NCRM National Centre for Research Methods

Automated zone design in practice

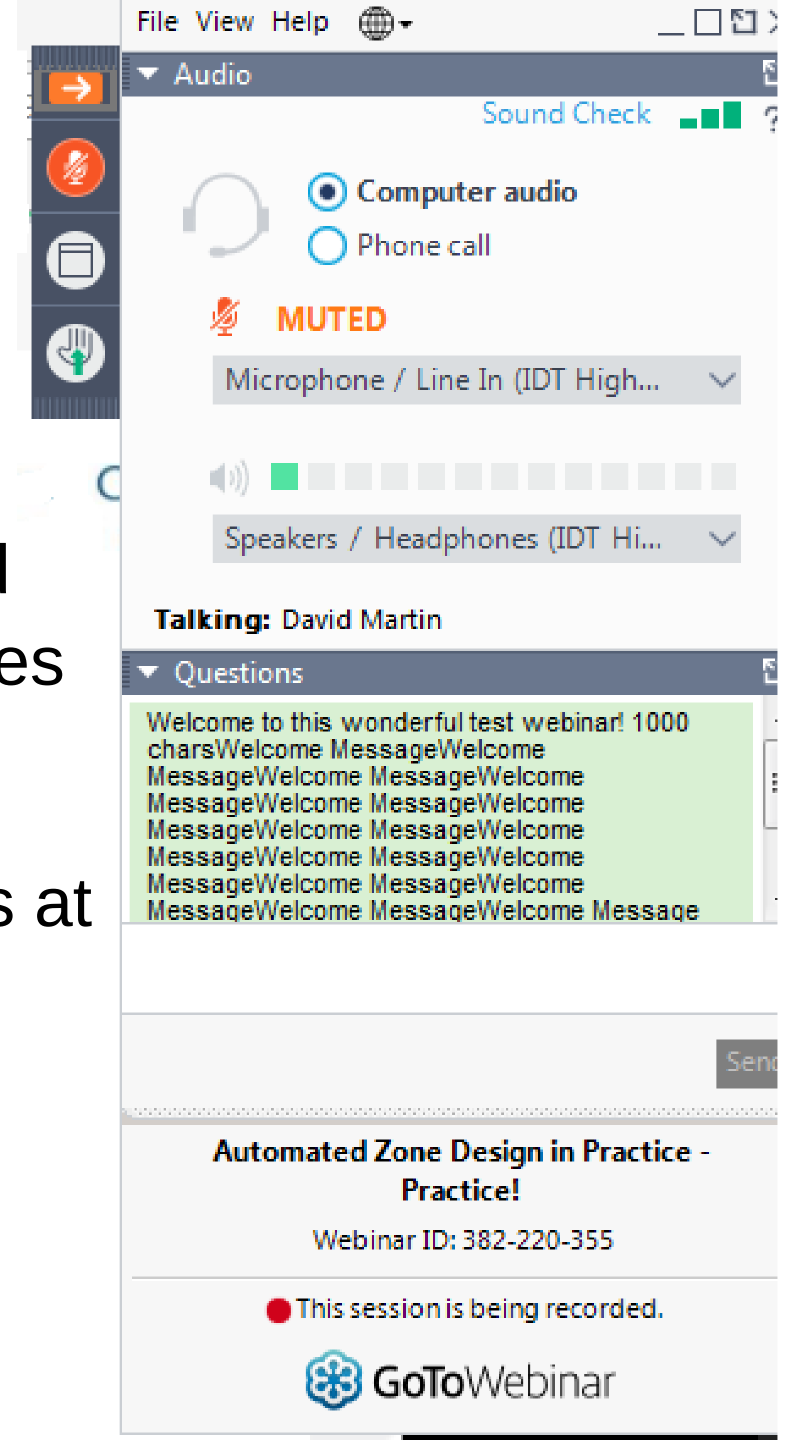
NCRM webinar - 6 June 2017

David Martin

E·S·R·C ECONOMIC & SOCIAL RESEARCH COUNCIL
UNIVERSITY OF Southampton
MANCHESTER 1824 The University of Manchester
THE UNIVERSITY OF EDINBURGH

GoToWebinar basics

- Can view control panel (right)
- Participant microphones muted
- Main webinar presentation slides and speaker audio will be recorded for website
- Participants can type questions at any time
- Questions either addressed immediately or after main presentation



NCRM National Centre for Research Methods

About the presenter

- David Martin
- Professor of Geography, University of Southampton
- Co-Director, NCRM

Automated zone design in practice - outline

- Quick poll
- (Very) brief overview of zone design
- Zone design in practice
- Using AZTool
- NCRM zone design research
- Q&A

Quick poll

- Question 1
  - Have you used the NCRM online resource “Principles of automated zone design?”
- Question 2
  - Have you actually used AZTool before?

NCRM National Centre for Research Methods

(Very) brief overview of zone design

What are zones?

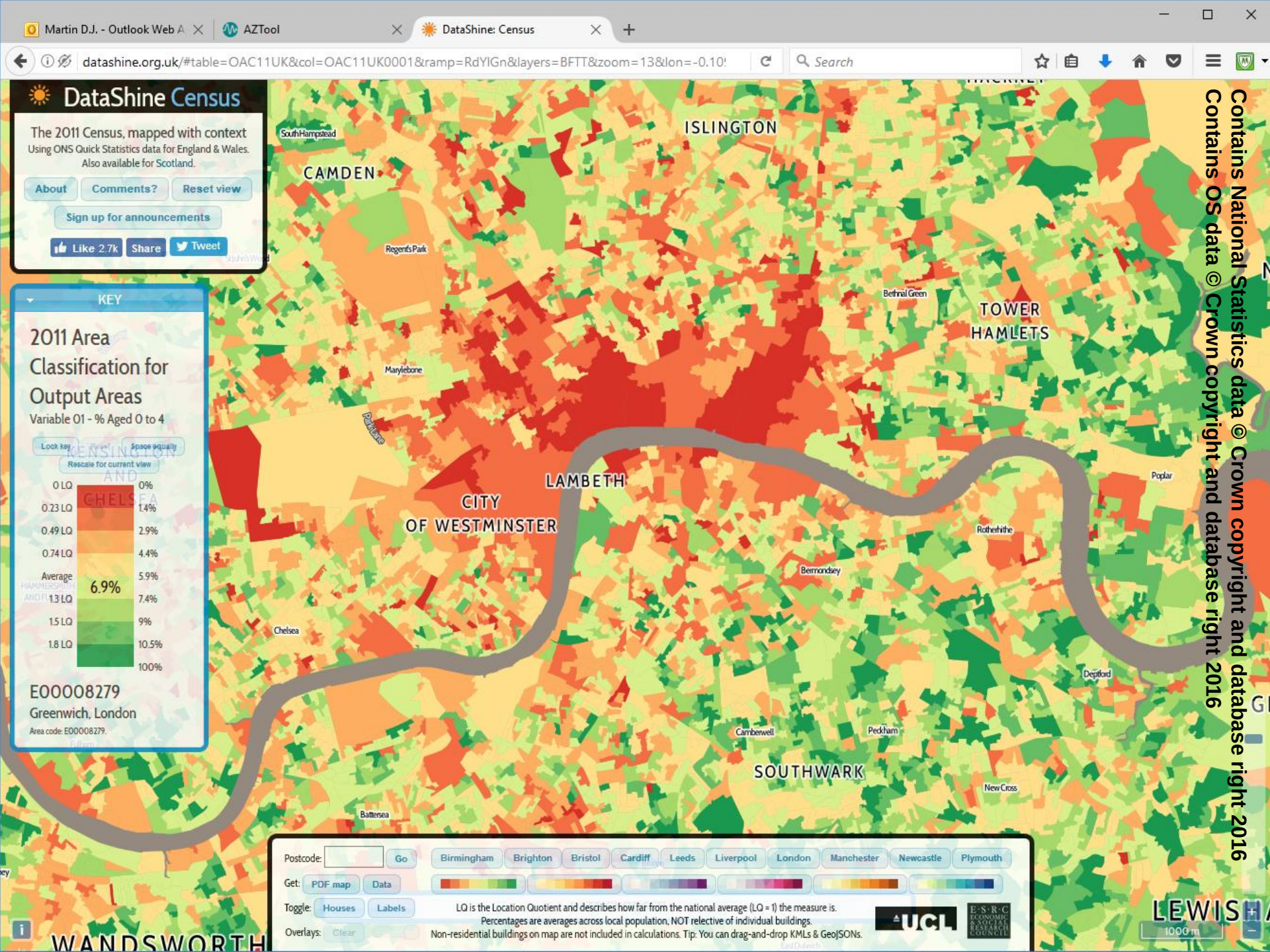
- Divisions of geographical space, usually defined in terms of polygons - often thought of as just the shaded areas on a map
- Usually represented by a tessellation of single polygons, although sometimes islands or separate parts
  - regions, counties, local authorities, wards, electoral districts, constituencies, states (US), communes (France), mesh blocks (Australia), postcode sectors, output areas (UK)

What is zone design?

- Choice of the number and configuration of zones
- Placement of boundaries on a map
- May be result of very careful consideration or a relatively arbitrary process
- Different combinations of historical, administrative processes or an algorithm
- Most obvious applications are publication of statistical results and political districting, but by no means limited to these

Example – 2001 census output areas (England and Wales)

- Matching as far as possible to unit postcodes
- Target size of 125 households
- Always having more than 100 persons and 40 households
- Control over shape (compactness) and social homogeneity (tenure and housing type)
- Used for the publication of small area census statistics


Martin D.J. - Outlook Web A
AZTool
DataShine: Census
datashine.org.uk/#table=OAC11UK&col=OAC11UK0001&ramp=RdYlGn&layers=BFTT&zoom=13&lon=-0.10
Search
DataShine Census
The 2011 Census, mapped with context
Using ONS Quick Statistics data for England & Wales.
Also available for Scotland.
About
Comments?
Reset view
Sign up for announcements
Like 2.7k
Share
Tweet
KEY
2011 Area Classification for Output Areas
Variable 01 - % Aged 0 to 4
Lock key
Space equally
Rescale for current view
0 LQ 0%
0.23 LQ 1.4%
0.49 LQ 2.9%
0.74 LQ 4.4%
Average 6.9% 5.9%
1.3 LQ 7.4%
1.5 LQ 9%
1.8 LQ 10.5%
100%
E00008279
Greenwich, London
Area code: E00008279.
ISLINGTON
CAMDEN
SouthHampstead
Regent's Park
Bethnal Green
TOWER HAMLETS
Marylebone
Paddington
LAMBETH
CITY OF WESTMINSTER
Poplar
Rotherhithe
Bermondsey
Chelsea
Deptford
Camberwell
Peckham
SOUTHWARK
New Cross
Battersea
WANDSWORTH
LEWISHA
Postcode:
Go
Birmingham
Brighton
Bristol
Cardiff
Leeds
Liverpool
London
Manchester
Newcastle
Plymouth
Get:
PDF map
Data
Toggle:
Houses
Labels
Overlays:
Clear
LQ is the Location Quotient and describes how far from the national average (LQ = 1) the measure is.
Percentages are averages across local population, NOT relective of individual buildings.
Non-residential buildings on map are not included in calculations. Tip: You can drag-and-drop KMLs & GeoJSONs.
UCL
E·S·R·C ECONOMIC & SOCIAL RESEARCH COUNCIL
1000 m
Contains National Statistics data © Crown copyright and database right 2016
Contains OS data © Crown copyright and database right 2016

Impact on statistical relationships

- Way in which counts are grouped will have a direct impact on measures such as census counts or election results (gerrymandering)
- Configuration of zone boundaries also affects observed relationships between variables and thus ecological associations
- Different relationships hold at different geographical scales, but also for different aggregations at the same scale

NCRM National Centre for Research Methods

Zone design in practice

Principles

- Need to achieve a set of zones that meet specific design criteria (e.g. not too big, not too small, sensible shapes, similar populations)
- Always a trade-off between competing objectives
- Natural human inclination to draw lines on a map into order to subdivide a larger area
- Alternative approach to put together many small areas as if completing a jigsaw - most computational approaches are closer to this

NCRM National Centre for Research Methods

AZTool: What is it for?

