

Best fit of Lower Super Output Areas to Built Up Areas (2011) ¹

The Office for National Statistics has previously published a definition of settlements for use with the 2011 Census of Population data. This defined Built Up Areas (BUA) at the Output Area (OA) level.

This statistical article accompanies a product that has been developed (a spreadsheet) that has the best fit Built Up Area for all 1,909 Lower Super Output Areas (LSOA) in Wales. This can be used to link with any data source available at LSOA level for further analysis.

This article provides a technical explanation of the methodology used to derive the best fit. While such a best fit is a useful analytical tool for various rural-urban analyses, the work in this article shows that it is not a definitive product. An important part of this article is to encourage users to explore how the classifications might be used and to comment on strengths, weakness and possible extensions.

An example of analysis using the best fit with Census of Population data has also been published as a separate Statistical Bulletin "Industrial sector of workers by size of Built Up Area, 2011".

Background

There are very few data sources available at Output Area (OA) level; the Census being the main source. To extend the usefulness of the ONS Built Up Area classification it would be helpful to have a best fit between the Built Up Areas and the Lower Super Output Area (LSOA) level. This would allow users to link with a range of data including not only the Census but also the range of other data and indicators (e.g. deprivation indices) available at LSOA level..

The best fit is presented for Wales only, although it covers those Built Up Areas that cross the Wales-England border. While this gives a useful tool there are many plausible ways to calculate the best fit and this cannot be claimed to be definitive.

This paper shows how the best fit was calculated and examines the strengths and weakness of the approach. Spreadsheets accompanying this paper show the calculations and give look ups between LSOA and Built Up Area.

Date of Publication: 20 May 2015

Author: Stuart Neil, Knowledge and Analytical Services

E-mail: stats.agric@wales.gsi.gov.uk

Telephone: 029 2082 6822

Twitter: www.twitter.com/statisticswales | www.twitter.com/ystadegaucymru

¹ Notes on the use of statistical articles can be found at the end of this document.

The paper also shows how this classification can be extended by linking with the National Statistics classification of rural and urban areas. In this way it is possible to produce a classification that uses both settlement size and a measure of “accessibility” and “remoteness”.

In connection with this paper an example of the sort of analysis the classification can be used for is has also been published. The second paper examines the distribution of workers by industrial sector taken from the Census of Population 2011.

Definition of Built Up Areas(BUAs)

Full documentation of the Built Up Areas has been published by ONS (see link in references). This section gives a quick introduction to the classification.

The areas defined are based on the physical characteristics of the man-made environment. They define areas of linked development. They do not reflect how people might use or think about a place.

The Built Up Areas for 2011 were defined in a project run by ONS, Defra, DCLG and Welsh Government and cover Wales and England. The analysis was performed by Ordnance Survey using the digital mapping data. Scotland have their own definition of settlements. For full documentation see link in references.

The definition uses mapping information to identify areas that have a man-made landscape (i.e. landscapes that contain buildings or man-made structures). From these the Built Up Areas are defined.

- Wales and England are divided into a grid of 50 metre squares (or cells). Each of these squares was assessed for the proportion of built up land it contained. This includes houses and their gardens, commercial or industrial property, roads, railways and similar structures. If any category exceeded a threshold the cell was classified as built up.
- The built up cells are joined to form contiguous polygons. Any polygon that covers at least 20 hectares (a hectare is a square 100 metres by 100 metres for a total of 10,000 square metres) is named and has a population calculated. In some cases the population of a polygon is zero. Examples might be an airfield or an industrial or commercial site separate enough to be its own polygon.
- The polygons are then linked to form the Built Up Areas. If two polygons are within 200 metres of each other at any point then they are joined into a single Built Up Area. A third polygon having a point within 200 metres of either of the first two polygons is also included. This linking of polygons continues for as long as a new polygon within 200 metres can be found.

The linking of the polygons into Built Up Areas ensures that this classification follows standard EU rules for settlement definition. These Built Up Areas are named. There are X Built Up Areas in Wales (or whose area extends into Wales).

The definition of Built Up Areas mean that they are not constrained to follow administrative boundaries. They can cross local authority boundaries and also the Wales-England border. A further confusion is that several local authorities in Wales have the same names as a Built Up Area such as Cardiff, Swansea, Newport, Bridgend and Wrexham.

Wider and Contiguous Built Up Areas

Results published by ONS show two different types of area.

- Built Up Areas. That may have a single or multiple polygons within them.
- Sub Divisions. The individual polygons within a Built Up Area.

Every Built Up Area is thus either a complete Built Up Area or a set of sub divisions that have been linked to form the overall Built Up Area. In Wales most Built Up Areas are a single unit without sub divisions. However, many of the larger Built Up Areas have multiple sub divisions.

To use the areas at the level of individual polygons we need a term to describe the set of all separate polygons – whether single polygon Built Up Area or a Sub Division of a composite Built Up Area. For the purposes of this work these will be known as “Contiguous Built Up Areas”. The combined Built Up Areas will be referred to as “Wider Built Up Areas”.

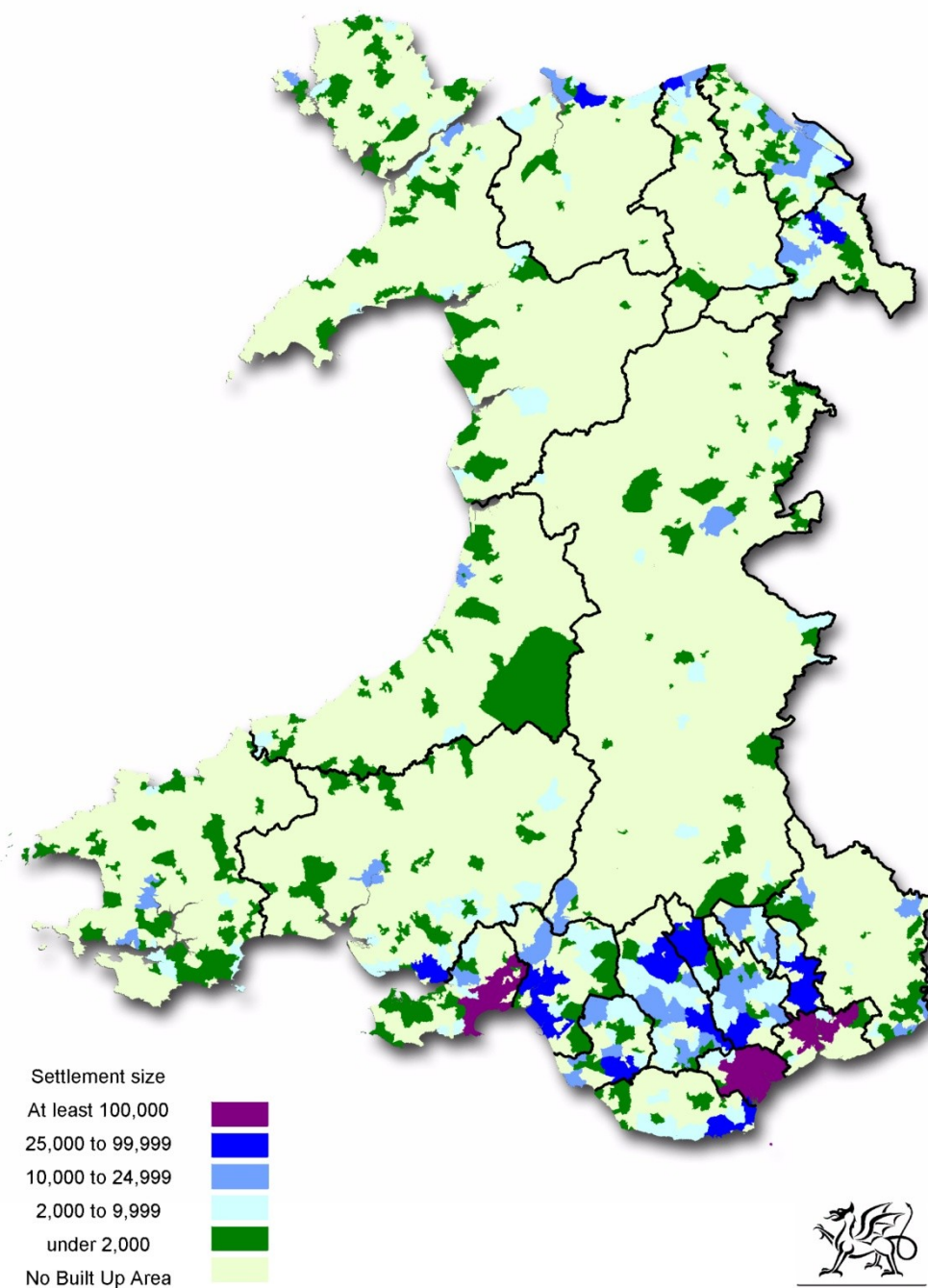
The reference code for a Contiguous Built Up Area will be the same as the Wider Built Up Area if it is a stand alone polygon. Otherwise it will be the code of the Sub Division.

