



UCL

WORKING PAPERS SERIES

Paper 148 - May 09

**3D cities and numerical
weather prediction models:
An overview of the methods
used in the LUCID project**

ISSN 1467-1298



CASA

3D cities and numerical weather prediction models: An overview of the methods used in the LUCID project

Stephen Evans
UCL Centre for Advanced Spatial Analysis,
1-19 Torrington Place, London WC1E 7HB
7th May 2009

Background

Transferring spatial data between different types of spatial models is often a much trickier process than GIS professionals would like to admit. This is further complicated if the models were generated for very different purposes and at different levels of spatial granularity and using different spatial projections. This is the case when you try to couple Ordnance Survey Mastermap data (which is at a 1:1,250 scale) with a numerical weather prediction model (NWP), (which until this project was based upon a grid of around 4km per grid cell). To achieve this, some new methods were developed to generate data for a localised NWP model at a 1000m and a 250m grid for the UK Meteorological Office as part of the LUCID project (the development of a Local Urban Climate Model and its Application to the Intelligent Development of Cities). These methods generated detailed urban morphology data to act as input data for the climate models, using Geographical Information System (GIS). The results form part of a chain of data that contribute to the main LUCID target: *to develop world leading methods for calculating local temperature and air quality in the urban environment*, in particular with reference to the heat island effect in large urban areas (in this case London).

LUCID, Meteorological modes, 3D urban data and spatial analysis

This is not the first time that CASA has been involved in coupling 3D models and data about the environment of London. Previous work (in collaboration with the Environmental Research Group at Kings College London) developed a visualisation system to view air quality in 3D alongside the CASA Virtual London model (for example see figure 1 below and the live site at: <http://www.londonair.org.uk/london/asp/virtualmaps.asp>). However, both models were independent. The air quality values were calculated independently using different data sources and so the 3D model was not coupled to the air quality model. In effect it was merely a visualisation of two models that occupied the same spatial location; one a model of the buildings, bridges, roads and parks of London, and the other a model of the air quality within that environment.

The LUCID project takes this work a step further, since it looks to model the interaction between the urban environment (buildings, roads, pavements, open spaces etc) with the various environmental models (numerical weather prediction models (NWP), air quality models and the like). The urban 3D model is no longer set to be a 'tailor's dummy' upon which to 'hang' another dataset, but has become *part of the environmental data model* itself.

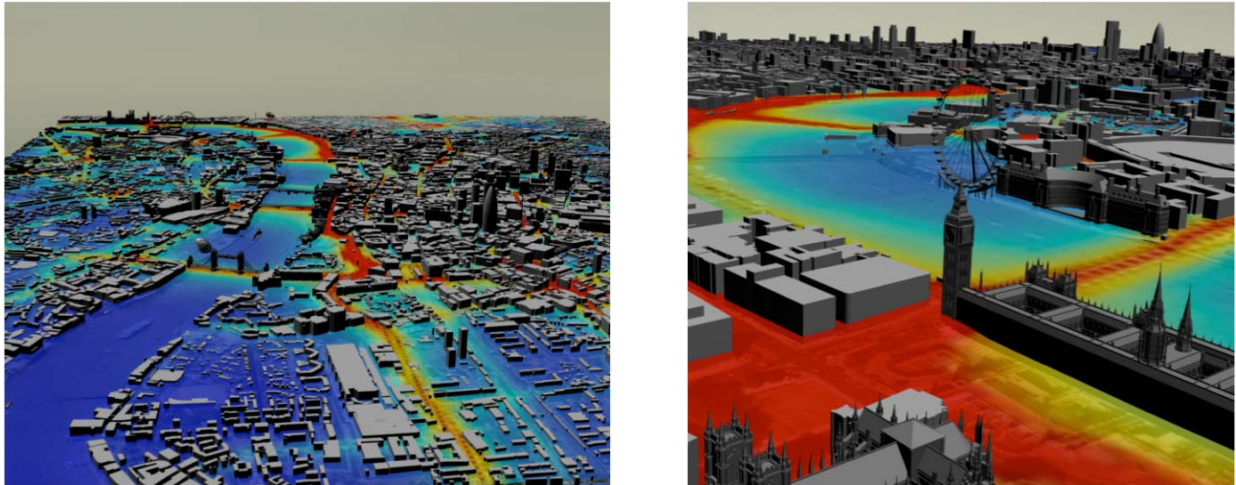


Fig. 1 – Virtual London and air quality – some visualisations from the author’s 2006 work. Nitrogen oxide (NOx) air pollution (annual average predicted for 2005) simulated as a layer within the model. This pollutant, often nicknamed ‘urban smog’ is largely derived from vehicle emissions. Red shows higher levels, whilst blue represents the lower levels (NOx data courtesy of Environmental Research Group, Kings College London and GLA).

At present, the LUCID project is just beyond the midway point. The project aims to calculate the impact on energy use and the consequences for health of a changing climate in a large urban area, and what the implications for future urban planning might be.

There are 3 core objectives to LUCID:

- To develop a new integrated tool to model the local climate in urban areas;
- To use the model to explore the complex relationships between the projected changes to regional climate and local urban climate and the impact on energy used in buildings;
- To evaluate the impacts of local temperature and air quality on health as the result of a changing climate.

In all of these objectives the critical zone for both modelling and measuring the impact of the climate model is the *boundary* where the NWP model meets the urban model. Modelling how the two interact is what we are most interested in (see Figure 2). In order to achieve this, the detail of the urban area is required and this is provided by the 3D model of London combined with details about land-use.