- Aggregation of “building block” polygons into “tract” polygons to best meet design criteria
- Iterative recombination of building blocks from many random starting points to produce a “best” solution, given a specified number of iterations
- One of a range of software implementations for automated zone design that have included Sage, ZDES, ZD2k, AZM

AZTool history

- Developed by David Martin, Samantha Cockings and Andrew Harfoot at the University of Southampton
- Originally based on Openshaw's (1977) Automated Zoning Procedure (AZP)
- Some of the functionality previously available as a Visual Basic 6 program called AZM
- Programmed in .NET environment – should run on any modern Windows PC, freely downloadable

Building blocks

NCRM National Centre for Research Methods

Building blocks

Design constraints

Building blocks

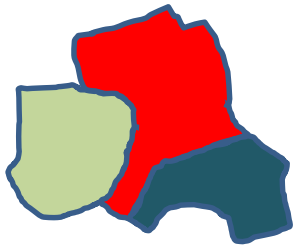
Design constraints

AZTool

Building blocks

Design constraints

AZTool

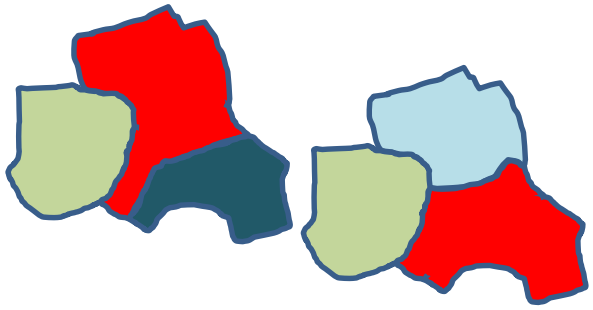


Iterative recombination
and evaluation

Building blocks

Design constraints

AZTool

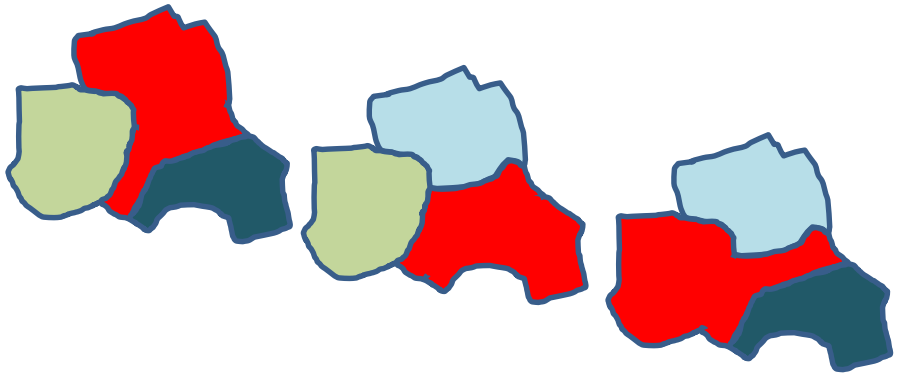


Iterative recombination
and evaluation

Building blocks

Design constraints

AZTool

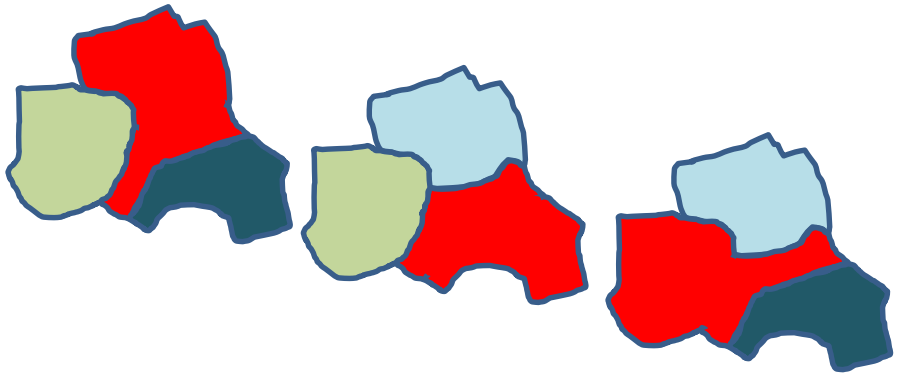


Iterative recombination
and evaluation

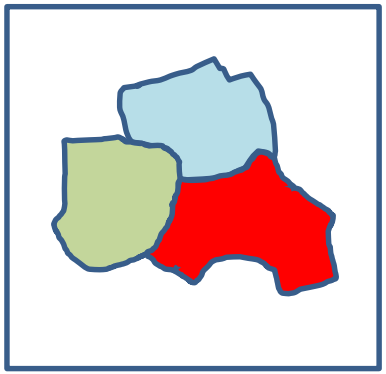
Building blocks

Design constraints

AZTool



Iterative recombination
and evaluation



Output tracts

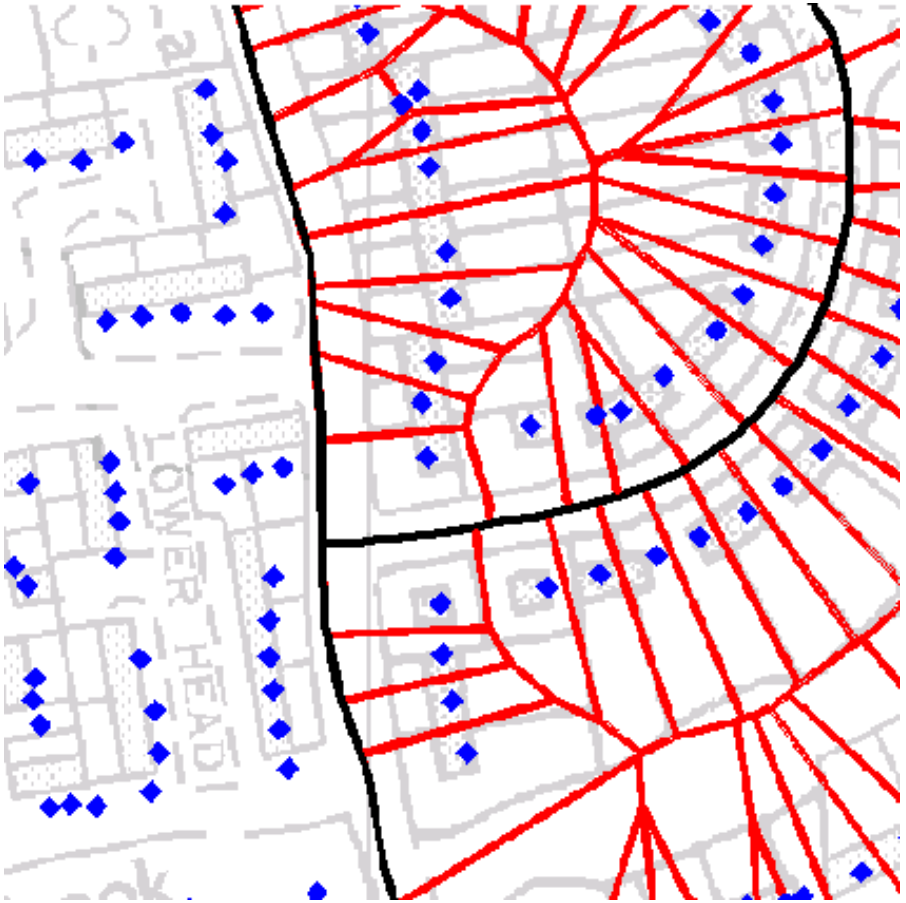
Building blocks

- Building block zones may come from many different sources (purpose-built or pre-existing)
- Need to be small relative to the output zones
- Most will be generated using a geographical information system (GIS)
- Statistical information to evaluate the design criteria needed for each building block
- Could e.g. use confidential data which are to be aggregated for publication

Address polygon generation

Use GIS to generate Thiessen (Voroni) polygons around each address location

These could be building blocks, or could be combined into building blocks



Postcode polygon creation

Dissolve boundaries between address polygons having common postcodes

Advice: remember that output tracts can only be built from these boundaries

Building block considerations

- Big impact on zone design solution
- Output zone boundaries drawn from building block boundaries: cannot be smoother, more realistic or better-aligned to real-world features
- As number of building blocks increases: more permutations, longer computation times and (probably) more good solutions
- The contiguity relationships are the key driver of the zone design process

Zone design criteria

- Hard constraints – must be met
  - e.g. more than 100 people and 40 households
  - e.g. must not cross local authority boundary
- Soft constraints – cannot be met exactly, but should be treated as objectives
  - e.g. zones should contain 125 households
  - e.g. zone should be as compact in shape as possible

Design criteria - measurement

- Suitable statistical measures that can be repeatedly recalculated for each of the design criteria
  - e.g. disallow any zone with a population less than threshold value
  - e.g. solve for all building blocks within an external boundary
  - e.g. minimize sum of {squared differences from target population size}
  - e.g. minimize sum of {perimeter squared/area}

One possible solution

Repeated swapping of building blocks between output tract and re-evaluation of constraints at each step: retain best-performing solution (Advice: don't apply too many constraints!)

Automated zone design in practice

- Automated zone design has been used by the Office for National Statistics in England and Wales:
  - 2001 Output Areas
  - 2004 Lower Layer Super Output Areas
  - 2011 Output Areas
  - 2011 Lower Layer Super Output Areas
  - 2011 Workplace Zones
- Northern Ireland 2001/11 output areas/small areas
- Northern Ireland/Scotland 2011 Workplace Zones

Schizophrenia Research

Available online 15 March 2017

In Press, Corrected Proof— Note to users

Ethnic density, urbanicity and psychosis risk for migrant groups – A population cohort study

Peter Schofield[a], Malene Thygesen[c, d, e], Jay Das-Munshi[b], Laia Becares[f], Elizabeth Cantor-Graae[g], Carsten Pedersen[c, d, e], Esben Agerbo[c, d, e]

Show more

https://doi.org/10.1016/j.schres.2017.03.032

2.5. Neighbourhood effects

Neighbourhood units were based on Danish parishes which vary considerably in size hindering model convergence. For small parishes we therefore combined adjacent units to arrive at an optimum size, using AZtool, the algorithm devised to create UK census area units (Cockings *et al.*, 2011 ; Martin, 2003). We set the algorithm to aim for an optimum parish size of 3000 inhabitants with no units < 200, collapsing 2114 parishes into 1135 units. We also split very large parishes (over 6500) into two, randomly

NCRM National Centre for Research Methods

Using AZTool

AZTool

1. Introduction

This software is provided as-is; use at your own risk! Neither the authors nor their employers can be held responsible for any damages resulting from its use.

AZTool is an automated zone design tool. It takes an input set of geographical "building blocks" and iteratively aggregates them into a number of larger zones optimised to meet user-specified constraints. The software is written in VB.NET using the .NET framework version 2. No GIS software is required to run AZTool, however data preparation and visualisation of the results may require such software.

These notes assume a basic familiarity with zone design problems and GIS operation.

2. History

The zone design approach implemented in AZTool is based on the automated zoning procedure (AZP), first developed by Openshaw (1977a; 1997b) and then enhanced by Openshaw and Rao (1995). AZP was further developed into the AZM software by Martin (2003), which is available here. The algorithm was subsequently used by the Office for National Statistics (ONS) to create the 2001 Census output geographies for England and Wales (Martin et al, 2001). The functionality of AZM was further developed by Cockings and Martin (2005) and has subsequently been used by a wide range of researchers for different applications in various countries. The current version of AZTool was developed by the ESRC-funded Census2011Geog project, in collaboration with ONS.

3. Download / Install

(a) AZTool

Download AZTool Current version 1.0.3 25/8/11

Change history

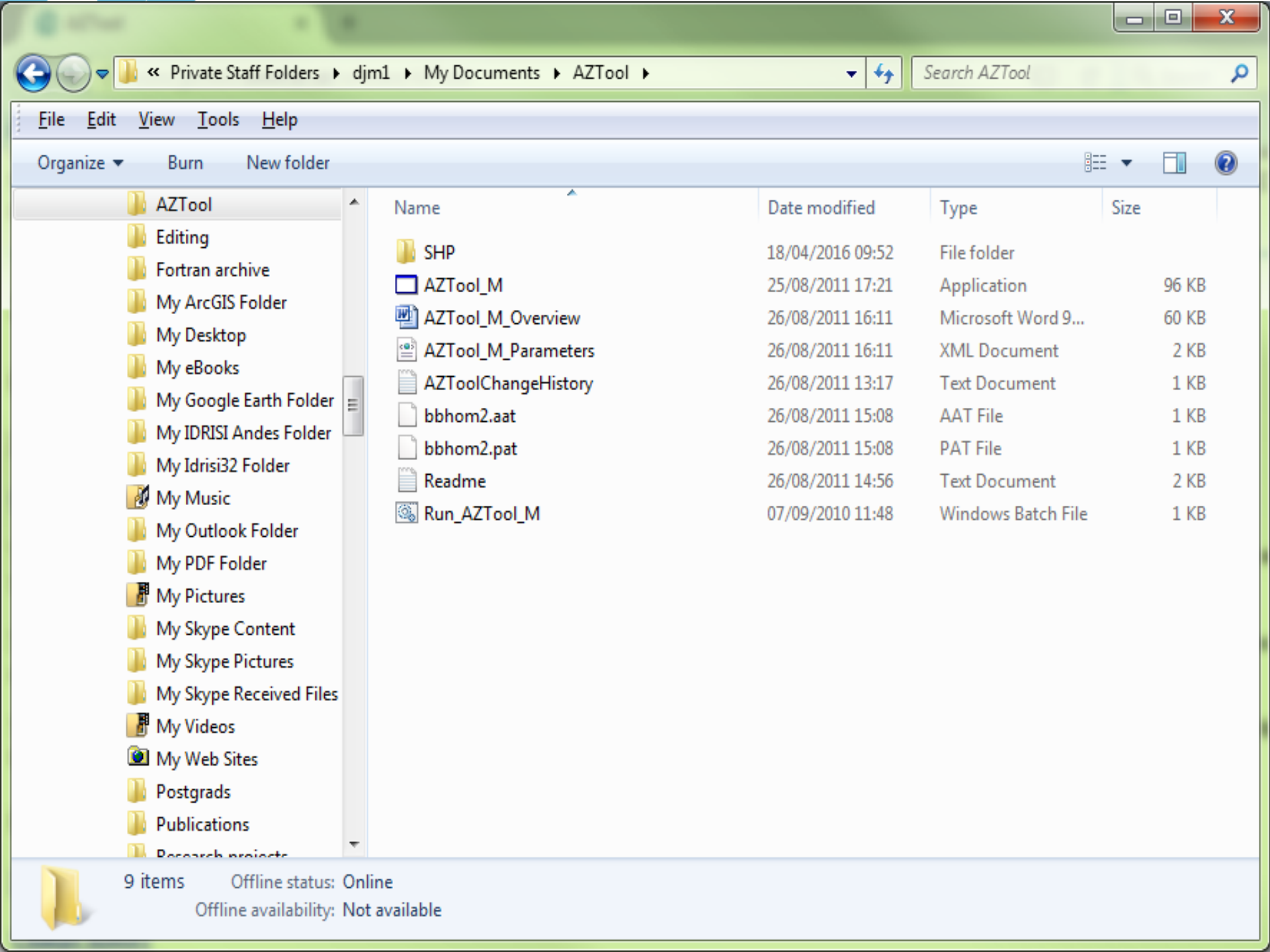
The zip file contains the latest AZTool executable, along with a sample dataset (shapefile), parameter file (.xml), command script (.bat) and description of the tool and parameters in a Word doc. Start with the Readme.txt file!