The rule for linking polygons together to form the Built Up Areas can lead to some unexpected results. This can be shown by considering the Cardiff Built Up Area. This includes not just the city itself. Because of the proximity rules it also includes areas close by such as Penarth and Dinas Powys. More surprisingly it also covers a much wider area extending to include Pontypridd and Caerphilly.

Maps 1 and 2 show the OA level classification for the Contiguous and Wider Built Up Areas. Map 3 summarises the Built Up Areas showing where a Built Up Area has sub divisions and where it does not. The Built Up Areas with sub divisions tend to be the larger ones, but not uniformly so.

Map 1

Contiguous Built Up Area in Wales at OA level by size band



Source: Knowledge and Analytical Services

© Crown Copyright and database right 2015.
Welsh Government. Licence number 100021874.

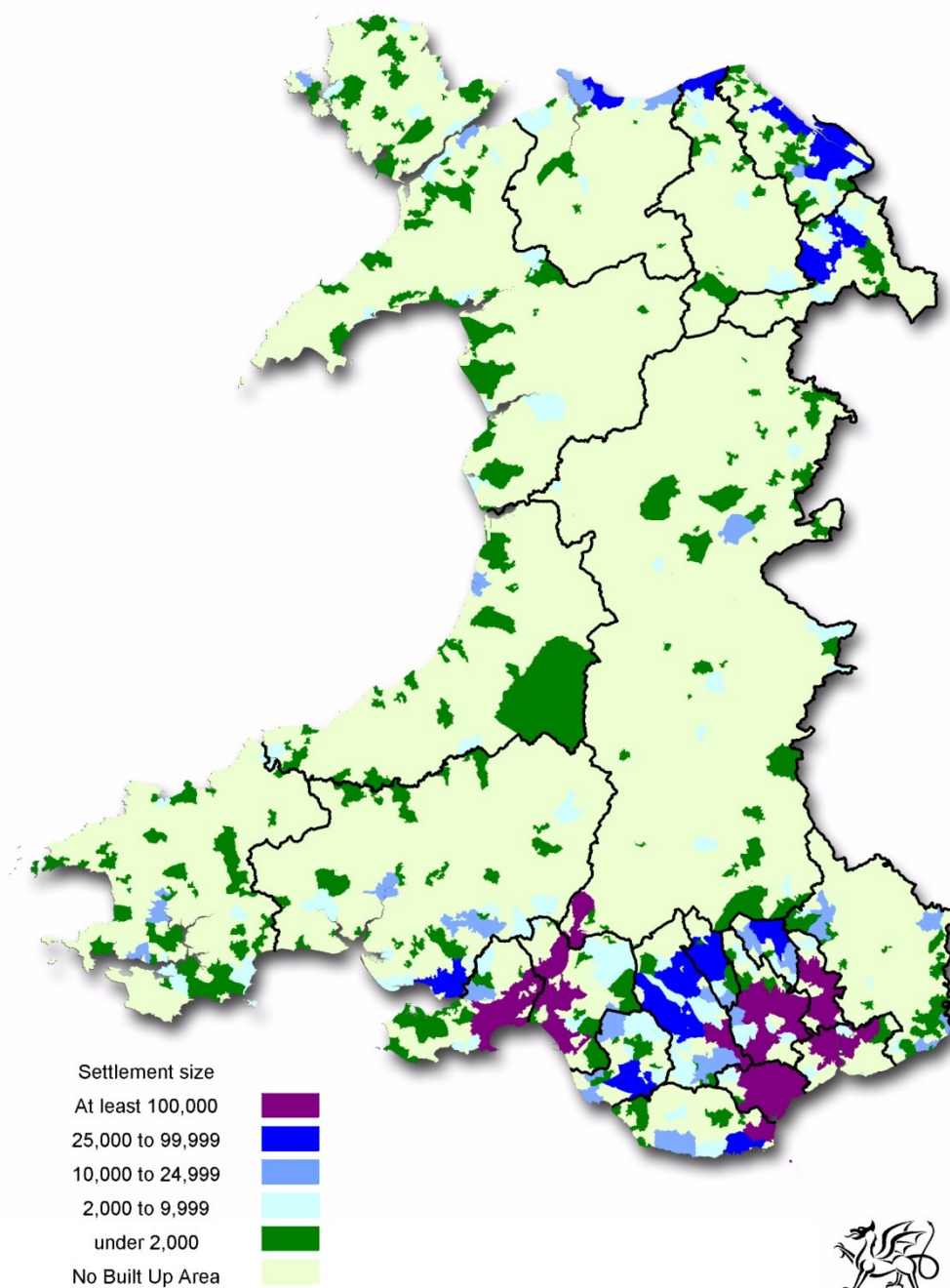


Llywodraeth Cymru
Welsh Government

029.15-16
Cartographics, FCS

Map 2

Wider Built Up Area in Wales by size band



Source: Knowledge and Analytical Services

© Crown Copyright and database right 2015.
Welsh Government. Licence number 100021874.