Urban areas have many ‘roughness’ elements such as buildings, trees, masts, street furniture and vehicles. As the wind blows over urban areas, the air interacts with these roughness elements and responds according to their size, shape, layout and distribution. Meteorologists and pollution dispersion modellers often use quantified parameters that describe the size and distribution of these roughness elements. Figure 3 shows the parameters of importance, their definition and an illustration of how they are determined.

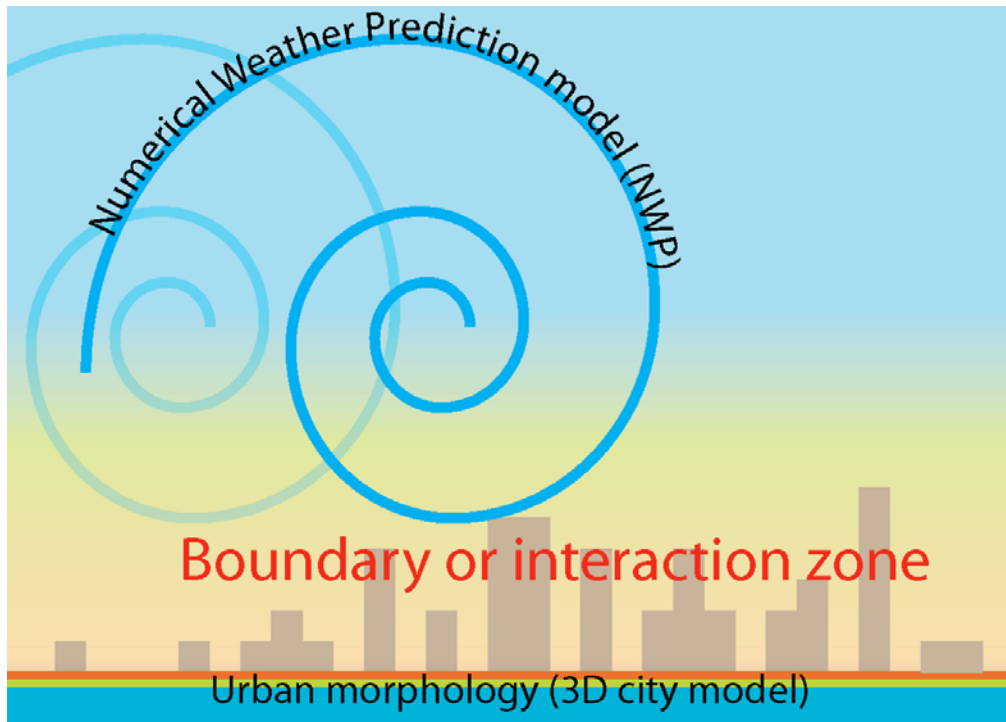


Fig. 2 – The interaction zone between a Numerical Weather Prediction model and the urban model

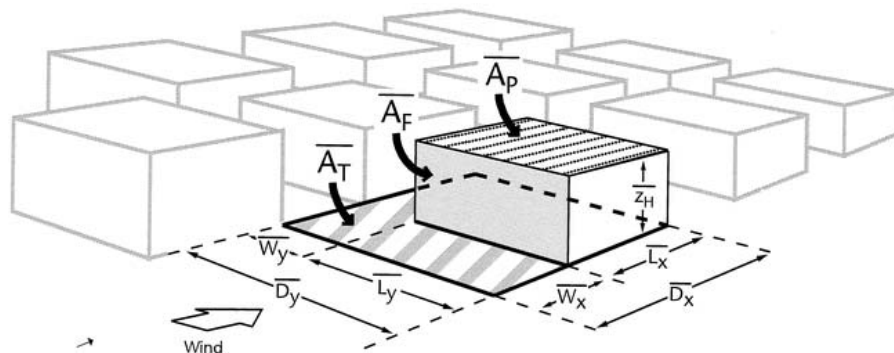


Fig. 3 - Definition of surface dimensions used in morphometric analysis.

The element portrayed has the characteristic mean dimensions, spacing, and total lot area (A_T) of the urban array. Using these measurements, the following nondimensional ratios are defined to characterize the morphometry:

$$\lambda_P = A_P/A_T = L_x L_y / D_x D_y; \lambda_F = A_F/A_T = z H L_y / D_x D_y; \lambda_S = z H / W_x = z H / (D_x - L_x), \text{ and} \\ \lambda_C = [L_x L_y + 2(L_y z H) + 2(L_x z H)] / D_x D_y.$$

Although drawn as building-like, the element is generic, representing all obstacles affecting the airflow. Similarly, the concept is not limited to a grid array. It could include scattered trees, differently shaped houses, and winding streets that are more typical of real cities. [Figure and caption reproduced from Grimmond and Oke (1999)].

At first this seems a representative view of an urban area and something that a 3D model could quickly produce. The A_P data is certainly straightforward to work out from a GIS simply by calculating the total area of all the building footprints within the defined area. Calculating A_F looks

slightly more complex, but still relatively simple in Figure 3. However, under closer scrutiny, it is clear that Figure 3 provides an over-simplified view of an urban area. In reality things are much more complicated, with individual buildings having courtyards, multiple roofs and so on, as can be seen in figure 4 below.



Fig. 4 - An example of A_F (blue) and A_P (orange) illustrated for a specific wind direction on some London buildings, showing how an urban area can be much more complex to work with in reality.