Note that full paths need to be specified in the AZTool parameter (.xml) file. The sample .xml file is set up to run from the c:\AZT_Demo directory. If you are installing or running AZTool from any other location you will need to specify the full pathnames in the .xml file.

Note that this version of AZTool does not have a full GUI front-end. The options for the program are driven by the .xml parameter file: this makes it more flexible and allows it to be used in batch mode (using the .bat file) for large jobs.

(b) AZTImporter

Download AZTImporter Current version 1.0.1 20/10/10


« Private Staff Folders ▸ djm1 ▸ My Documents ▸ AZTool ▸
Search AZTool
File Edit View Tools Help
Organize Burn New folder
AZTool
Editing
Fortran archive
My ArcGIS Folder
My Desktop
My eBooks
My Google Earth Folder
My IDRISI Andes Folder
My Idrisi32 Folder
My Music
My Outlook Folder
My PDF Folder
My Pictures
My Skype Content
My Skype Pictures
My Skype Received Files
My Videos
My Web Sites
Postgrads
Publications
Name
Date modified
Type
Size
SHP 18/04/2016 09:52 File folder
AZTool_M 25/08/2011 17:21 Application 96 KB
AZTool_M_Overview 26/08/2011 16:11 Microsoft Word 9... 60 KB
AZTool_M_Parameters 26/08/2011 16:11 XML Document 2 KB
AZToolChangeHistory 26/08/2011 13:17 Text Document 1 KB
bbhom2.aat 26/08/2011 15:08 AAT File 1 KB
bbhom2.pat 26/08/2011 15:08 PAT File 1 KB
Readme 26/08/2011 14:56 Text Document 2 KB
Run_AZTool_M 07/09/2010 11:48 Windows Batch File 1 KB
9 items Offline status: Online
Offline availability: Not available

Input files (1)

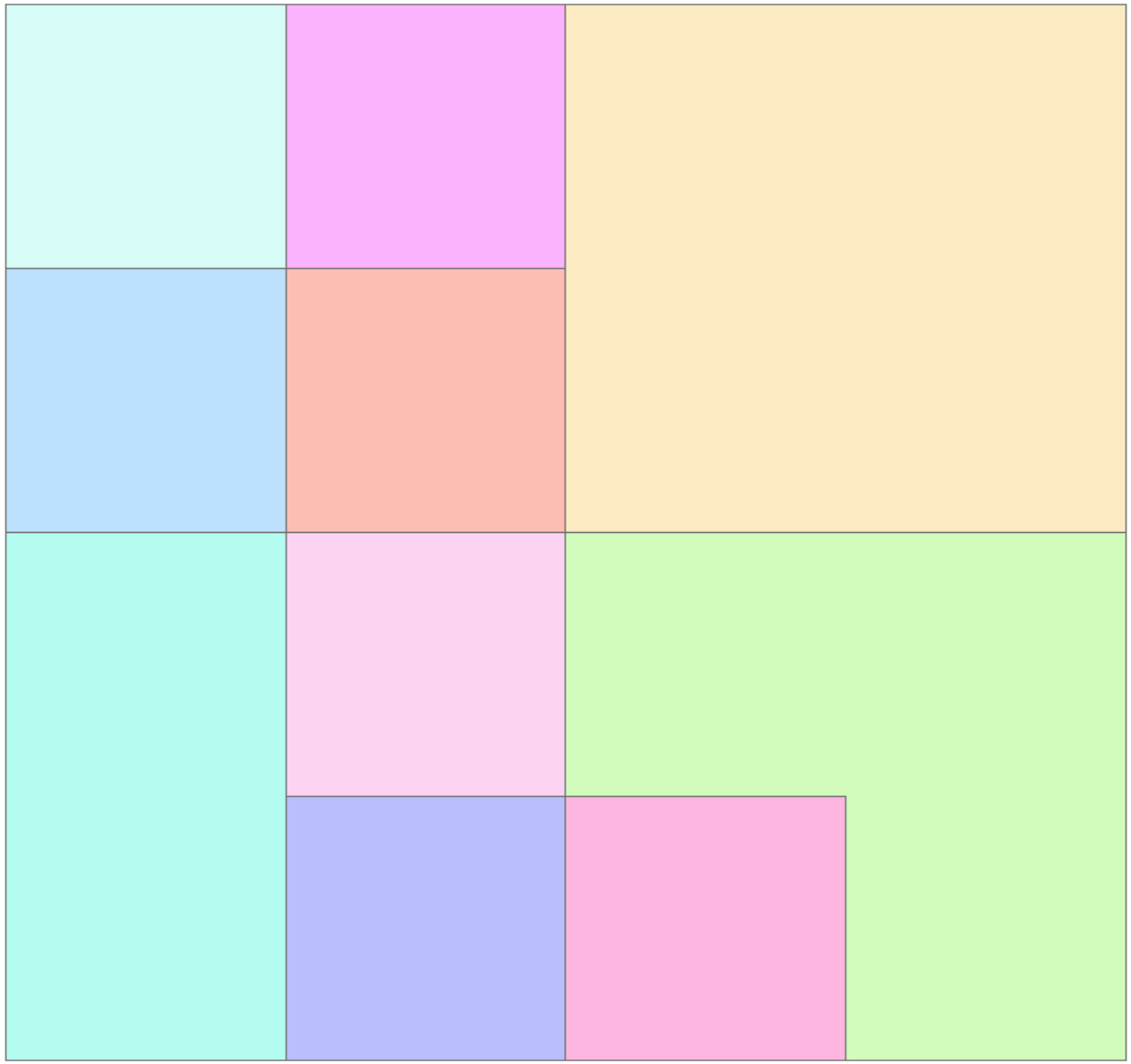
- A set of building blocks and associated data. These are specified as .aat and .pat files
  - The arc attributes describe which building blocks are contiguous
  - No coordinates are needed, but the contiguity information and attributes of each polygon relevant to the design criteria are required

Untitled - ArcMap

File Edit View Bookmarks Insert Selection Geoprocessing Customize Windows Help

Table Of Contents

- Layers
  - \\soton.ac.uk\ude\personalf
    - bbhom2
      - <all other values>
      - BBHOM2_ID
      - 2
      - 3
      - 4
      - 5
      - 6
      - 7
      - 8
      - 9
      - 10
      - 11



-56.086 132.308 Unknown Units

bbhom2 - Notepad

File Edit Format View Help

```
0,1,100
0,3,100
0,4,0
0,10,200
1,2,100
1,3,0
1,4,100
1,10,100
2,4,100
2,6,0
2,7,200
2,10,400
3,4,100
3,5,100
3,6,0
3,10,100
4,5,0
4,6,100
4,7,0
5,6,100
5,8,100
5,10,300
6,7,100
6,8,100
6,9,0
7,8,0
7,9,200
7,10,300
8,9,100
8,10,100
9,10,100
```

bbhom2 - Notepad

File Edit Format View Help

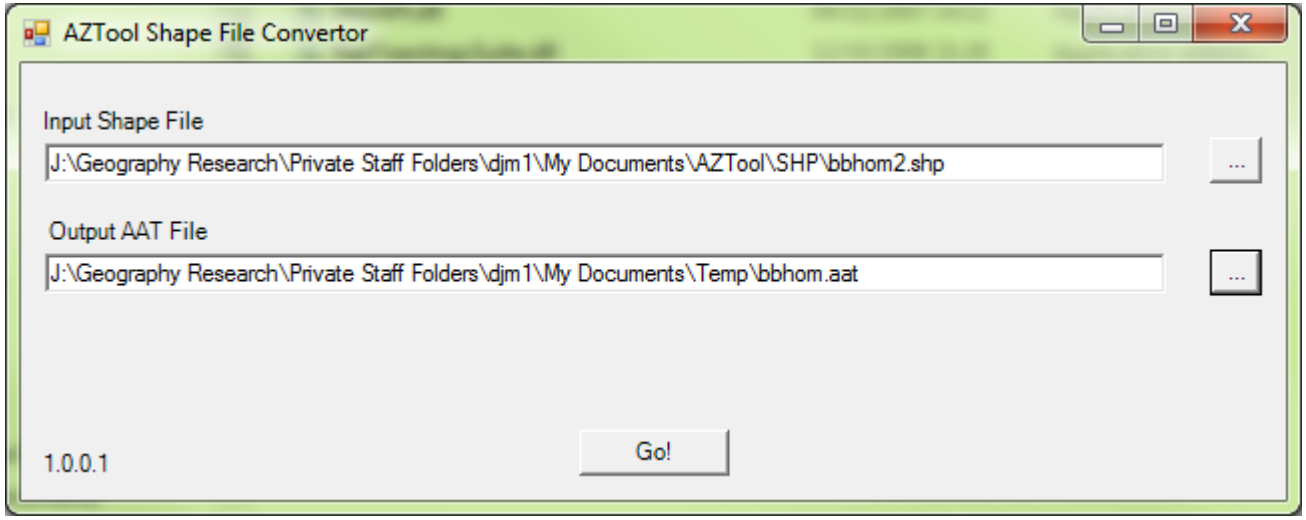
```
AZM_ID,AREA,PERIMETER,BBHOM2_ID,BBPOP,OWNOCC,PRENT,HARENT,DET,SEMI,FLAT,AZM_Area
0,10000,400,2,100,80,10,10,100,0,0,10000
1,10000,400,3,100,78,12,10,90,5,5,10000
2,40000,800,6,100,0,100,0,15,10,75,40000
3,10000,400,4,100,75,15,10,80,10,10,10000
4,10000,400,5,100,70,10,20,90,0,10,10000
5,20000,600,7,50,20,20,10,15,15,20,20000
6,10000,400,9,25,10,10,5,10,10,5,10000
7,30000,800,11,75,0,75,0,10,60,5,30000
8,10000,400,8,25,10,5,10,10,10,5,10000
9,10000,400,10,25,10,5,10,5,10,10,10000
10,0,0,0,0,0,0,0,0,0,0,-1
```