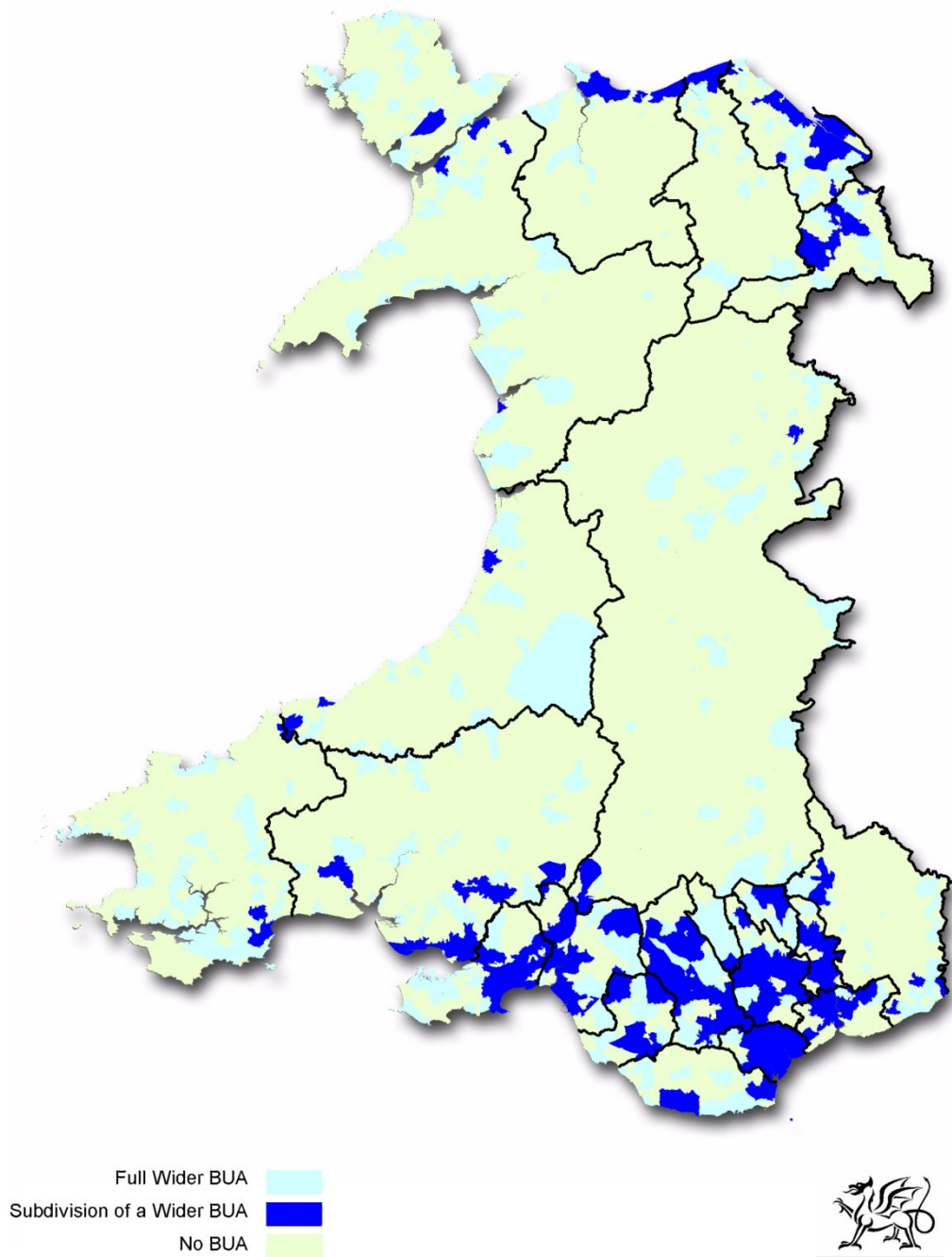


Llywodraeth Cymru
Welsh Government

029.15-16
Cartographics, FCS

Map 3

Type of Built Up Area at OA level in Wales



Source: Knowledge and Analytical Services

© Crown Copyright and database right 2015.
Welsh Government. Licence number 100021874.



Llywodraeth Cymru
Welsh Government

029.15-16
Cartographics, FCS

How the Wider Areas are formed from the Contiguous ones is explored for the Ammanford Wider Built Up Area in Annex 1. Ammanford is picked because it is a rather unusual example and so demonstrates the issues.

The classification is based on the characteristics of the man-made environment. It does not take into account how people use the settlements or how they think about them. While it might be good to have a classification that would take account of these factors, clearly it vastly increases the difficulty of the process and removes the objectivity.

For general analysis the default is to use the Wider Built Up Areas. This is driven partly by wanting consistency with previous classifications and also the link with standard definitions elsewhere across the EU. However, in a Welsh context, the Contiguous Built Up Areas seem to give a better match to what would be seen as “separate places” in Wales.

For this article both sets of Built Up Areas are considered so that users can make their own judgement. As is often the case there is going to be a trade off between consistency and local relevance.

Areas not within a Built Up Area

Not every Output Area is classified as being part of a named Built Up Area. A named Built Up Area must cover at least 20 hectares and this will exclude many of the smallest villages and hamlets. Output Areas that have no named Built Up Area are called “Dispersed” in this paper. Note that in the process of linking the Contiguous Built Up Areas together to form the Wider Built Up Areas, the Dispersed OA are unaffected.

For use in the best fit calculations we can think of all the Dispersed Output Areas forming a single unit equivalent to a Built Up Area. An LSOA that has one or more Dispersed OA in effect has one more Built Up Area. When looking at the size groups for Built Up Areas the Dispersed OA will be counted as being in the smallest size group.

Distribution of population by size of Built Up Area

This section summarises the results for Wales from the published Output Area level. Table 1 shows the share of the Welsh population by size band for Wider and Contiguous Built Up Areas. The results are mapped in Maps 2 and 3.

Note that there are a number of Built Up Areas that cross the border between Wales and England. Examples include Chester, Hay on Wye and Chepstow. The size of these Built Up Areas is based on the whole settlement, not just the part in Wales. Thus we treat Chester as a settlement of some 90,000 people, with some 5,000 of them living in Wales.

Table 1**Welsh population by size of Built Up Area, 2011**

	People ('000)		Share	
	Contig	Wider	Contig	Wider
Under 2,000	622	579	20%	19%
2,000 to 9,999	617	427	20%	14%
10,000 to 24,999	580	403	19%	13%
25,000 to 99,999	602	600	20%	20%
At least 100,000	643	1,054	21%	34%
Total	3,063	3,063	100%	100%

Source: ONS classification of Built Up Areas at Output Area level with usual resident population from Census of Population, 2011

The smallest category in the table includes both small Built Up Areas (with a population of under 2,000 people) and all the Dispersed OA that have no named Built Up Area.

The size bands are selected to be relevant in Welsh circumstances. For the contiguous Built Up Areas this particular choice of size bands gives similar populations in each band. A similar table for England would need extra categories for larger Built Up Areas. For England as a whole around 80 per cent of the people live in settlements of at least 10,000 people. In Wales the share is around 60 per cent.

Best fit methods

The main principle is that we use the published Census OA level population estimates to make the best fit. To ensure that users could replicate the working it uses this source directly without extra calculations using other GIS data.

In this paper we use two methods

- Simple
- Threshold

Simple best fit methodology

For simple best fit we find all the Built Up Areas in an LSOA. Then we find the Built Up Area that has the largest share of the LSOA population by adding up the OA population. There is a risk that this could leave an LSOA with more than one largest Built Up Area. To happen this would need two sets of OA with exactly the same overall population from two different Built Up Areas. However, this does not happen in Wales using the published data. The spreadsheets that accompany this paper include a list of all Built Up Areas identified in each LSOA in Wales, together with the relevant populations.

Threshold best fit methodology

The threshold best fit is a similar method but does not use full set of published Built Up Areas. Instead it treats all the Built Up Areas with a population below the chosen threshold as belonging to a "Small" class. The dispersed OA are also included in this category. The Small category is, like Dispersed in the simple best fit, treated as if it was a single Built Up Area. With this modification the rest of the calculation follows the simple best fit method.

The idea behind the threshold is to try and simplify the treatment of the smaller and fringe settlements. This becomes clearer as the numbers are developed below.

In this paper a threshold of 2,000 people is used to define the Small Built Up Areas. This is an arbitrary threshold but the reasoning behind the figure is demonstrated when we assess the goodness of fit.

Mix of Built Up Areas within an LSOA

A Welsh LSOA usually has 4 or 5 OA although there is quite wide variation (between 3 and 7). Each of these OA can be in a Built Up Area or be Dispersed. There can be multiple Built Up Areas present in an LSOA combined with multiple Dispersed OA.

The first step is to give a summary of how many LSOA have a single Built Up Area or multiple Built Up Areas in them. In this analysis all the Dispersed OA in an LSOA count as the equivalent of a single Built Up Area. Table 2 shows the distribution for both Contiguous and Wider Built Up Areas and for the Simple and Threshold best fit.