Further still, the building and the exposed wall areas (A_F) are rarely aligned with the wind direction. This means that for each building a more complex calculation needs to take place to extract the 'projected' wall area that is exposed at right angles to the wind direction, as shown in Figure 5.

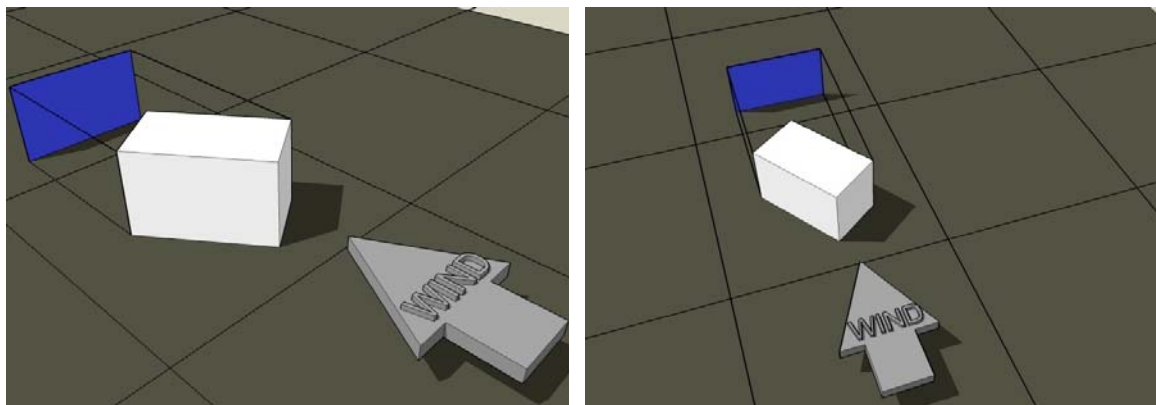


Fig. 5 - An example of A_F (blue) projected to calculate the exposed area for a building (white) that is not aligned to the wind direction.

Climate and meteorological models are normally grid-based (using a regular grid with each cell representing a value, such as temperature, air pressure and so on), whereas most city models are usually vector based (constructed of lines, points and polygons). Whilst it is relatively simple to move between the two, it can also be detrimental to the data, with generalisations occurring and fine scale data being 'lost' in the process. For this reason, it is not always advisable to convert data from its 'native' format prior to any major processing or calculations. In order to maintain data quality it became clear that calculations should be carried out on the 3D model in its native vector format.

To add to this already complicated process, the meteorological models usually work in a very specific projection (the Unified model used by LUCID uses a rotated latitude and longitude coordinate system

in which the computational North Pole is shifted to a position of 37.5 deg. N, 177.5 E. This is done to yield a uniform horizontal grid resolution above London.). National spatial data is often held in a specific projection system (for Great Britain it is a transverse Mercator projection, with a central meridian at longitude 2 degrees west). Re-projecting data is fairly straightforward, but can take up large amounts of processing time. Since a large part of LUCID was set to work with spatial data in the transverse Mercator projection, we chose to keep the London data in its native projection so that any data changes or new data sets could be added and edited without re-projection issues. The grid used by the meteorological model could be projected into transverse Mercator to provide the structure and gateway for transferring data across to the meteorological model. However, a re-projected grid is almost always distorted and the result is that the grid that is north-south aligned in one projection will appear to be skewed when re-projected into the other projection (see figure 6). This has serious implications since it rules out using raster modelling techniques, since a raster grid has to be aligned vertically and horizontally with the data in the projection in which it is used. The only way to generate the ‘raster’ outputs from the vector data in OSGB coordinates (suitable for the NWP model) is to use vector techniques, i.e. develop a ‘vector’ grid lattice, and use this to identify and select buildings and objects that fall within each lattice ‘tile’.

	NWP model	Urban morphology model
Scale / resolution	1km grid	1:1,250 scale data
Data type	Grid data (Raster)	Line, polygons (vector)
Projection	Met Office specific projection	Transverse Mercator

Table. 1 – A summary of some of the conflicts between the two models

It was initially envisaged that the 3D urban data for London would come from the CASA model that was originally named ‘Virtual London’. This model was first created in 2004 and showcased at the ESRI International User Conference in 2004 (to help launch the ESRI ArcGlobe product). Details about the construction of the model can be found in [Evans, Hudson-Smith and Batty \(2005\)](#). In 2006 the model was expanded to include all London Boroughs. In essence, it is a simple coupling of two commercial datasets: Ordnance Survey Mastermap data (OSMM), and **Light Detection and Ranging** (LiDAR) height data. This model used LiDAR height data to calculate average heights of the building polygons supplied by the OSMM topographic layer. To enhance the visual appeal, key London landmarks and streets were replaced with more detailed models created using CAD packages (for example the British Museum, Houses of Parliament...). Some of the techniques are described in Batty et al (2007).

However, as a block model, relatively little spatial analysis has ever been carried out with it. It is attractive, but dumb. As such it is often coupled with other data sets (see Figure 1), but only as ‘tailor’s dummy’, and rarely as part of the analysis itself. This is partly because it is stored and managed in a format that lacks topology.

Topology is a mathematical procedure for explicitly defining spatial relationships. The principle in practice is simple. Topology expresses different types of spatial relationships as lists of features (e.g. an area is defined by the arcs comprising its border). The ability to create and store topological relationships has a number of advantages. Topology stores data more efficiently. This allows processing of larger data sets and faster processing (ESRI, 1995). When topological relationships

exist, you can perform analyses such as modelling flow through connecting lines in a network or analysing the polygons that sit either side of an arc. This is particularly useful in GIS because many spatial modelling operations don't require coordinates, only topological information.