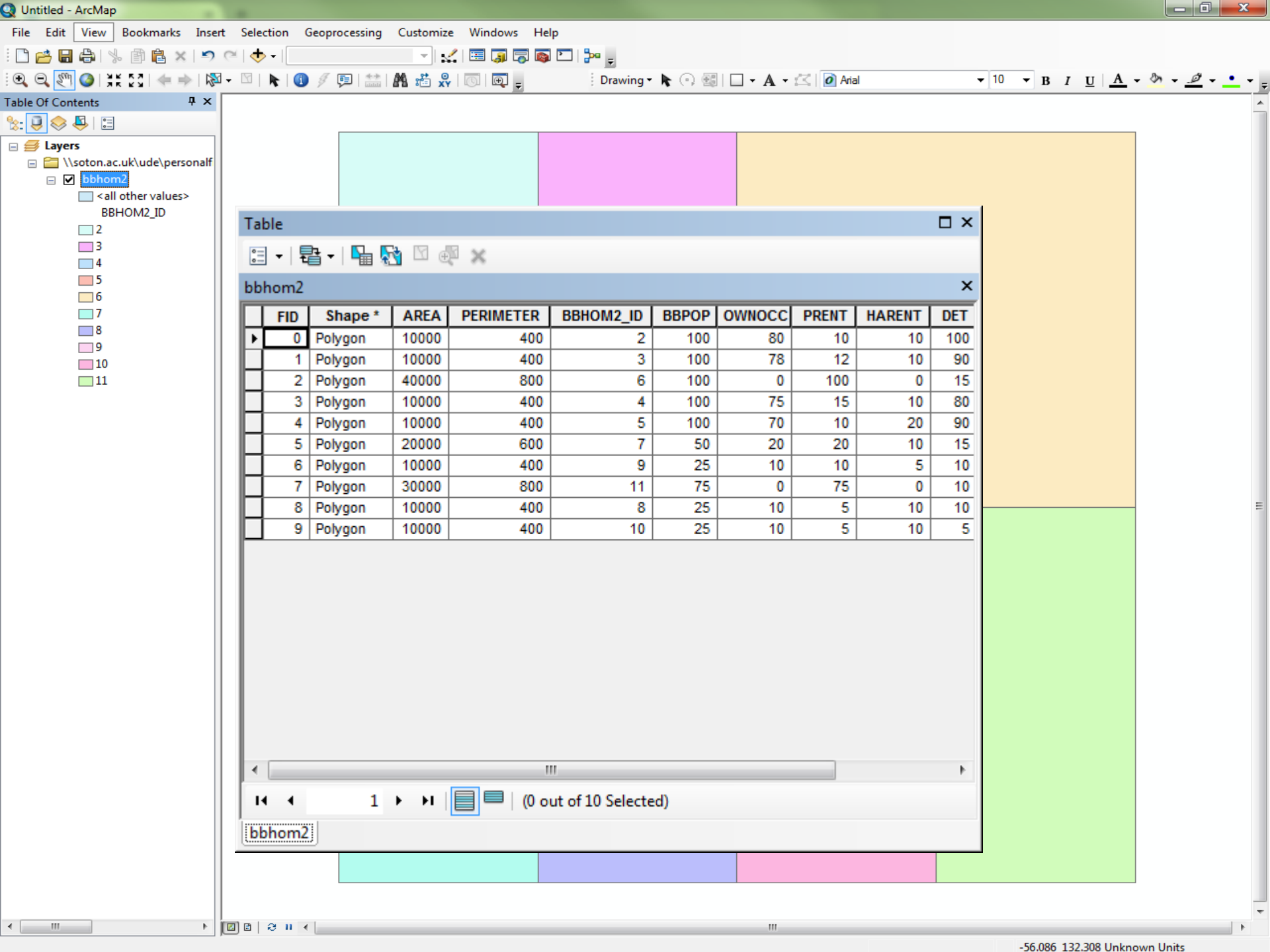
Input files (2)

- Attributes for each building block might include:
  - Population (to be used as a target and/or min/max thresholds)
  - Region (e.g. a larger area within which zones are to be constrained)
  - Homogeneity variables (e.g. tenure or accommodation type, for designing zones which are as internally homogenous as possible)

AZTImporter

- If needed, the AZTImporter program will generate .aat and .pat files from ESRI Shapefile GIS format



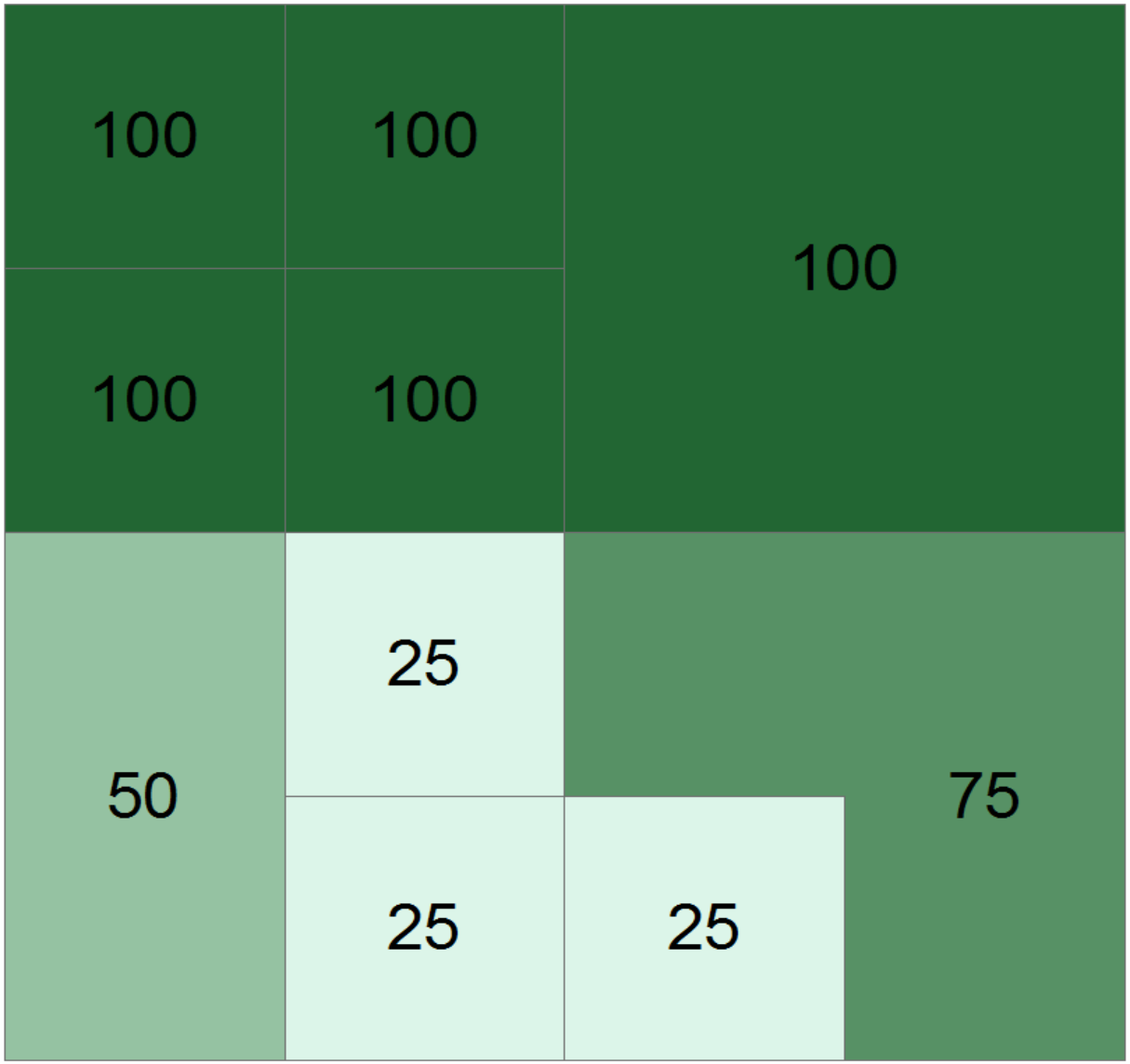

Untitled - ArcMap
File Edit View Bookmarks Insert Selection Geoprocessing Customize Windows Help
Table Of Contents
Layers
\\soton.ac.uk\ude\personalf
bbhom2
<all other values>
BBHOM2_ID
2
3
4
5
6
7
8
9
10
11
Table
bbhom2
FID Shape * AREA PERIMETER BBHOM2_ID BBPOP OWNOCC PRENT HARENT DET
0 Polygon 10000 400 2 100 80 10 10 100
1 Polygon 10000 400 3 100 78 12 10 90
2 Polygon 40000 800 6 100 0 100 0 15
3 Polygon 10000 400 4 100 75 15 10 80
4 Polygon 10000 400 5 100 70 10 20 90
5 Polygon 20000 600 7 50 20 20 10 15
6 Polygon 10000 400 9 25 10 10 5 10
7 Polygon 30000 800 11 75 0 75 0 10
8 Polygon 10000 400 8 25 10 5 10 10
9 Polygon 10000 400 10 25 10 5 10 5
1
(0 out of 10 Selected)
bbhom2
-56.086 132.308 Unknown Units

Untitled - ArcMap

File Edit View Bookmarks Insert Selection Geoprocessing Customize Windows Help

Table Of Contents

Layers
\\soton.ac.uk\ude\personalf
bbhom2
BBPOP
25
26 - 50
51 - 75
76 - 100



-54.593 173.62 Unknown Units

Parameter file

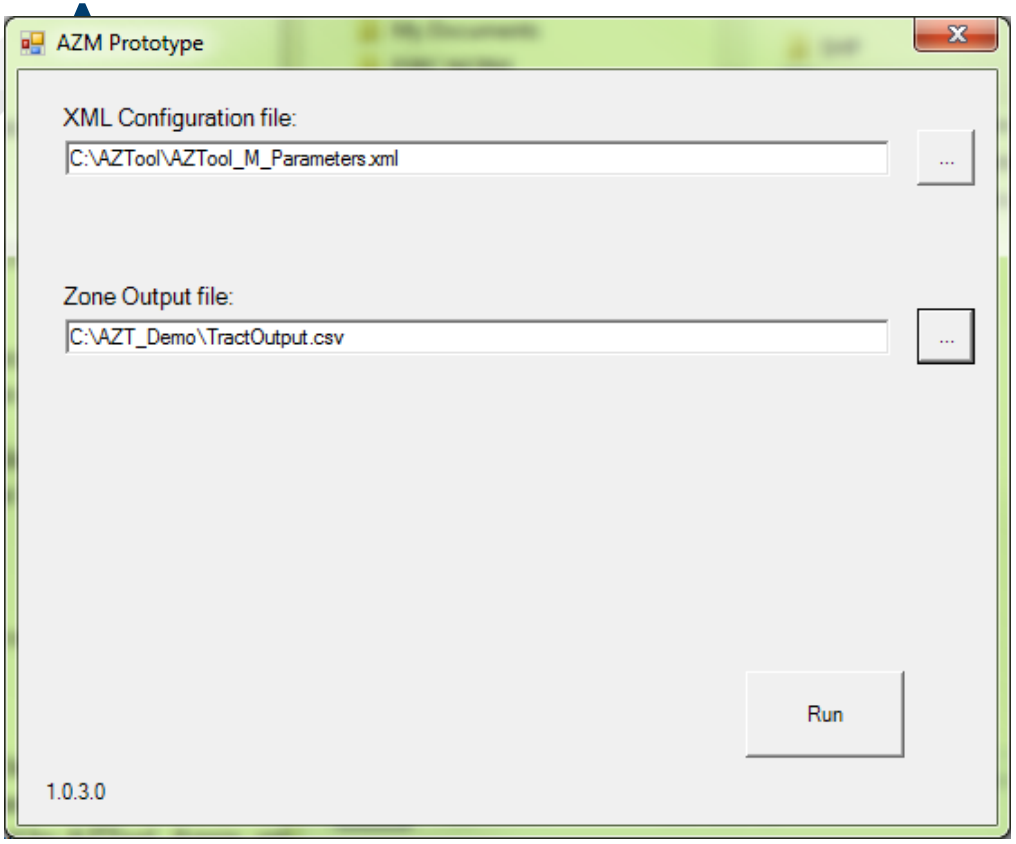
- An XML file containing the program run parameters. This can be edited, saved and re-used.
- Contains all necessary program control parameters for setup, specification and output
- For use in batch mode using a Windows Batch File

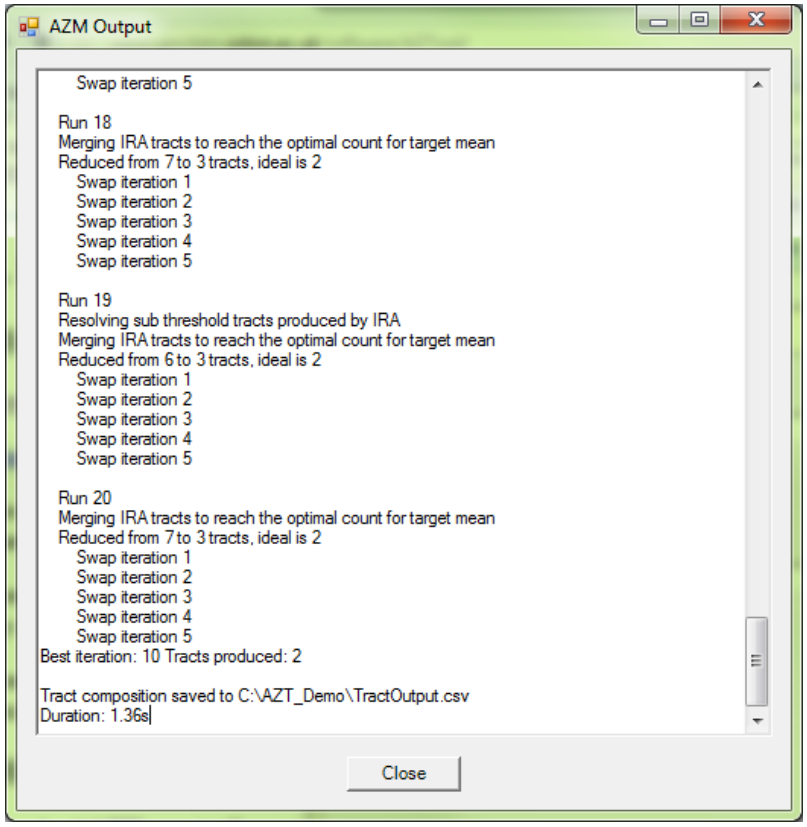
J:\Geography Research\Private Staff Folders\djm1\My Documents\AZTool\AZTool_M_Parar