Table 2

Distribution of LSOA by number of individual BUA they contain

	Number of different Built Up Areas in an LSOA				Total
	One	Two	Three	Four	
<u>Number of LSOA</u>					
<u>Contiguous BUA</u>					
Simple Best Fit	1,481	336	86	6	1,909
Threshold Best Fit	1,689	208	12	0	1,909
<u>Wider BUA</u>					
Simple Best Fit	1,517	321	68	3	1,909
Threshold Best Fit	1,713	191	5	0	1,909
<u>Share of LSOA</u>					
<u>Contiguous BUA</u>					
Simple Best Fit	77.6%	17.6%	4.5%	0.3%	100.0%
Threshold Best Fit	88.5%	10.9%	0.6%	0.0%	100.0%
<u>Wider BUA</u>					
Simple Best Fit	79.5%	16.8%	3.6%	0.2%	100.0%
Threshold Best Fit	89.7%	10.0%	0.3%	0.0%	100.0%

Source: ONS classification of Built Up Areas

In all cases less than a quarter of the LSOA contain multiple Built Up Areas where the best fit is not an exact match. This can be improved if we group all the smaller Built Up Areas together if they fall below a threshold of 2,000 people. With this threshold in place just over 1 in 10 LSOA contain more than one Built Up Area. The share of LSOA with multiple Built Up Areas is also reduced when we use Wider rather than Contiguous Built Up Areas.

Table 3 extends Table 2 to demonstrate why the threshold is useful in improving the fit. The Built Up Areas are split between "small" Built Up Areas with a population of under 2,000 people (including Dispersed OA) and "other". The table also shows the number of

“matched” people. That is people in OA that have the same Built Up Area as the LSOA best fit.

Each LSOA can be categorised by the mix of Built Up Areas it has in it. For the small Built Up Areas under the 2,000 person threshold there might be none, one or multiple areas. This is also true for the other Built Up Areas above the threshold. It is rare for an LSOA to have more than one “other” Built Up Area, but it can happen.

Table 3

Summary by number of individual Built Up Areas in an LSOA

Individual BUA		LSOA	All people (‘000)	Matched people (‘000)		Share matched	
Small	Other			Simple	Threshold	Simple	Threshold
Contiguous BUA							
One	None	110	169	169	169	100.0%	100.0%
None	One	1,371	2,177	2,177	2,177	100.0%	100.0%
Multiple	None	208	349	219	349	62.6%	100.0%
None	Multiple	25	44	30	30	68.3%	68.3%
One	One	154	250	193	193	77.0%	77.0%
One	Multiple	11	20	11	11	55.8%	55.8%
Multiple	One	29	52	30	35	56.3%	67.6%
Multiple	Multiple	1	2	1	2	51.2%	74.5%
Total		1,909	3,063	2,829	2,966	92.3%	96.8%
Wider BUA							
One	None	103	158	158	158	100.0%	100.0%
None	One	1,414	2,249	2,249	2,249	100.0%	100.0%
Multiple	None	196	330	208	330	63.3%	100.0%
None	Multiple	10	19	13	13	68.2%	68.2%
One	One	159	259	201	201	77.5%	77.5%
One	Multiple	5	9	5	5	55.2%	55.2%
Multiple	One	22	40	23	27	58.9%	66.9%
Multiple	Multiple	0	0	0	0	na	na
Total		1,909	3,063	2,858	2,982	93.3%	97.3%

Source: ONS classification of Built Up Areas

If we use a threshold then the LSOA with multiple “small” Built Up Areas but no “other” will become fully matched. This is just over 200 LSOA for Contiguous Built Up Areas and just under for Wider. Notice also that there are very few LSOA that have multiple Built Up Areas with at least 2,000 people.

An LSOA that has a simple best fit of either Dispersed or a BUA with a population of under 2,000 people will automatically have a threshold best fit of Small BUA. An LSOA can change from a simple best fit of a named Built Up Area of at least 2,000 to Small if there is a combination of Dispersed OA and small Built Up Areas that together exceed the people in the larger Built Up Area. There are only two such LSOA in Wales.

Table 4 looks at the share of people in an LSOA who are in OA that have the same Built Up Area as the best fit. The table again considers both Contiguous and Wider Built Up Areas and also simple and threshold best fit.

Table 4 confirms that a large majority of LSOA have only a single Built Up Area. It is only LSOA where this is not true that will have mismatch between the actual OA and Best Fit LSOA allocation. It also shows that, at worst, less than 50 LSOA out of 1,909 do not have a majority of their population in the best fit Built Up Area. If the threshold best fit is used the number without a majority falls to less than 5.

However, we also see that LSOA in the range from at least 90 but under 100 per cent are relatively rare. Most of the LSOA that are not a complete fit fall into the range from at least 50 to under 90 percent of the LSOA population.

As another way of looking at the fit within an LSOA, Table 5 categorises the LSOA by the size band for the best fit Built Up Area. It shows how many LSOA fall into each of the categories and the populations. It also shows how many people live in OA that have the same Built Up Area as the LSOA – those that are matched. These results are shown for Contiguous and Wider Built Up Areas and for simple and threshold best fit.

Table 5 makes clear the advantage of using threshold best fit method for the smallest Built Up Areas. The share of the population that is matched increases from just over 73 per cent for simple best fit to just under 97 per cent with a threshold best fit at 2,000 people. There is a similar improvement for the Wider Built Up Areas.

Taking Tables 2 to 5 together we can see that overall the best fit for LSOA is good. A large majority of the LSOA contain only a single Built Up Area. For these LSOA the Best Fit is a perfect match.

The tables show the the problems with the best fit are generally for the smallest Built Up Areas and, particularly, the Dispersed areas. This is why using a threshold for the smallest Built Up Area we will use is so helpful.

The fit is a little better using the Wider rather than the Contiguous Built Up Areas. It is also improved by using the Threshold rather than Simple best fit.

There are however, a relatively small number of LSOA that are very much mixed. The best fit for these covers only a small proportion of the overall LSOA population.

Table 4**Summary by share of the LSOA population in the Best Fit BUA**

Share in Best Fit	LSOA	People ('000)			Share matched
		Matched	Not	Total	
Contiguous BUA with Simple Best Fit					
Under 50%	46	33	46	79	41.7%
50% to 60%	103	95	77	172	55.0%
60% to 70%	83	91	50	140	64.7%
70% to 80%	95	114	38	152	75.0%
80% to 90%	79	112	20	132	84.9%
90% to 100%	22	38	3	42	91.7%
Exactly 100%	1,481	2,346	0	2,346	100.0%
Total	1,909	2,829	235	3,063	92.3%
Contiguous BUA with Threshold Best Fit					
Under 50%	4	3	3	6	46.0%
50% to 60%	42	38	31	69	55.4%
60% to 70%	36	41	23	64	64.4%
70% to 80%	57	69	22	91	75.4%
80% to 90%	61	85	16	100	84.5%
90% to 100%	20	35	3	38	91.8%
Exactly 100%	1,689	2,695	0	2,695	100.0%
Total	1,909	2,966	98	3,063	96.8%
Wider BUA with Simple Best Fit					
Under 50%	39	27	39	67	41.0%
50% to 60%	83	75	61	136	54.9%
60% to 70%	75	82	44	126	65.1%
70% to 80%	93	113	38	151	75.0%
80% to 90%	80	115	20	135	85.3%
90% to 100%	22	38	3	42	91.7%
Exactly 100%	1,517	2,408	0	2,408	100.0%
Total	1,909	2,858	206	3,063	93.3%
Wider BUA with Threshold Best Fit					
Under 50%	2	1	1	2	45.0%
50% to 60%	32	29	24	53	55.1%
60% to 70%	29	33	17	50	65.3%
70% to 80%	54	65	22	86	75.0%
80% to 90%	59	82	14	96	85.0%
90% to 100%	20	35	3	38	91.8%
Exactly 100%	1,713	2,737	0	2,737	100.0%
Total	1,909	2,982	81	3,063	97.3%