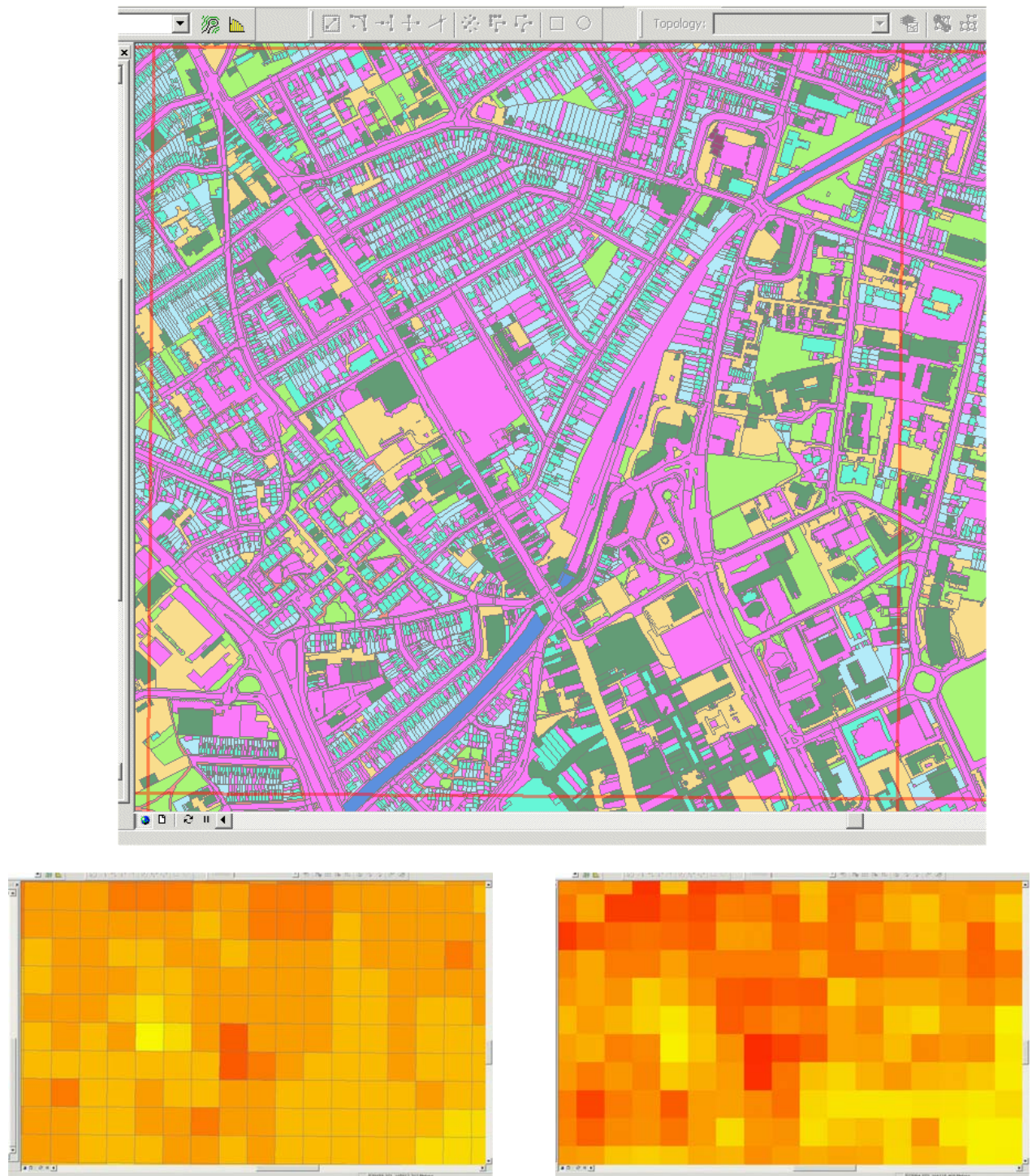


Fig. 6 – Some details showing the NWP model grid (shown in red in top image) alongside OS Mastermap data. Close inspection should show how the grid is slightly distorted and not aligned north-south due to the re-projection from its original projection. The lower two images show some sample data. On the left, the vector grid is used, and the distortion should be apparent, whilst on the right, a raster grid is used, which aligns perfectly in north-south direction.

One of the downsides of creating topology is that it can be (computationally) time consuming to create and maintain, particularly for massive datasets such as is the case for the London data. When a dataset like this is largely used for visual purposes (creating maps or 3D visualisations), then topology is surplus to requirements and takes up unnecessary time to create and maintain. However, *if advanced spatial analysis is to be carried out then topology plays a fundamental role* and also ensures the quality of the GIS database is maintained.

With reference to LUCID, one of the downsides of storing Virtual London data without topology is that the polygons have no idea 'who' their neighbours are or whether they face the street or a courtyard. When this model is extruded into a block model some similar flaws can result. One of these is that all buildings are extruded from the ground, even if a building is multi-faceted. This can cause the creation of unnecessary (hidden) 3D objects (buildings within buildings) which can create a significant overhead on graphics cards during any 3D visualisations. More importantly this could cause the over-calculation of the surface area of walls within the model, and this, (as A_F) is one of the key inputs already noted in figure 3. This issue is illustrated in figure 7 below.