J:\Geography Research\Priv...

```
<?xml version="1.0" encoding="UTF-8"?>
- <ProgramOptions xmlns:xsd="http://www.w3.org/2001/XMLSchema" xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance">
    <InputPATFile>C:\AZT_Demo\bbhom2.pat</InputPATFile>
    <InputAATFile>C:\AZT_Demo\bbhom2.aat</InputAATFile>
    <Header>true</Header>
    <IDIndex>1</IDIndex>
    <RegionIndex>0</RegionIndex>
    <RegionToUse>ALL</RegionToUse>
    <RespectRegions>false</RespectRegions>
  - <TargThreshVars>
    - <TargetThresholdVar>
        <Name>Population</Name>
        <FileIndex>5</FileIndex>
        <TargetSet>true</TargetSet>
        <Target>300</Target>
        <Tolerance>1000000</Tolerance>
        <Weight>100</Weight>
        <MinThreshSet>true</MinThreshSet>
        <MinThresh>100</MinThresh>
        <MaxThreshSet>true</MaxThreshSet>
        <MaxThresh>625</MaxThresh>
      </TargetThresholdVar>
    </TargThreshVars>
    <IACStartIndex>6</IACStartIndex>
    <IACSet>true</IACSet>
    <IACWeight>100</IACWeight>
  - <IACvarGroups>
    - <IACvarGroup>
        <Name>Tenure</Name>
        <Weight>100</Weight>
        <CategoryCount>3</CategoryCount>
      </IACvarGroup>
    - <IACvarGroup>
        <Name>AccomType</Name>
        <Weight>100</Weight>
        <CategoryCount>3</CategoryCount>
      </IACvarGroup>
    </IACvarGroups>
    <AreaIndex>12</AreaIndex>
    <P2ASet>true</P2ASet>
    <P2AWeight>100</P2AWeight>
    <MinBdyLenSet>false</MinBdyLenSet>
    <MinBdyLenPerc>10</MinBdyLenPerc>
    <IgnoreBishopsContig>true</IgnoreBishopsContig>
    <AllowDonuts>false</AllowDonuts>
    <IRATargetBasedTractCount>true</IRATargetBasedTractCount>
    <TestSpreadsheetReqd>false</TestSpreadsheetReqd>
    <ReportStatistics>false</ReportStatistics>
    <NumberSwapIterations>5</NumberSwapIterations>
    <NumberRuns>20</NumberRuns>
    <UseLogDomainScores>false</UseLogDomainScores>
    <IgnoreTractsWithUnbreachedBB>false</IgnoreTractsWithUnbreachedBB>
    <RandomSeed>0</RandomSeed>
</ProgramOptions>
```

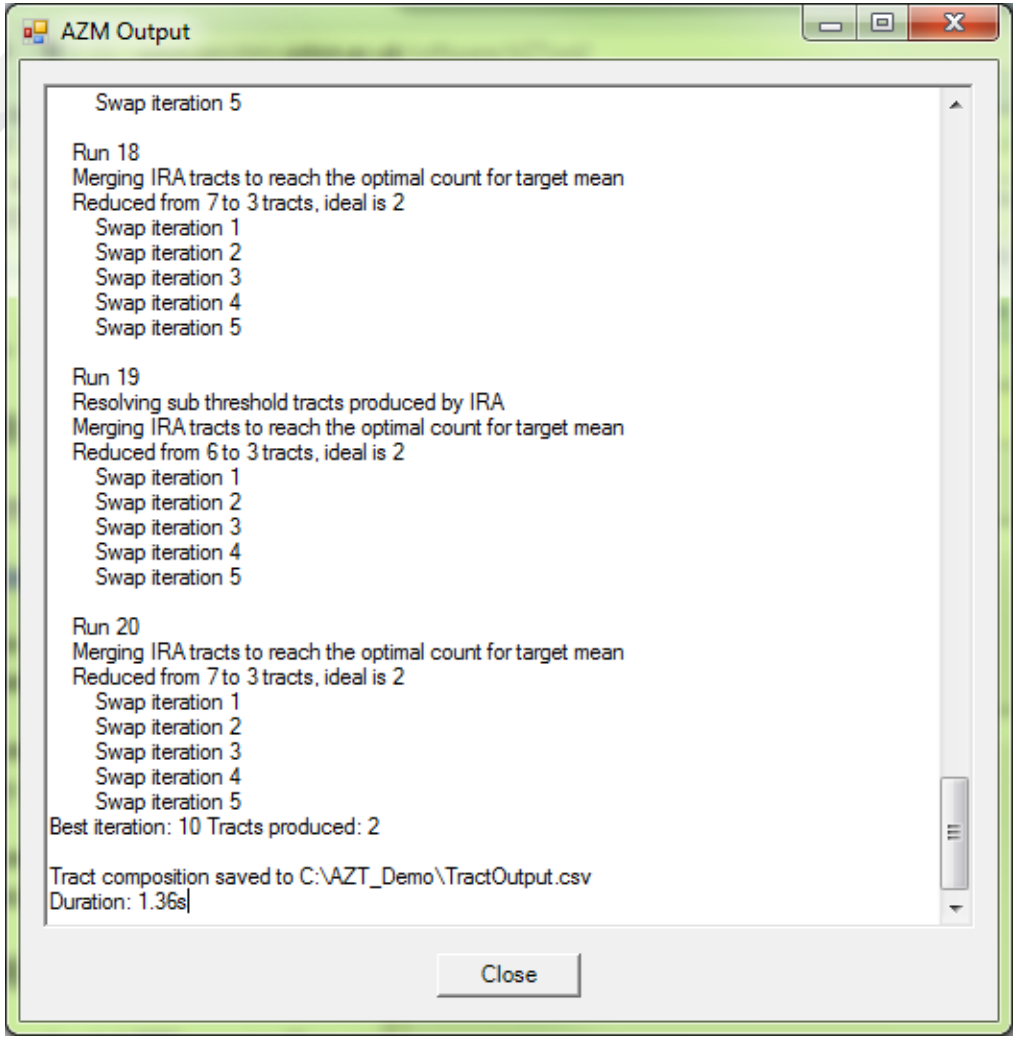
```
- <TargThreshVars>
    - <TargetThresholdVar>
          <Name>Population</Name>
          <FileIndex>5</FileIndex>
          <TargetSet>true</TargetSet>
          <Target>300</Target>
          <Tolerance>1000000</Tolerance>
          <Weight>100</Weight>
          <MinThreshSet>true</MinThreshSet>
          <MinThresh>100</MinThresh>
          <MaxThreshSet>true</MaxThreshSet>
          <MaxThresh>625</MaxThresh>
      </TargetThresholdVar>
  </TargThreshVars>
```

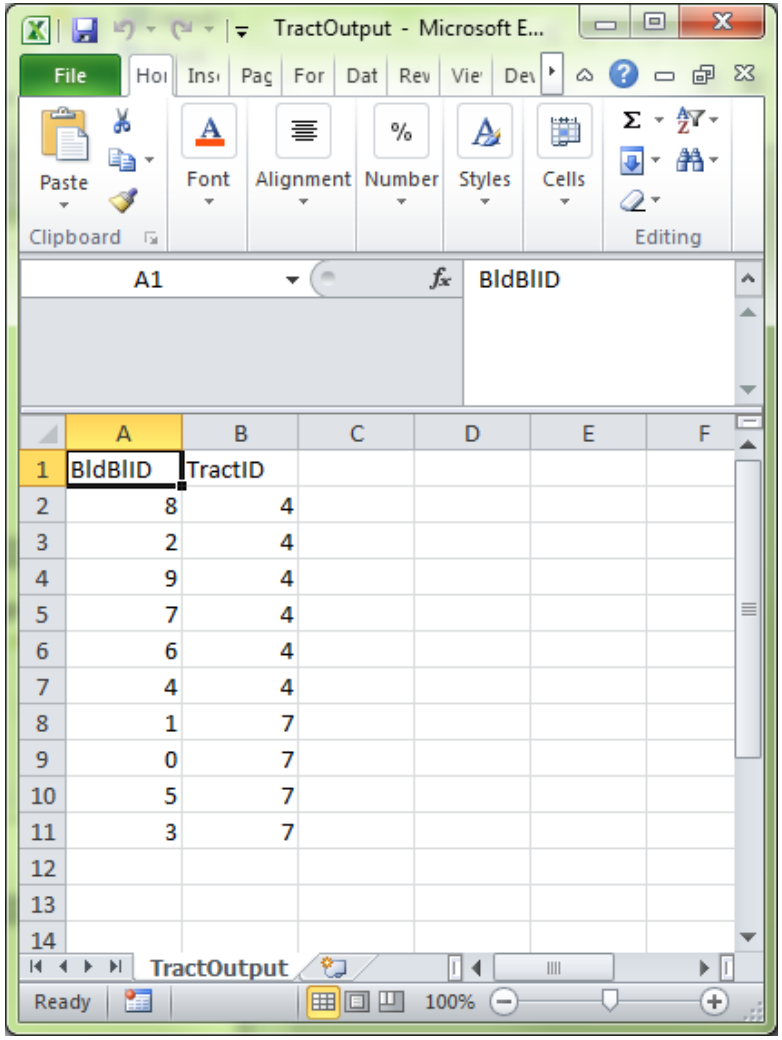

AZM Prototype
XML Configuration file:
C:\AZTool\AZTool_M_Parameters.xml
...
Zone Output file:
C:\AZT_Demo\TractOutput.csv
...
Run
1.0.3.0