Source: ONS classification of Built Up Areas

Table 5

Best fit for LSOA by size band for Contiguous and Wider BUA

BUA size band & fit type	LSOA	People ('000)			Share matched
		All	Matched	Others	
Contiguous BUA - Simple					
Under 2,000	354	578	424	154	73.4%
2,000 to 9,999	405	649	601	49	92.5%
10,000 to 24,999	363	586	568	18	97.0%
25,000 to 99,999	385	606	596	10	98.4%
At least 100,000	402	645	640	5	99.3%
Total	1,909	3,063	2,829	235	92.3%
Contiguous BUA - Threshold					
Under 2,000	356	581	562	19	96.7%
2,000 to 9,999	404	648	600	48	92.6%
10,000 to 24,999	362	584	567	17	97.1%
25,000 to 99,999	385	606	596	10	98.4%
At least 100,000	402	645	640	5	99.3%
Total	1,909	3,063	2,966	98	96.8%
Wider BUA - Simple					
Under 2,000	325	529	393	136	74.3%
2,000 to 9,999	284	455	422	33	92.6%
10,000 to 24,999	252	412	397	14	96.5%
25,000 to 99,999	382	607	595	12	98.0%
At least 100,000	666	1,060	1,050	10	99.1%
Total	1,909	3,063	2,858	206	93.3%
Wider BUA - Threshold					
Under 2,000	327	532	519	13	97.5%
2,000 to 9,999	283	454	421	33	92.8%
10,000 to 24,999	251	410	397	13	96.7%
25,000 to 99,999	382	607	595	12	98.0%
At least 100,000	666	1,060	1,050	10	99.1%
Total	1,909	3,063	2,982	81	97.3%

Source: ONS classification of Built Up Areas

Assessing the fit for Built Up Areas

As well as assessing the best fit across the 1,909 LSOA in Wales we also need to make the assessment across the Built Up Areas. The first stage of this is to look at Built Up areas by size group.

Table 6 identifies four categories of people by adding up different sets of OA level population.

- Total – all OA in Built Up Areas in a size band.
- Matched – Adding up the OA level populations for those OA that have the same Built Up Area as the best fit for the LSOA they are in.
- False in – OA in an LSOA that have a different Built Up area to the best fit
- False out – OA that are in the Built Up Area named as the best fit but are outside the group of LSOA
- Gross mismatch – adding together the false in and false out to give the total mismatched. Table shows both a count of the people and expresses this as a share of the true population.

The total population of a size group is the matched population plus the false out. The total population in the best fit is the matched plus false in.

The table shows that the improvement in fit from the simple best fit to the threshold best fit is particularly marked for the smallest Built Up Areas of under 2,000 people. Changes for the larger Built Up Areas are much less marked.

The table also shows a small improvement in fit for the Wider Built Up Areas compared to the Contiguous.

The table finally shows that the goodness of fit improves with settlement size.

The main issue with using a threshold best fit is how to set the threshold. The value of 2,000 people was largely a pragmatic decision. With a threshold at 2,000 people there are no missing Built Up Areas. If the threshold was lowered to, say, 1,500 then a few of the relevant Built Up Areas would have no LSOA allocated in the Best Fit.

Table 7 shows a similar set of variables for the threshold best fit for individual Wider Built Up Areas in Wales that have a population of at least 25,000 people. Note that in this table Chester is omitted. This is a Built Up Area of around 80,000 people but only 5,000 of these live in Wales.

Table 6

Fit for Built Up Areas by size group and fit method

BUA size group, type and fit method	People ('000)				Gross mismatch	
	Total	Matched	False in	False out	People	Share
Contiguous BUA - Simple Best Fit						
Under 2,000	622	424	154	198	352	57%
2,000 to 9,999	617	601	49	16	65	11%
10,000 to 24,999	580	568	18	12	30	5%
25,000 to 99,999	602	596	10	6	16	3%
At least 100,000	643	640	5	2	7	1%
Total	3,063	2,829	235	235	469	15%
Contiguous BUA - Threshold Best Fit						
Under 2,000	622	562	19	60	79	13%
2,000 to 9,999	617	600	48	17	65	10%
10,000 to 24,999	580	567	17	12	29	5%
25,000 to 99,999	602	596	10	6	16	3%
At least 100,000	643	640	5	2	7	1%
Total	3,063	2,966	98	98	196	6%
Wider BUA - Simple Best Fit						
Under 2,000	579	393	136	186	322	56%
2,000 to 9,999	427	422	33	5	39	9%
10,000 to 24,999	403	397	14	6	20	5%
25,000 to 99,999	600	595	12	5	17	3%
At least 100,000	1,054	1,050	10	4	14	1%
Total	3,063	2,858	206	206	411	13%
Wider BUA - Threshold Best Fit						
Under 2,000	579	519	13	60	74	13%
2,000 to 9,999	427	421	33	6	38	9%
10,000 to 24,999	403	397	13	6	20	5%
25,000 to 99,999	600	595	12	5	17	3%
At least 100,000	1,054	1,050	10	4	14	1%
Total	3,063	2,982	81	81	163	5%

Source: ONS classification of Built Up Areas

Table 7**Threshold best fit for Wider Built Up Areas of at least 25,000 people**

Name	Size	People			Gross mismatch	
		Matched	False in	False out	People	Share
Cardiff BUA	447,287	445,786	2,098	1,501	3,599	0.8%
Newport (Newport) BUA	306,844	305,399	3,718	1,445	5,163	1.7%
Swansea BUA	300,352	298,944	4,074	1,408	5,482	1.8%
Wrexham BUA	65,692	65,258	2,992	434	3,426	5.2%
Buckley BUA	63,576	62,870	1,056	706	1,762	2.8%
Tonypandy BUA	62,545	62,253	1,413	292	1,705	2.7%
Bridgend BUA	58,380	57,349	882	1,031	1,913	3.3%
Barry BUA	54,673	54,673	228	0	228	0.4%
Llanelli BUA	49,591	48,989	534	602	1,136	2.3%
Rhyl/Prestatyn BUA	46,267	45,721	0	546	546	1.2%
Merthyr Tydfil BUA	43,820	43,820	0	0	0	0.0%
Colwyn Bay BUA	34,284	33,837	993	447	1,440	4.2%
Ebbw Vale BUA	33,068	32,275	0	793	793	2.4%
Aberdare BUA	31,135	31,135	326	0	326	1.0%
Flint BUA	26,442	26,442	1,884	0	1,884	7.1%
Rhosllanerchrugog BUA	25,362	25,362	1,854	0	1,854	7.3%

Source: ONS classification of Built Up Areas

Excluding the portion of Chester that is in Wales

The name Rhosllanerchrugog is a separate Built Up Area just to the south of Wrexham Built Up Area It includes Ruabon which may be a more familiar name.

Extending the Built Up Area classification

For the smaller Built Up Areas it may also be useful to split between those that are in some way “remote” and those that are “accessible”. This is quite an intuitive way to classify the more rural areas. In Scotland the main rural-urban classification works in this way. Areas are first classified by settlement size. Areas outside settlements of at least 10,000 people are then also classified by the time it takes to drive to the nearest settlement of at least 10,000 people. If this drive time is less than 30 minutes the area is said to be “accessible”, if it is over 30 minutes the area is “remote”.

For Wales it is possible to use the existing National Statistics rural-urban classification to add a level of remoteness to the classification. The National Statistics classifications divides Wales and England into the Sparsest Context and Less Sparse Context. The Sparsest Context can be used as a proxy for “accessibility” or “remoteness” even though it is not actually calculated in that way.