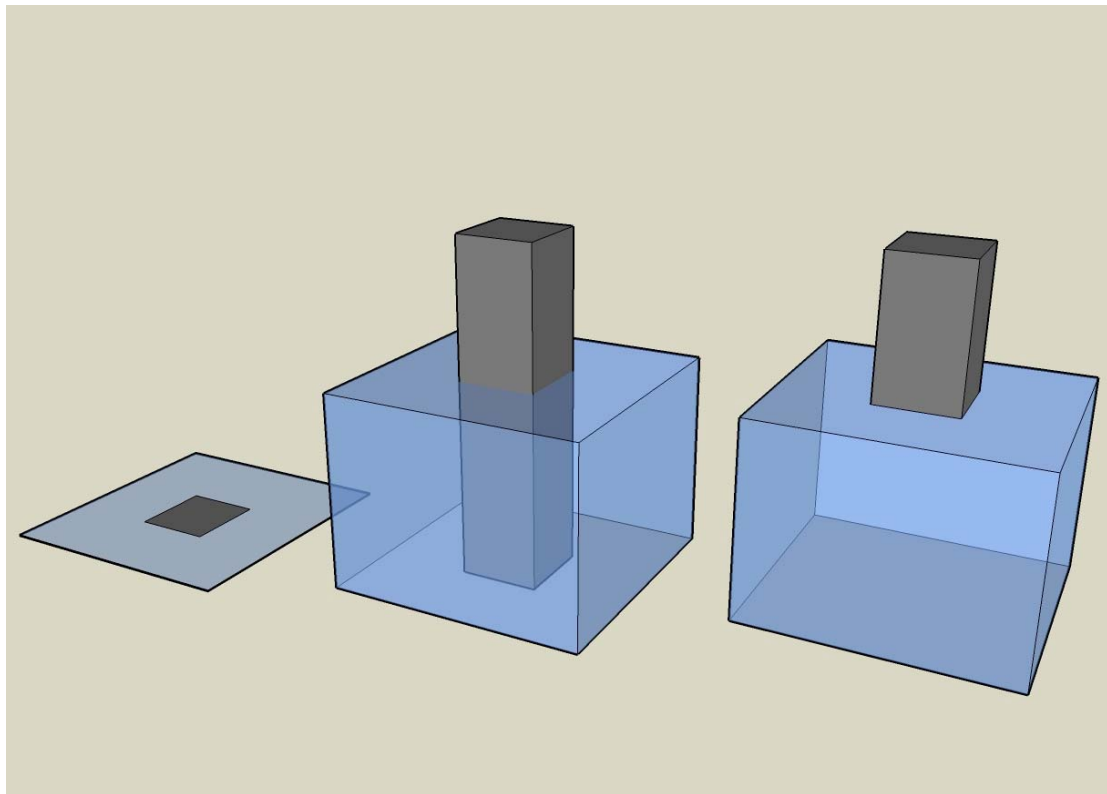


Fig. 7 – A simple example showing the different 3D extrusion methods for a multifaceted building. On the far left, two polygons are shown, representing the footprint of a building that has a main block and a tower above. To the right are two different extrusion methods (the outer polygon is made transparent for illustrative purposes). The middle example shows both polygons being extruded from the ground level (as is the case in Virtual London). The result is that too many 3D surfaces are created, and in the case of LUCID, an over-calculation of wall surface area could result. To the far right, the middle polygon has a base at the top of the main block, and as a result, a correct calculation of the wall surface area would be possible.

Creating topology for a huge dataset like this is time-consuming and it can be onerous to maintain the data in this state. However, some careful use of topology can address some of these issues raised above and this is discussed in more detail in the next section.

Overview of the methods used

It has already been mentioned that the Virtual London model, which was the starting point for this project, is simply based upon Ordnance Survey Mastermap (OSMM) data, combined with LiDAR data. Combining the two datasets in a GIS allows you to calculate height statistics for each building polygon based upon the LiDAR points that fall within the polygon. These statistics typically include average, maximum and minimum heights. The result is that building polygons can be extruded to a set height (for example average height of LiDAR points falling within that building polygon footprint) for visualisation purposes.

From here two grids were developed by the UK Meteorological Office that they would use in their NWP model. One was at approximately 1000m cell size and the other was at 250m cell size. Both were designed to be used in the specific projection used by the NWP model and hence were distorted when projected into any other projection. Each grid cell was assigned a unique code.

Next the OSMM was joined with the Generalised Land-Use Database (GLUD), developed by what was then known as the Office of the Deputy Prime Minister. This dataset attempts to classify all OSMM polygons into the following categories:

- Water
- Domestic buildings
- Non-domestic buildings
- Roads
- Paths
- Rail
- Greenspace
- Domestic gardens
- Other (mainly hardstanding)
- Unclassified

Using these classifications allowed us to develop a selection query to classify A_T (see figure 3)

A_T

There was some discussion over what should classify as A_T and what should not. For LUCID we used the following GLUD classifications:

- Domestic buildings
- Non-domestic buildings
- Domestic gardens
- Roads
- Paths
- Rail
- Other (mainly hardstanding)

In addition all ‘greenspace’ polygons with an area of less than 7854m^2 were added to this selection. This represented those small ‘greenspace’ areas that are not classified as domestic gardens, but have the same effect, in that they should be classified as part of the urban area, since they are too small to have a significant impact on the local environment. The decision to set the area to 7854m^2 was made by the meteorological modelling team based upon the calculation of a circular area with a diameter of 100m ($\pi * (100/2) * (100/2) = 7854$)

Having made these selections, the boundaries between these polygons were ‘dissolved’ to remove polygon boundaries and to create a layer with A_T polygons.

