detected devices Signal strength and number of packets 										
	1000 sensors to be located at approximately 81 towns and cities across the UK.										
	This contr	olled dataset is	s only availa	ble for acade	emic research	n on a contrac	t basis and the a	pplication	forms and n	nore	
License	details can	i be found <u>here</u>	<u>e</u> .								
Other (Not Open)	Data and Resources										
	Terms of Use - LDC SmartStreetSensor Controlled										
🖻 Social											
Soogle+								 More infor 	rmation \bigcirc Do	ownload	
Twitter	Secure	footfall	location	mobile	network	sensor w	ifi				
Facebook	Additional Info										
	Field					Value					
	Product					LDC Footfall S	Sensor Data				
	Date Rar	ige				2015-present					
	Collectin	g Date		Continuously	Jously						

SmartStreetSensor footfall data Southampton



Traffic England data

Southampton



New and emerging forms of data



Google Places – Opening Hours data Southampton



Google Places – Popular Times data Southampton



New and emerging forms of data



New signals/proxies for ...

- **Retail** retail centres, supermarkets, high streets, markets
- Leisure restaurants?, cinemas, theatres, tourist sites, music venues, sports stadia
- **Transport** railway stations, bus stations?, airports?, cruise terminals?
- **One-off/unusual events** festivals (Glastonbury), sports (Wimbledon, Boat Race, Marathons), Bristol Balloon Festival, shows (RHS), concerts ... Bank Holidays, snow?!

Acknowledgements

- Economic and Social Research Council Awards: ES/G031304/1, ES/P010768/1
- Samantha Cockings, Samuel Leung, Glen Hart, William Holmes, Tom Charnock, Nick Gibbins (current and past coinvestigators)
- Alan Smith, Becky Martin (past PhD researchers)