This paper has shown how we can classify each LSOA by settlement size. Each LSOA can also be allocated to one of the two contexts. The way that the Sparsest Context is calculated means that it will not contain the largest settlements. In Wales the largest Built Up Areas in the Sparsest Context are Aberystwyth, Carmarthen, Newtown and Holyhead all of at least 10,000 people. Aberystwyth is the largest at around 18,000 usual residents. As a town that attracts a lot of tourists the actual population of Aberystwyth will change significantly in and out of season.

Table 8 shows summary figures combining the size bands from Table 1 with the Sparsest and Less sparse context. For simplicity a single best fit calculation is used. This is the best fit for the Contiguous BUA using a **threshold** of 2,000 people for the small settlements. As before the “matched” population refers to people in OA that have the same classification as the LSOA.

Table 8

Classification using size of Built Up Area and Context

	LSOA	People ('000)		Share of Wales	Share matched
		Total	Matched		
Under 2,000 - sparsest	169	273	272	8.9%	99%
Under 2,000 - less sparse	187	307	290	10.0%	94%
2,000 to 9,999 - sparsest	57	90	84	3.0%	93%
2,000 to 9,999 - less sparse	347	557	516	18.2%	93%
10,000 to 24,999 - sparsest	34	57	56	1.9%	98%
10,000 to 24,999 - less sparse	328	527	511	17.2%	97%
25,000 to 99,999 - less sparse	385	606	596	19.8%	98%
At least 100,000 - less sparse	402	645	640	21.1%	99%
All sparsest	260	421	412	13.7%	98%
All less sparse	1,649	2,643	2,554	86.3%	97%
Wales	1,909	3,063	2,966	100.0%	97%

Source: ONS classification of Built Up Areas and National Statistics classification of rural and urban areas

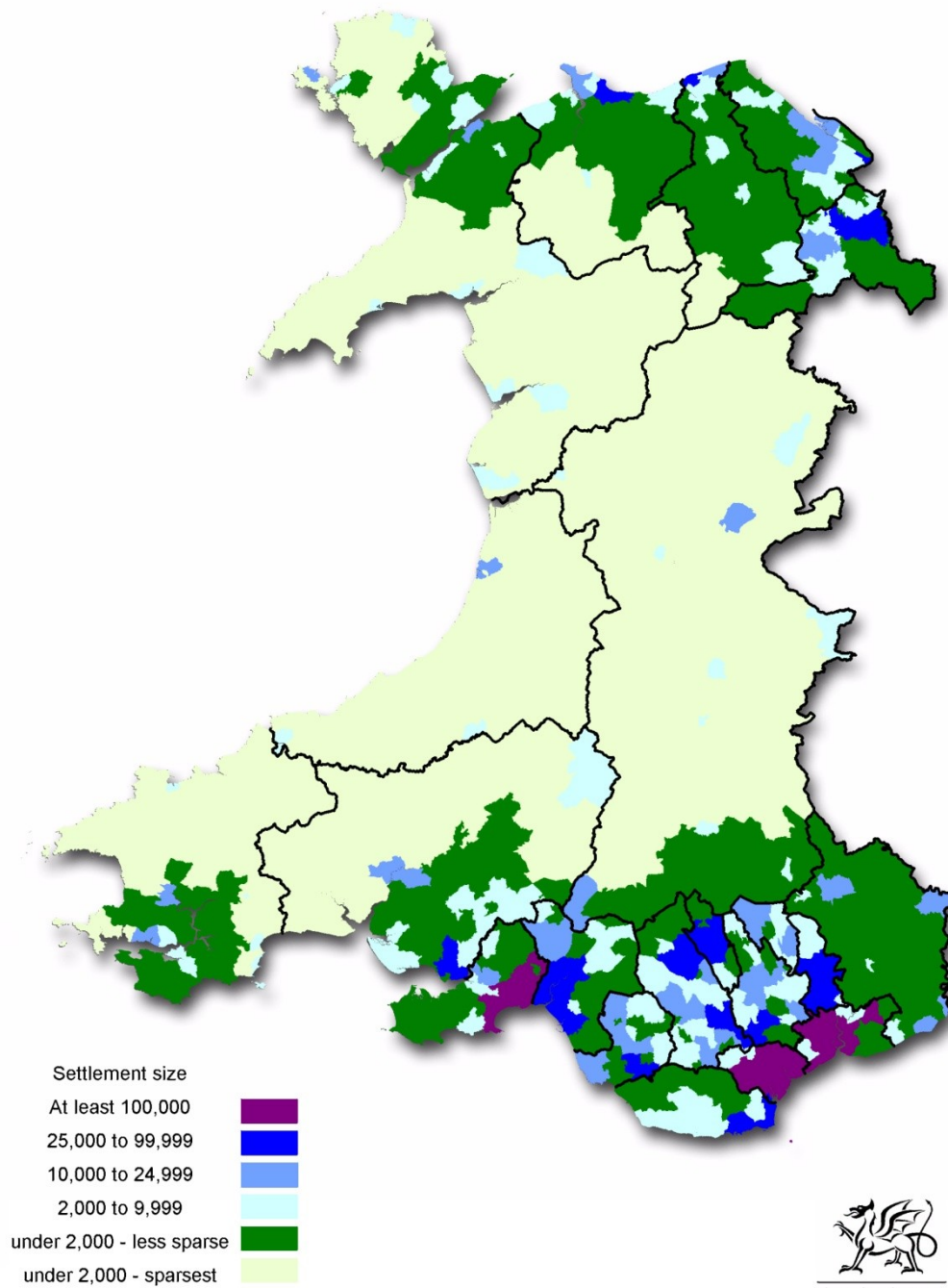
Using best fit to Contiguous BUA with a threshold of 2,000 for "small" settlements

Most of the categories identified have reasonably large populations that may be useful for further analysis. However, the categories for Built Up Areas of 2,000 to 9,999 and 10,000 to 24,999 people in the Sparsest Context may be rather small. This could be overcome by combining them – either into an at least 2,000 people in the Sparsest Context category or with Built Up Areas of similar size in the Less Sparse Context.

This classification, with the smallest settlements split by Context, is shown in Map 4. The population shares in each classification band for each local authority in Wales are shown in Table 9.

Map 4

Extended best fit to Contiguous Built Up Areas by size band



Source: Knowledge and Analytical Services

© Crown Copyright and database right 2015.
Welsh Government. Licence number 100021874.



Llywodraeth Cymru
Welsh Government

029.15-16
Cartographics, FCS

Table 9**Classification using size of Built Up Area, Context and Local Authority**