The meteorological modelling team were then concerned that some of the areas classified as A_T were too small and isolated to contribute significantly to the meteorological model and that these should be removed. These were termed ‘patchy’ areas, and included, for example, small clusters of urbanised areas set within an area of open space. This would include, for example, a small cluster of housing surrounded by fields, or a small warehouse surrounded by open space. Based upon this we made a subsequent selection of the dissolved ‘ A_T ’ polygons to find those polygons that were smaller than 7854m^2 and these were removed from A_T .

This layer was then combined with the vector grid, and from here the total area of A_T per grid cell was calculated.

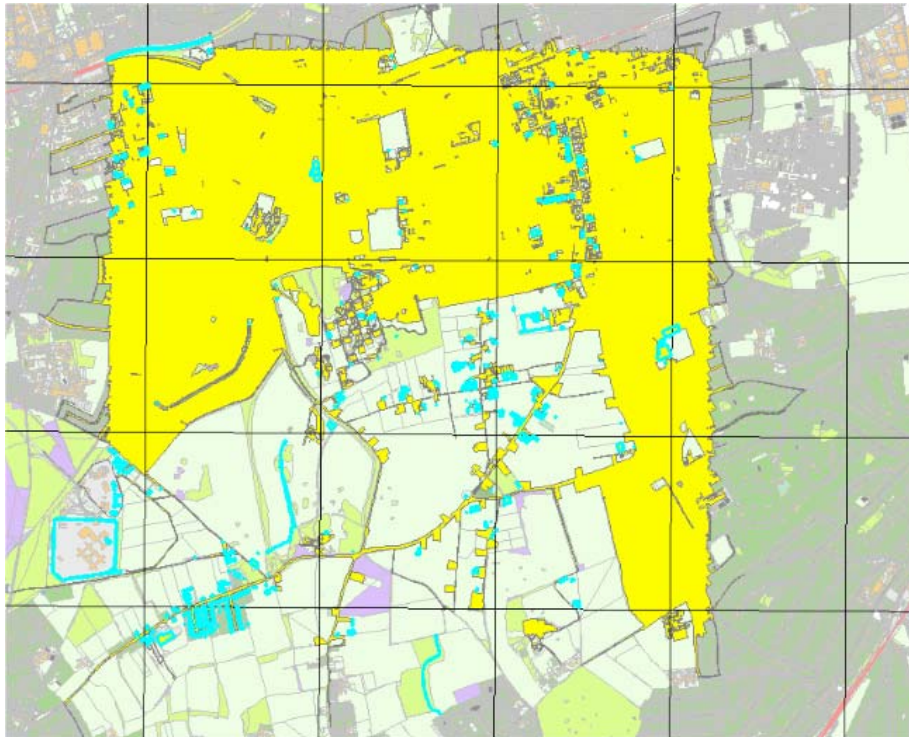


Fig. 8 – A_T is shown in bright yellow for an area on the outskirts of Greater London, with clusters of housing as well as fields and large areas of open space (shown in pale greens and purple). A 1km grid is overlaid on the data. A_T (urbanised) is shown in yellow, whilst some of this is highlighted in light blue. These are A_T polygons that are smaller than the minimum level (7854m^2) and are subsequently removed from A_T .

A_P

Once A_T had been calculated, it made it relatively simple to calculate A_P (see figure 3). A_P is the sum of the footprints of the buildings that are spatially contained by A_T , calculated per vector grid cell. All that is required is to select all OSMM polygons classified as ‘buildings’ and then to filter this selection with a spatial selection using A_T . Again the output was generated for each vector grid cell.

A_F

Calculating A_F is a more complex process. Since this is in effect the wind resistance of the built environment, it is first important to agree a wind direction. For the purpose of LUCID it was agreed to test a southerly wind (180°) and a westerly wind direction (270°). We then calculated, after several experiments, that for a building that is represented as an extruded, closed polygon, the A_F value for that building for a particular wind direction, is identical to the A_F value for the opposite direction, no matter what the shape of the building, provided you consider that no walls of the building shelter any other walls from the wind.

The next step was to generate topology for the OSMM buildings, both for the polygon coverage and the arc (line) coverage. The polygon coverage by this stage had the height statistics calculated as an attribute for each polygon, which could be identified by a unique identifier known as a TOID (TOpographic IDentifier). One of the by-products of generating topology is that the lines (arcs) that make up a polygon are given two attributes (LPOLY#, RPOLY#) which contain the unique code for the polygon that falls to the left, and to the right of that arc. Using this link a join can be made back to the polygon layer, the polygon TOID, and hence for each arc it is possible to calculate a height (from the polygon) for the building immediately to the left of the arc and immediately to the right of the arc, if one exists (see figure 9). Since the original data was filtered to include only building polygons (everything else was deleted), many of either the LPOLY# or RPOLY#'s of the arcs are a reference to the Universe polygon (empty space that continues potentially to infinity). Those walls that have either their RPOLY# or LPOLY# as the Universe polygon we know to be walls that face out into open space (i.e. the street, a garden, field etc). These walls can be assigned the height of the building to which they are joined. Other walls will reference a building both through their LPOLY# and RPOLY# joins. In these cases we know that they are to a large extent ‘party walls’, for example dividing a row of terraced houses, or separating a semi-detached property. However, as is often the case, one of the buildings may be taller than the other, in which case this difference is calculated as the exposed section of the party wall. It is simply calculated as the difference between the height of the lower building and the taller building that share the wall. A third case is that either the LPOLY# or RPOLY# references neither the Universe polygon nor another building. In these cases, these are walls facing into a courtyard area, and in a similar way to the Universe polygon facing wall, it can be assigned the full height of the building. As part of this process each wall was assigned a parent polygon (building). In the case of the ‘party walls’ this was assigned as the taller of the two buildings between which it was shared. The result was that a one-to-one relationship could be generated whereby each wall was set to be the ‘child’ of one ‘parent’ building polygon.