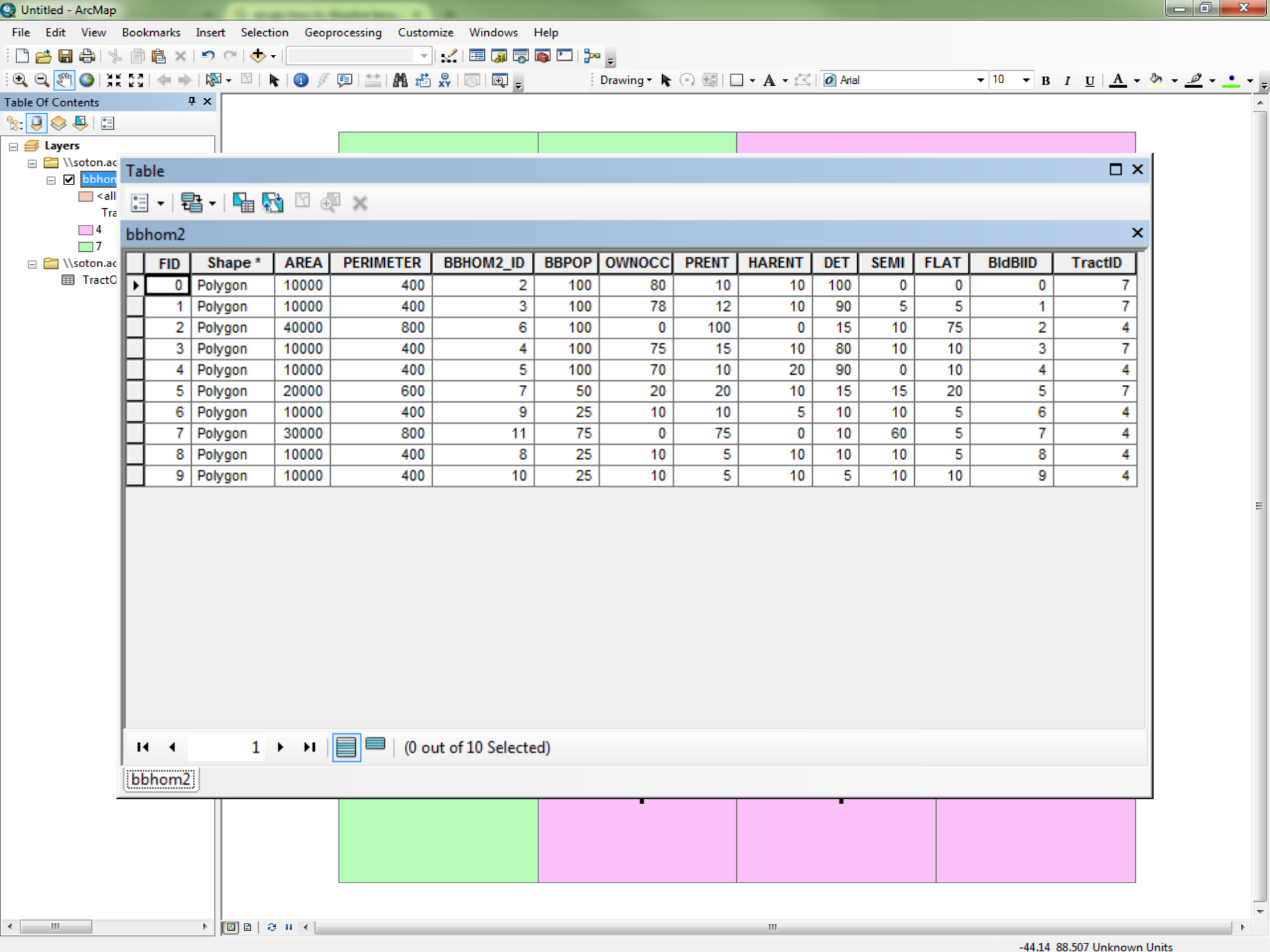

AZM Output
Swap iteration 5
Run 18
Merging IRA tracts to reach the optimal count for target mean
Reduced from 7 to 3 tracts, ideal is 2
Swap iteration 1
Swap iteration 2
Swap iteration 3
Swap iteration 4
Swap iteration 5
Run 19
Resolving sub threshold tracts produced by IRA
Merging IRA tracts to reach the optimal count for target mean
Reduced from 6 to 3 tracts, ideal is 2
Swap iteration 1
Swap iteration 2
Swap iteration 3
Swap iteration 4
Swap iteration 5
Run 20
Merging IRA tracts to reach the optimal count for target mean
Reduced from 7 to 3 tracts, ideal is 2
Swap iteration 1
Swap iteration 2
Swap iteration 3
Swap iteration 4
Swap iteration 5
Best iteration: 10 Tracts produced: 2
Tract composition saved to C:\AZT_Demo\TractOutput.csv
Duration: 1.36s
Close

Output files

- A .txt format log file, reporting progress of the program run and identifying any problems, e.g. with the input data
- A .csv format results file, showing the output tract to which each building blocks has been assigned
- Zoning results can be re-imported to GIS and used to dissolve boundaries between building blocks


AZM Output
Swap iteration 5
Run 18
Merging IRA tracts to reach the optimal count for target mean
Reduced from 7 to 3 tracts, ideal is 2
Swap iteration 1
Swap iteration 2
Swap iteration 3
Swap iteration 4
Swap iteration 5
Run 19
Resolving sub threshold tracts produced by IRA
Merging IRA tracts to reach the optimal count for target mean
Reduced from 6 to 3 tracts, ideal is 2
Swap iteration 1
Swap iteration 2
Swap iteration 3
Swap iteration 4
Swap iteration 5
Run 20
Merging IRA tracts to reach the optimal count for target mean
Reduced from 7 to 3 tracts, ideal is 2
Swap iteration 1
Swap iteration 2
Swap iteration 3
Swap iteration 4
Swap iteration 5
Best iteration: 10 Tracts produced: 2
Tract composition saved to C:\AZT_Demo\TractOutput.csv
Duration: 1.36s
Close


TractOutput - Microsoft E...
File
Paste
Font
Alignment
Number
Styles
Cells
Editing
Clipboard
A1
BldBlID
BldBlID TractID
8 4
2 4
9 4
7 4
6 4
4 4
1 7
0 7
5 7
3 7
TractOutput
Ready
100%

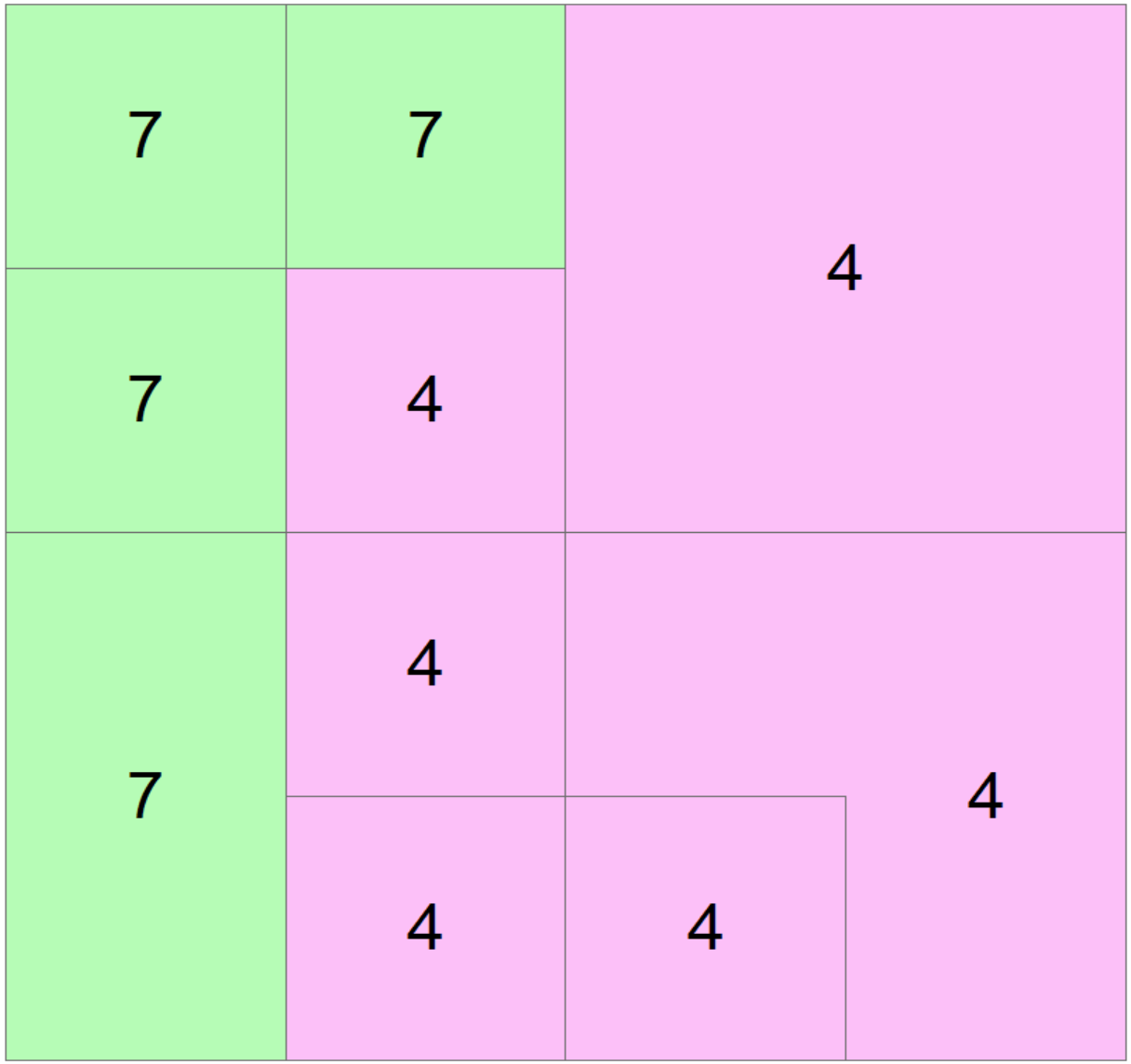

Untitled - ArcMap
File Edit View Bookmarks Insert Selection Geoprocessing Customize Windows Help
Table Of Contents
Layers
Table
bbhom2
FID Shape * AREA PERIMETER BBHOM2_ID BBPOP OWNOCC PRENT HARENT DET SEMI FLAT BldBlID TractID
0 Polygon 10000 400 2 100 80 10 10 100 0 0 0 7
1 Polygon 10000 400 3 100 78 12 10 90 5 5 1 7
2 Polygon 40000 800 6 100 0 100 0 15 10 75 2 4
3 Polygon 10000 400 4 100 75 15 10 80 10 10 3 7
4 Polygon 10000 400 5 100 70 10 20 90 0 10 4 4
5 Polygon 20000 600 7 50 20 20 10 15 15 20 5 7
6 Polygon 10000 400 9 25 10 10 5 10 10 5 6 4
7 Polygon 30000 800 11 75 0 75 0 10 60 5 7 4
8 Polygon 10000 400 8 25 10 5 10 10 10 5 8 4
9 Polygon 10000 400 10 25 10 5 10 5 10 10 9 4
1
(0 out of 10 Selected)
bbhom2
-44.14 88.507 Unknown Units

Untitled - ArcMap

File Edit View Bookmarks Insert Selection Geoprocessing Customize Windows Help

Table Of Contents

- Layers
  - \\soton.ac.uk\ude\personalf
    - bbhom2
      - <all other values>
      - TractID
      - 4
      - 7
  - \\soton.ac.uk\ude\personalf
    - TractOutput.csv



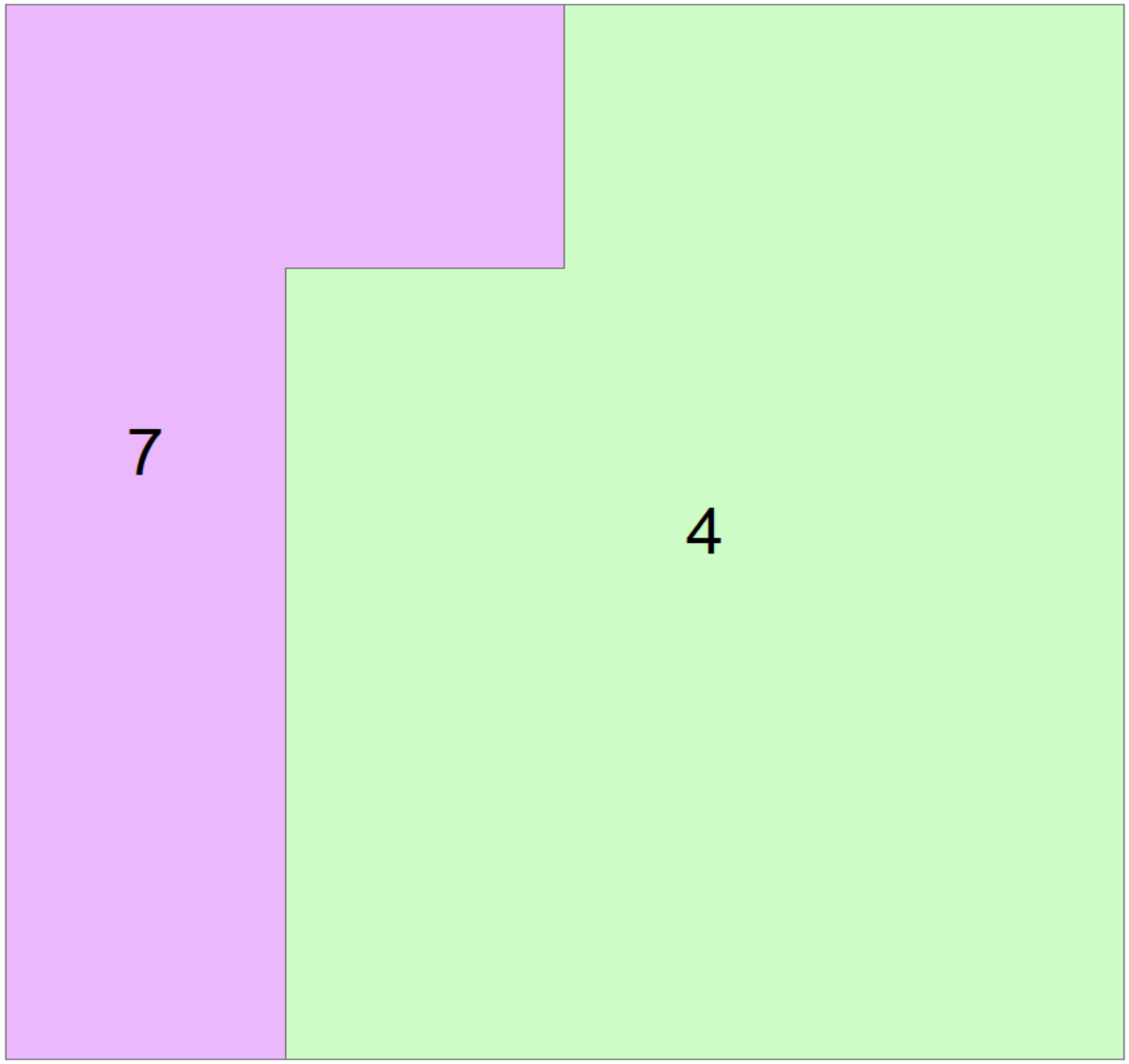
-44.14 88.507 Unknown Units

Untitled - ArcMap

File Edit View Bookmarks Insert Selection Geoprocessing Customize Windows Help

Table Of Contents

- Layers
  - \\soton.ac.uk\ude\personalf
    - bbhom2_Dissolve
      - <all other values>
        - TractOutput_csv_Trac
      - 4
      - 7
  - \\soton.ac.uk\ude\personalf
    - bbhom2
      - <all other values>
  - \\soton.ac.uk\ude\personalf
    - TractOutput.csv