	Under 2,000		2,000 - 9,999	10,000 - 24,999	25,000 - 99,999	At least 100,000	Total ('000 people)
	Sparsest	Less sparse					
Ceredigion	67.0%	none	9.4%	23.5%	none	none	75.9
Powys	51.8%	7.8%	26.9%	13.6%	none	none	133.0
Gwynedd	38.6%	16.2%	30.8%	14.4%	none	none	121.9
Isle of Anglesey	30.2%	19.9%	33.5%	16.4%	none	none	69.8
Pembrokeshire	33.2%	16.1%	28.0%	22.7%	none	none	122.4
Carmarthenshire	20.2%	17.7%	30.2%	9.0%	23.0%	none	183.8
Monmouthshire	none	34.7%	12.6%	52.7%	none	none	91.3
Denbighshire	2.6%	21.6%	32.2%	17.4%	26.2%	none	93.7
Flintshire	none	21.4%	23.8%	51.4%	3.4%	none	152.5
Conwy	4.6%	13.6%	32.8%	23.0%	26.0%	none	115.2
Wrexham	none	13.7%	28.5%	10.3%	47.5%	none	134.8
Neath Port Talbot	none	11.5%	12.9%	11.6%	63.9%	none	139.8
Vale of Glamorgan	none	9.3%	27.3%	none	63.4%	none	126.3
Caerphilly	none	8.8%	29.2%	39.5%	22.5%	none	178.8
Swansea	none	7.8%	10.8%	6.4%	none	75.1%	239.0
Bridgend	none	7.7%	15.8%	43.9%	32.6%	none	139.2
Merthyr Tydfil	none	4.9%	20.5%	none	74.5%	none	58.8
Blaenau Gwent	none	4.7%	32.2%	63.0%	none	none	69.8
Rhondda Cynon Taf	none	4.0%	33.6%	36.0%	26.4%	none	234.4
Newport	none	2.8%	7.8%	none	none	89.4%	145.7
Torfaen	none	none	16.0%	none	84.0%	none	91.1
Cardiff	none	none	2.3%	none	0.8%	96.9%	346.1
Total	8.9%	10.0%	21.1%	19.1%	19.8%	21.1%	3,063.5

Source: ONS classification of Built Up Areas and National Statistics

classification of rural and urban areas

Using best fit to Contiguous BUA with a threshold of 2,000 for "small" settlements

Local authorities ranked by the share in settlements of under 2,000 people

Using the best fit Built Up Areas

The main reason to develop the 'best fits' was to allow users to analyse the range of datasets available at LSOA level using settlement size.

This can be useful for presentation because settlement size is an easy concept to explain (even if the detail of how this is done is not so easy).

The above analysis has shown that it is worthwhile producing an LSOA best fit to Built Up Areas.

However, it does not give a definitive solution. Users can see what has been done and decide to modify these results to fit their needs. The penalty for this is the loss of consistency with what other users are doing and this may, or may not, be important.

Using a best fit with a threshold of 2,000 people with the Contiguous Built Up Areas seems to be a good starting point for analysing Welsh data. Using a threshold makes a significant improvement to the overall fit of the classification.

The Contiguous Built Up Areas seem to fit better with Welsh expectations and settlement pattern. For example, the way that a town like Caerphilly is part of the Cardiff Wider Built Up Area. The disadvantage of using the Contiguous Built Up Areas is that it is the Wider areas that are generally used in England and for EU comparisons. If the focus of the analysis is on Wales then this disadvantage may not be so important.

Using the Built Up Area classification

The LSOA best fit provides a useful tool for statistical analysis. Using the attached spreadsheet we can link the classification with many existing data sources. This can be a useful tool for “rural proofing” and generally for understanding how settlement pattern effects various social, economic, demographic or deprivation variables.

In all cases the classification is a statistical tool. It assumes that we can accept a “margin of error” in the analyses. This is acceptable for broad statistical analyses. However, it would not be appropriate to use the classification directly for funding purposes – where the inclusion or exclusion of an area sets the eligibility of an area for funding. It might be feasible to use the classification as a starting point, with an opportunity for discussions about a small number of marginal cases, leading to a final calculation of eligibility. .

Identifying a single definition that should be used in all cases is, as we have seen, problematic. However, if the focus of the analysis is within Wales then using the best fit for Contiguous Built Up Areas of at least 2,000 people provides a sensible starting point. If the results need to be compared with England or other parts of the EU then the Wider Built Up Areas (with the same threshold) is a good starting point.

For users with a focus on particular parts of Wales it may be sensible to trade some consistency to teak the classification for local purposes. Using local knowledge in this way may be a useful way forward, but the cost is, obviously, a loss of consistency.

On line data for analysis

To accompany this article a spreadsheet has been made available on line. It shows the best fit Built Up Area for each of the 1,909 LSOA in Wales. It shows this for both Wider and Contiguous Built Up Areas for best fit with a threshold of 2,000 people.

For convenience it also includes the categories from the National Statistics rural-urban classification.

The spreadsheet that shows the full set of calculations used in this article can also be made available, but is rather large. This can be sent out on CD on request for those wishing to delve deeper.

Feedback and further information

A key reason for doing Statistical Articles is to try and prompt feedback from users. Has the Article been useful, what might you use it the look-ups for and what might have been done differently? Also interesting to have feedback on what the classification has been used for and what were the results.

Please feel free to make (constructive) comments at the address below.

Stuart Neil, Agriculture and Rural Affairs Statistician

statsagric@wales.gsi.gov.uk

029 2082 6822

References

Built Up Areas documentation

<http://www.ons.gov.uk/ons/guide-method/geography/beginner-s-guide/census/built-up-areas---built-up-area-sub-divisions/index.html>

National Statistics rural-urban classification

<http://www.ons.gov.uk/ons/guide-method/geography/products/area-classifications/2011-rural-urban/index.html>

Statistical Focus on Rural Wales (from 2008, so a bit dated, but section on “rural definitions and how to choose between them” is still relevant)

<http://gov.wales/statistics-and-research/statistical-focus-rural-wales/?lang=en>

Example analysis on the industrial sectors people work in by Built Up Area size (SB 29/2015)

<http://gov.wales/statistics-and-research/industrial-sector-workers-size-built-up-area/?lang=en>

Annex 1

Contiguous and Wider Built Up Area classifications for Ammanford

To show the build up of the Wider Built Up Areas from the Contiguous Built Up Areas this annex provides an example of Ammanford in Carmarthenshire. This is selected because it is something of an unusual example and so shows the issues involved.

The Wider Built Up Area of Ammanford has around 23,000 people. However, all the Contiguous Built Up Areas that are part of the Wider area have less than 10,000 people. The largest is the Contiguous area of Ammanford itself with a population of just under 8,000 people.

For this Annex we select all the LSOA that have at least one OA in the Ammanford Wider Built Up Area. This contains all the constituent Contiguous Built Up Areas. It also contains other Contiguous Built Up Areas that are not part of Ammanford. Finally it includes areas that have no named Built Up Area but are part of one of the selected LSOA. Table A1 shows the separate areas identified.

Table A1

BUA around Ammanford Wider BUA

Code	Name	People
Ammanford Wider BUA		
W37000391	Ammanford BUA	23,709
Contiguous BUA in the Ammanford BUA		
W38000018	Ammanford	7,945
W38000136	Pen-y-groes	5,717
W38000097	Tumble	4,302
W38000137	Tycroes	3,752
W38000054	Penybanc	1,126
W38000038	Saron	867
Other contiguous BUA in selection		
W37000182	Llannon	869
W37000217	Blaenau	425
W37000321	Llandybie	2,813
Dispersed OA in selection		
Total		4,177

Source: ONS classification of Built Up Areas

"People" - sum of usual residents in an OA as at Census 2011

Selecting all LSOA that have at least one OA in the Ammanford Wider Built Up Area

Population figures for the "other" Built Up Area are for the full Built Up Area

Map A1 shows how the various Built Up Area relate to each other. The map shows the polygons for the Contiguous Built Up Areas with the OA and LSOA boundaries. This shows the way that the Contiguous areas are separate enough not to be a single unit using the 50 metre grid but are linked when the threshold is set at 200 metres separation.

To examine the best fit we can consider the 17 LSOA that have at least one OA in the Ammanford Wider Built Up Area. Table A2 shows all the LSOA with the population in each of Built Up Areas and the population outside Built Up Areas.