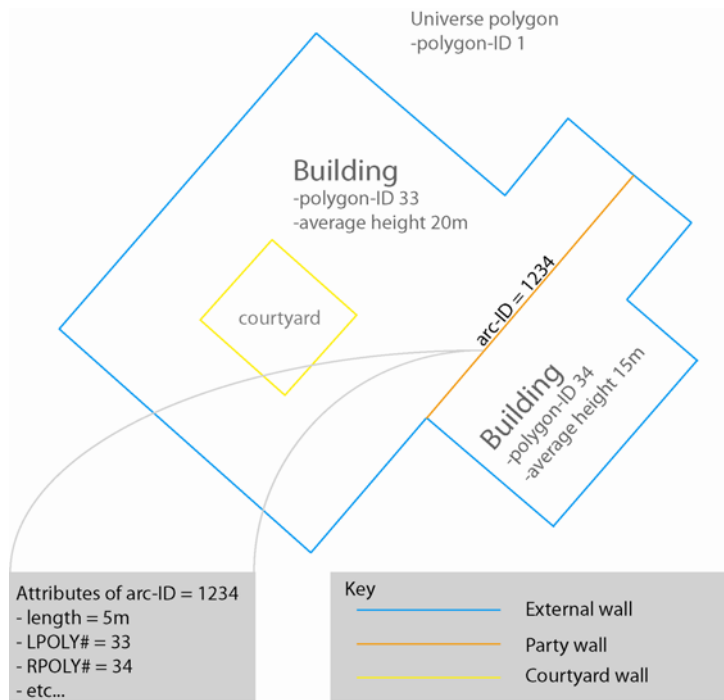


Fig. 9 – Once topology is created, a relationship can be made between arcs (walls) and polygons (buildings). This allows the classification of walls into various categories and in the case of party walls that separate two buildings of different heights, it can be used to calculate the exposed height of the party wall. In the example above, the wall with arc-ID = 1234 can be seen to relate to the building-ID 33 and building-ID 34. Using this link, it can be calculated that the exposed height is 20m – 15m, which leaves 5m of exposed wall. Courtyard facing walls and Universe polygon facing walls are assigned the full height of their parent building.

The next step was to break all the arcs into straight line sections and then calculate the orientation of these straight sections. From here it was possible to use basic trigonometry to calculate the area of wall that was exposed to the specified wind direction (see figure 10).

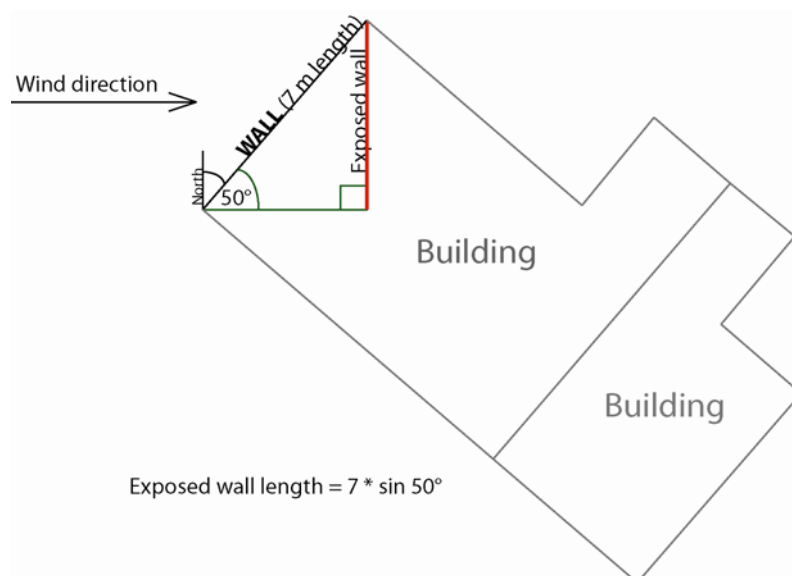


Fig. 10 – The length (and area) of each specific wall is calculated in terms of the exposure to a particular wind direction. This is calculated by working out the length of wall that is exposed when

viewed from the wind direction (illustrated in red). This can be calculated using basic trigonometry as illustrated.

Finally, using A_T as a spatial selection, it was then possible to select all walls that fall within a grid cell and within A_T and calculate A_F for the specified wind direction.

Results

The final results from the LUCID project are still in progress so it is too early to draw any detailed conclusions. Early results have shown that the coupling of the urban model and the numerical weather prediction model has worked well, and that some detailed meteorological effects can be seen to take place in the simulations, as a direct result of the finer scale urban data. However, the details of these effects need to be tested thoroughly before we can publish any results. It has been clear from all the tests so far that including the urban model in the NWP model does make a significant difference to the temperatures produced by the NWP model. The figures below show some of the outputs from the A_T , A_F , A_P and land-use data, and an illustration of the effects of these in the NWP model.