Untitled - ArcMap

File Edit View Bookmarks Insert Selection Geoprocessing Customize Windows Help

Table Of Contents

Layers

\\soton.ac.uk\ude\personalfiles\user

bbhom2_Dissolve

Field2

350

\\soton.ac.uk\ude\personalfiles\user

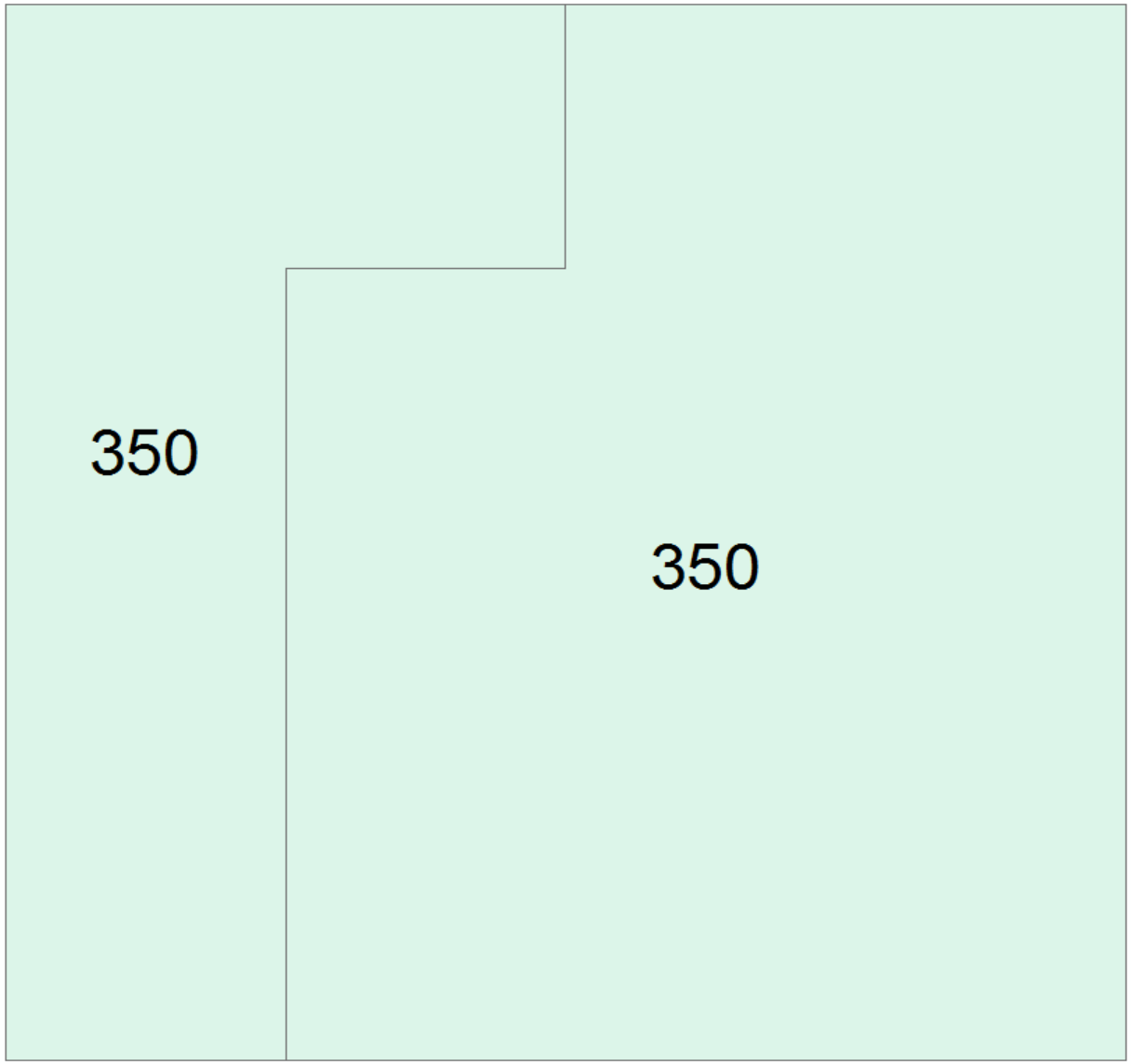
bbhom2

<all other values>

\\soton.ac.uk\ude\personalfiles\user

TractOutput.csv

sums.csv



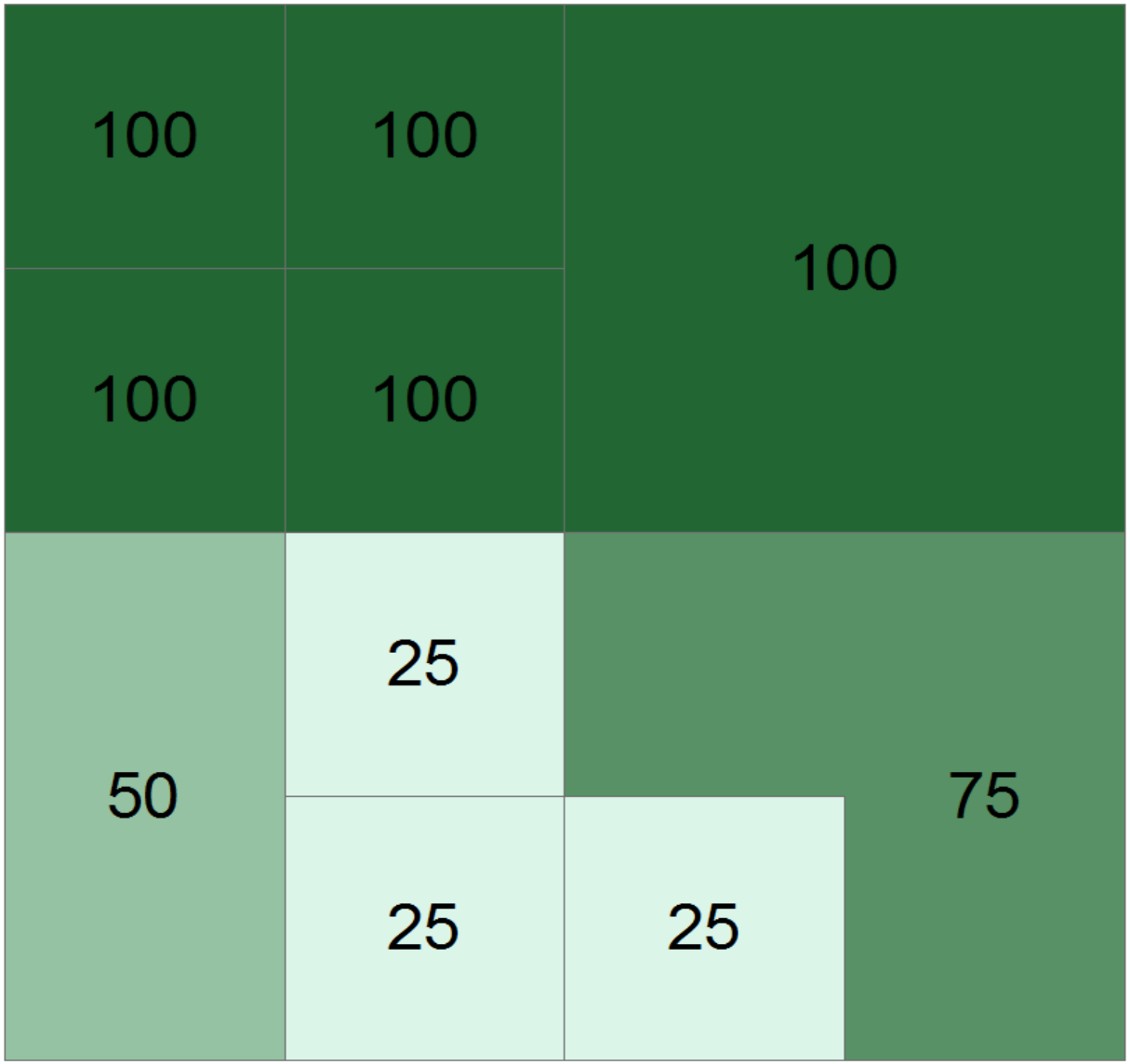
78.303 86.516 Unknown Units

Untitled - ArcMap

File Edit View Bookmarks Insert Selection Geoprocessing Customize Windows Help

Table Of Contents

Layers
\\soton.ac.uk\ude\personalf
bbhom2
BBPOP
25
26 - 50
51 - 75
76 - 100



-54.593 173.62 Unknown Units

AZTool design constraints

- Constraint within higher level regions
- Population targets and thresholds
- Shape compactness
- Intra-area correlation measures
- New accessibility/network connectivity measures