Map A1

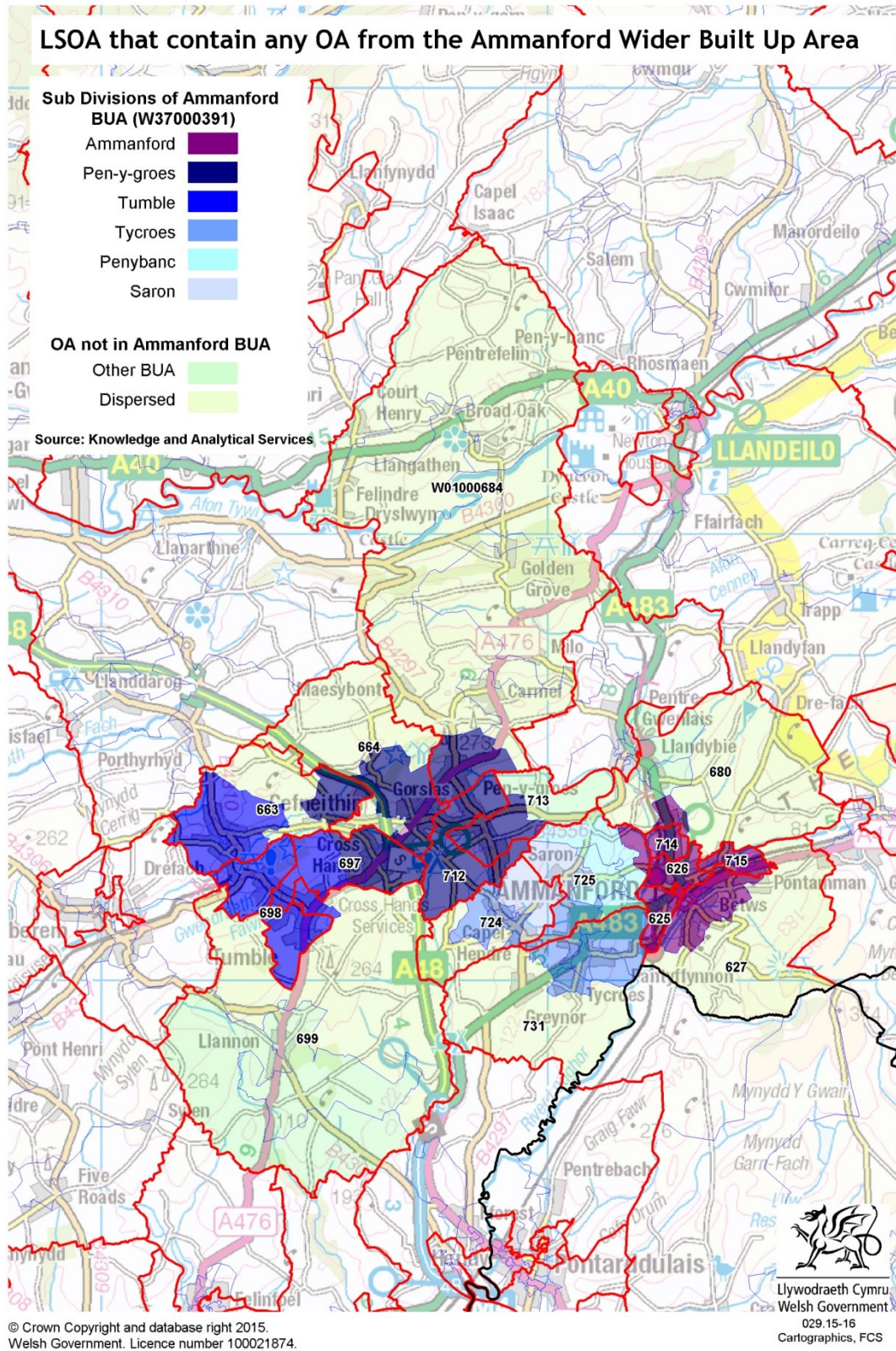


Table A2

LSOA that have at least one OA in the Ammanford Wider Built Up Area

	Within the Ammanford Wider Built Up Area							Outside Ammanford Wider Built Up Area				LSOA total	
	Ammanford	Pen-y-groes	Tumble	Tycroes	Penybanc	Saron	Total	Llanon	Blaenau	Llandybie	No BUA		Total
W01000625	1,200	0	0	0	0	0	1,200	0	0	0	0	0	1,200
W01000626	1,462	0	0	0	0	0	1,462	0	0	0	0	0	1,462
W01000714	1,452	0	0	0	0	0	1,452	0	0	0	0	0	1,452
W01000715	1,297	0	0	0	0	0	1,297	0	0	0	0	0	1,297
W01000627	1,960	0	0	0	0	0	1,960	0	0	0	215	215	2,175
W01000680	376	0	0	0	0	0	376	0	0	1,042	593	1,635	2,011
W01000725	198	0	0	362	1,126	513	2,199	0	0	0	0	0	2,199
W01000712	0	1,496	0	0	0	0	1,496	0	0	0	0	0	1,496
W01000684	0	349	0	0	0	0	349	0	0	0	1,502	1,502	1,851
W01000713	0	968	0	0	0	0	968	0	425	0	0	425	1,393
W01000664	0	1,161	0	0	0	0	1,161	0	0	0	644	644	1,805
W01000663	0	775	1,204	0	0	0	1,979	0	0	0	282	282	2,261
W01000697	0	968	866	0	0	0	1,834	0	0	0	0	0	1,834
W01000698	0	0	1,906	0	0	0	1,906	0	0	0	0	0	1,906
W01000699	0	0	326	0	0	0	326	869	0	0	335	1,204	1,530
W01000731	0	0	0	2,146	0	0	2,146	0	0	0	292	292	2,438
W01000724	0	0	0	1,244	0	354	1,598	0	0	0	314	314	1,912
Total	7,945	5,717	4,302	3,752	1,126	867	23,709	869	425	1,042	4,177	6,513	30,222

Source: ONS classification of Built Up Areas (population figures add up the OA populations within the selection)

The largest element in an LSOA is shown in a box, if it is the only element the box is shaded

Notes on the use of statistical articles

Statistical articles generally relate to one-off analyses for which there are no updates planned, at least in the short-term, and serve to make such analyses available to a wider audience than might otherwise be the case. They are mainly used to publish analyses that are exploratory in some way, for example:

- Introducing a new experimental series of data;
- A partial analysis of an issue which provides a useful starting point for further research but that nevertheless is a useful analysis in its own right;
- Drawing attention to research undertaken by other organisations, either commissioned by the Welsh Government or otherwise, where it is useful to highlight the conclusions, or to build further upon the research;
- An analysis where the results may not be of as high quality as those in our routine statistical releases and bulletins, but where meaningful conclusions can still be drawn from the results.

Where quality is an issue, this may arise in one or more of the following ways:

- being unable to accurately specify the timeframe used (as can be the case when using an administrative source);
- the quality of the data source or data used; or
- other specified reasons.

However, the level of quality will be such that it does not significantly impact upon the conclusions. For example, the exact timeframe may not be central to the conclusions that can be drawn, or it is the order of magnitude of the results, rather than the exact results, that are of interest to the audience.

The analysis presented does not constitute a National Statistic, but may be based on National Statistics outputs and will nevertheless have been subject to careful consideration and detailed checking before publication. An assessment of the strengths and weaknesses in the analysis will be included in the article, for example comparisons with other sources, along with guidance on how the analysis might be used, and a description of the methodology applied.

Articles are subject to the release practices as defined by the release practices protocol, and so, for example, are published on a pre-announced date in the same way as other statistical outputs.

Missing value symbols used in the article follow the standards used in other statistical outputs, as outlined below.

- .. The data item is not available
- . The data item is not applicable
- The data item is not exactly zero, but estimated as zero or less than half the final digit shown
- * The data item is disclosive or not sufficiently robust for publication



All content is available under the Open Government Licence v3.0, except where otherwise stated. <http://www.nationalarchives.gov.uk/doc/open-government-licence/version/3/>