NCRM National Centre for Research Methods

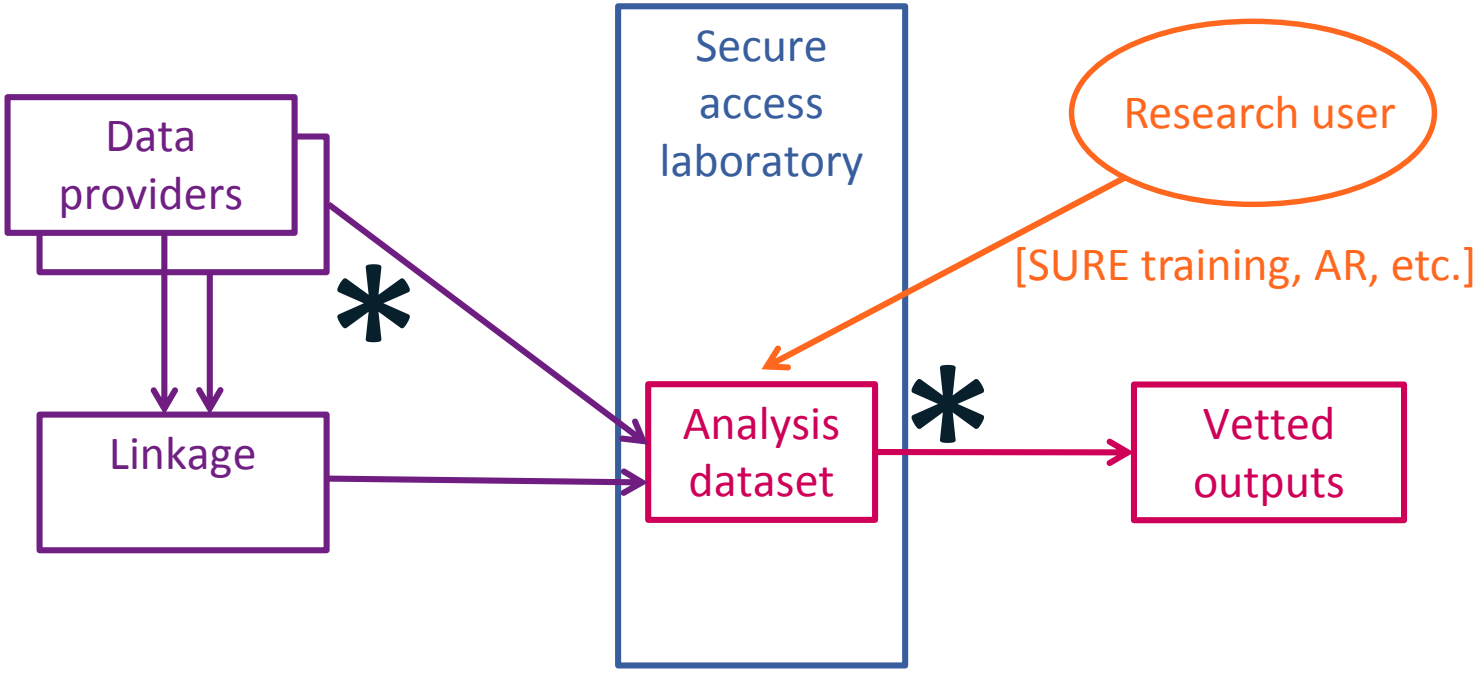
NCRM zone design research

Current zone design research

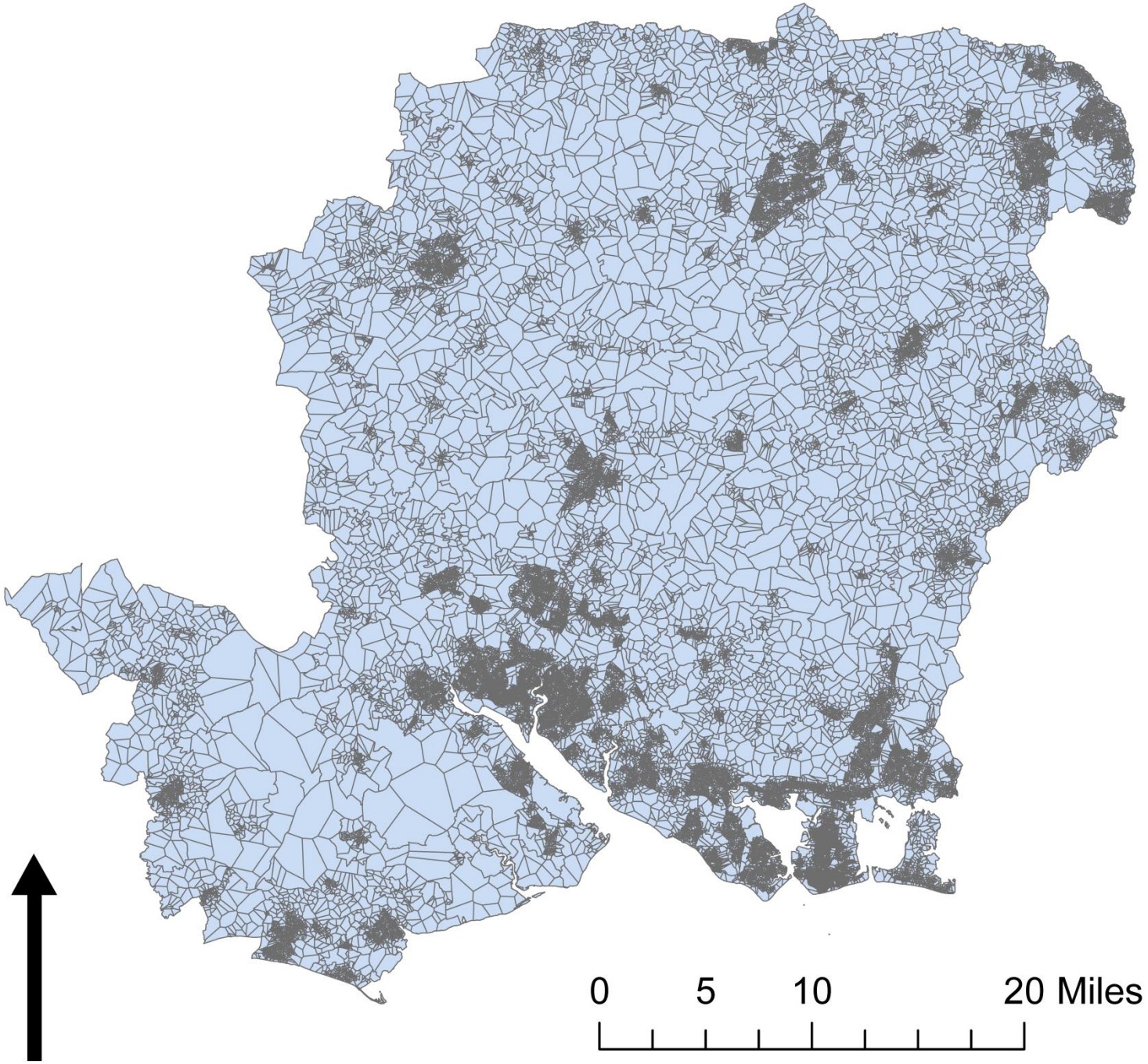
- Rezoning of a synthetic spatial microdataset to explore
  - Modifiable areal unit effects
  - Impact of statistical disclosure control measures
- Work with ONS on redesign of data collection geographies
  - Potential new census workload areas
  - Potential new survey areas

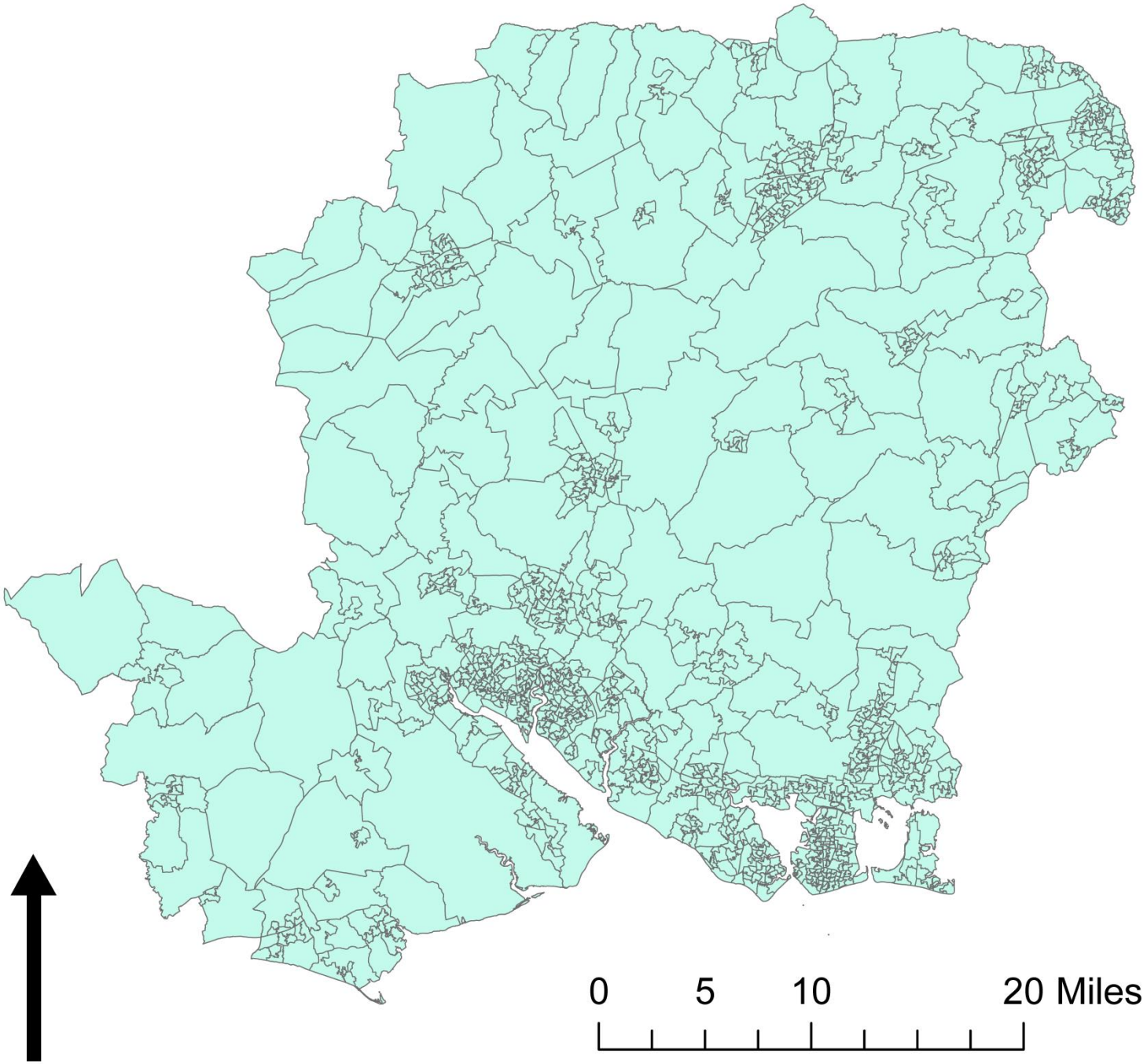
Statistical Disclosure Control challenge

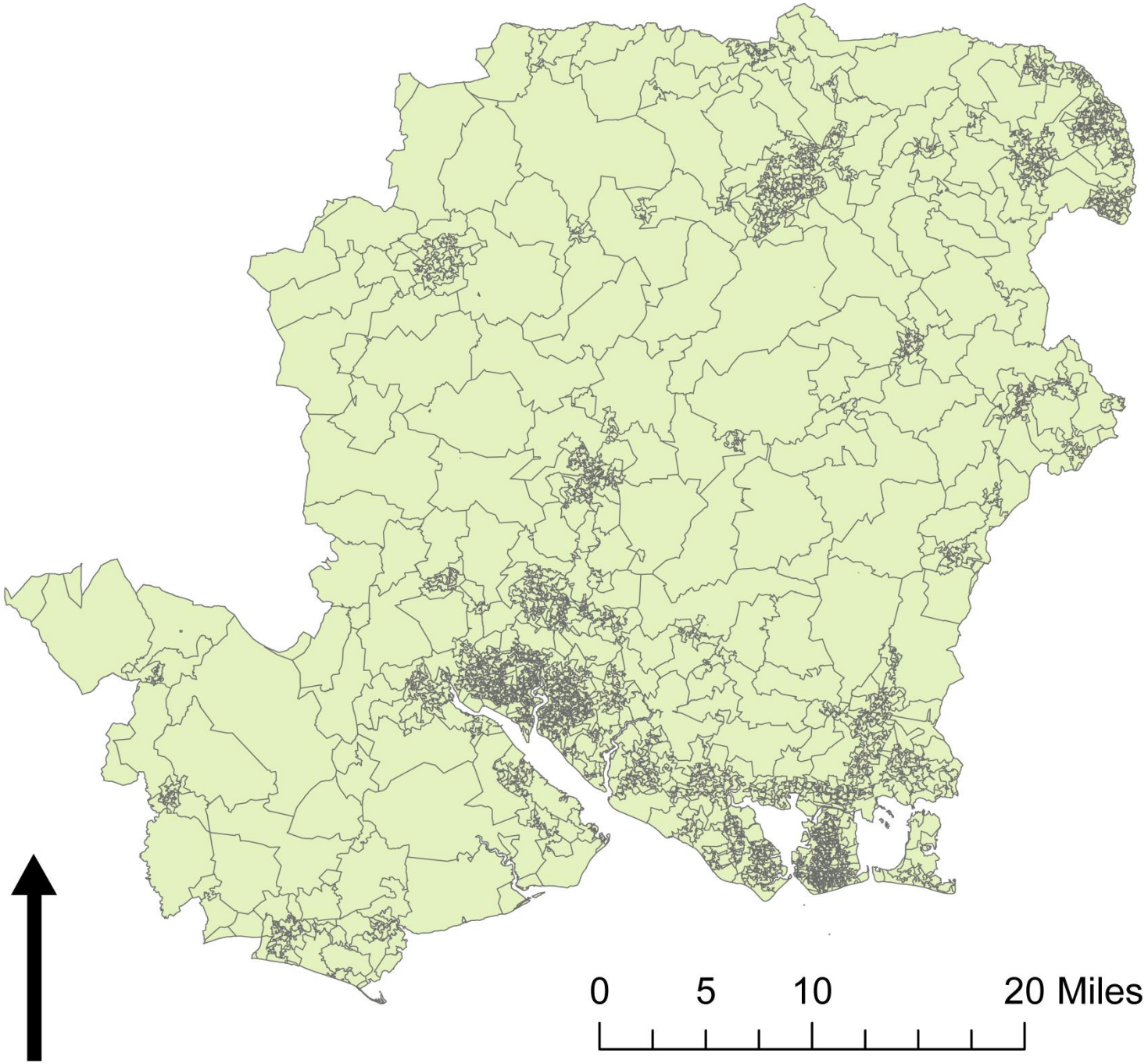
- Individual records are potentially identifiable
- Locational information is an important key to disclosure.
- Standard approach is to aggregate over geographical areas to meet a required level of comfort for data release – for aggregated data (census output areas).
- Additional protection from record swapping, collapsing classes, minimum thresholds
- More admin data-based research means more consideration of these issues



* Data disclosure design decisions (geographical codes, variable coding, cell suppression, aggregation, other measures…)


0
5
10
20 Miles


0
5
10
20 Miles


0
5
10
20 Miles

NCRM National Centre for Research Methods

Conclusions

- Range of application problems
- AZTool available and being further developed
- Importance of understanding the implications of building blocks and constraints in determining characteristics of overall solutions
- Significant proportion of work usually involved in data setup and testing stages (as with so many other things!)

Useful resources

- NCRM online learning resource “Principles of automated zone design” https://www.ncrm.ac.uk/resources/online/automated_zone_design/
- AZTool program home page, including software download https://www.geodata.soton.ac.uk/software/AZTool/

NCRM National Centre for Research Methods

Any questions?

Acknowledgements

- David Martin and James Robards are supported by NCRM, ESRC Award ES/L008351/1
- David Martin and Chris Gale are supported by ADRC-E, ESRC Award ES/L007517/1

- AZTool http://www.geodata.soton.ac.uk/software/AZTool/
- The data for this research have been provided by the Consumer Data Research Centre, an ESRC Data Investment, under project ID CDRC 025, ES/L011840/1; ES/L011891/1.
- Murdock, A.P., Harfoot, A.J.P., Martin, D., Cockings, S. and Hill, C. (2015) OpenPopGrid: an open gridded population dataset for England and Wales. GeoData, University of Southampton.
- Nomis data source: Office for National Statistics.