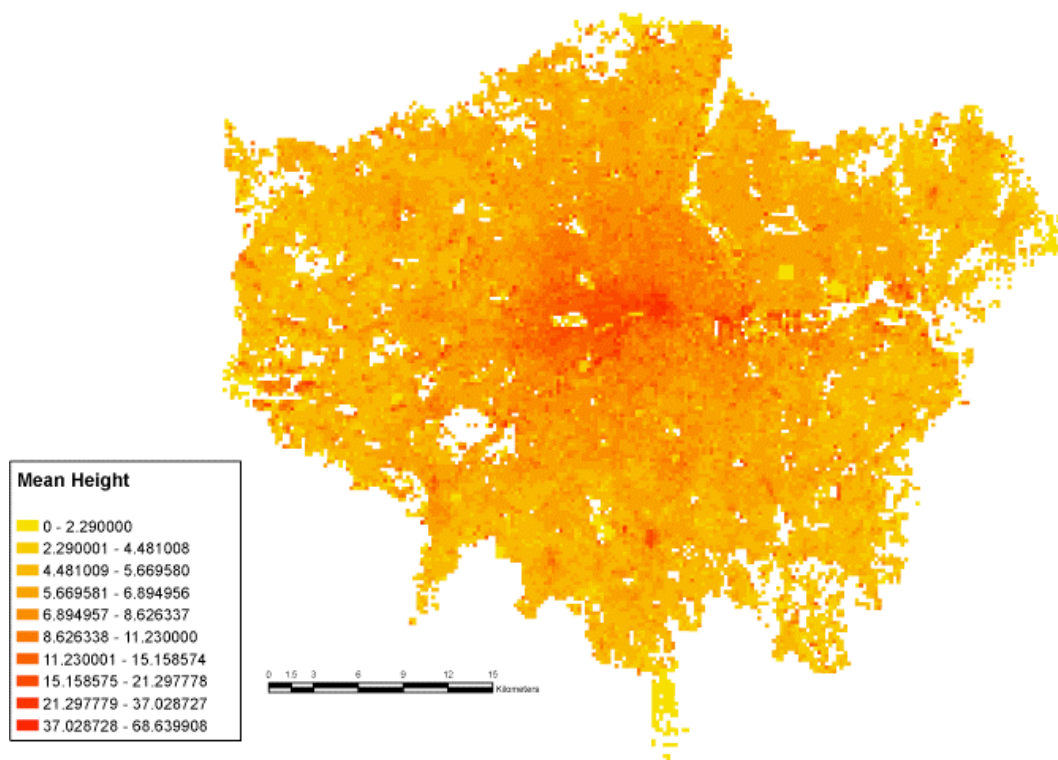


Fig. 10 – Average height (m above local ground level) of buildings, per 250m grid cell for London

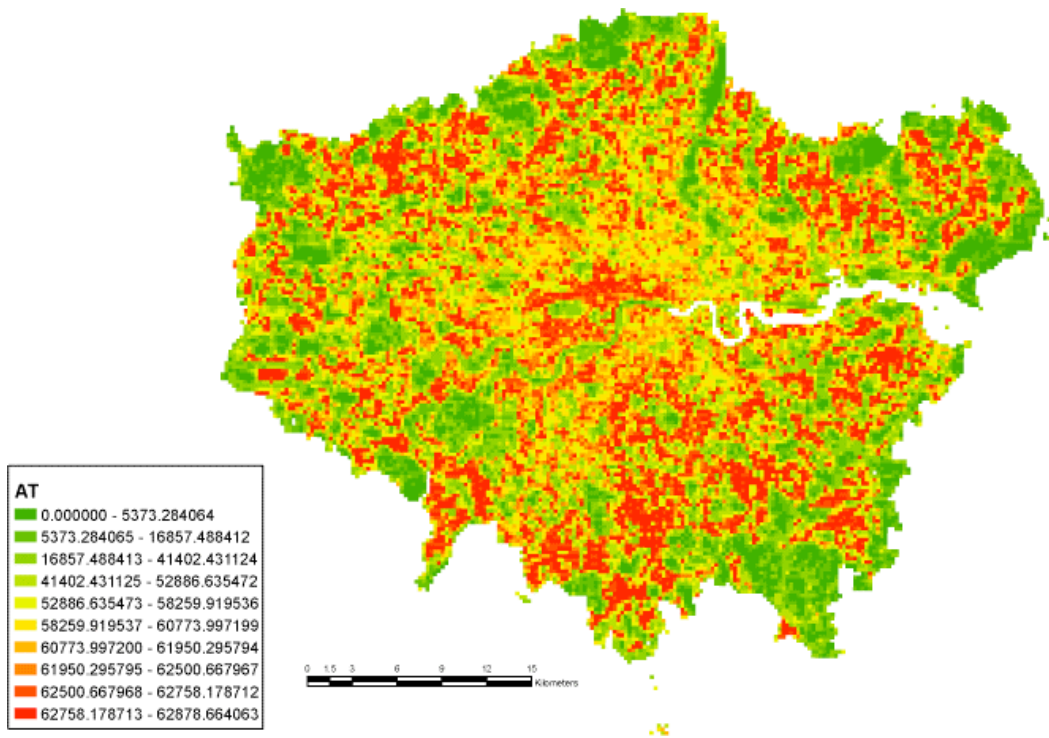


Fig. 11 – A_T for London (m^2) (above).

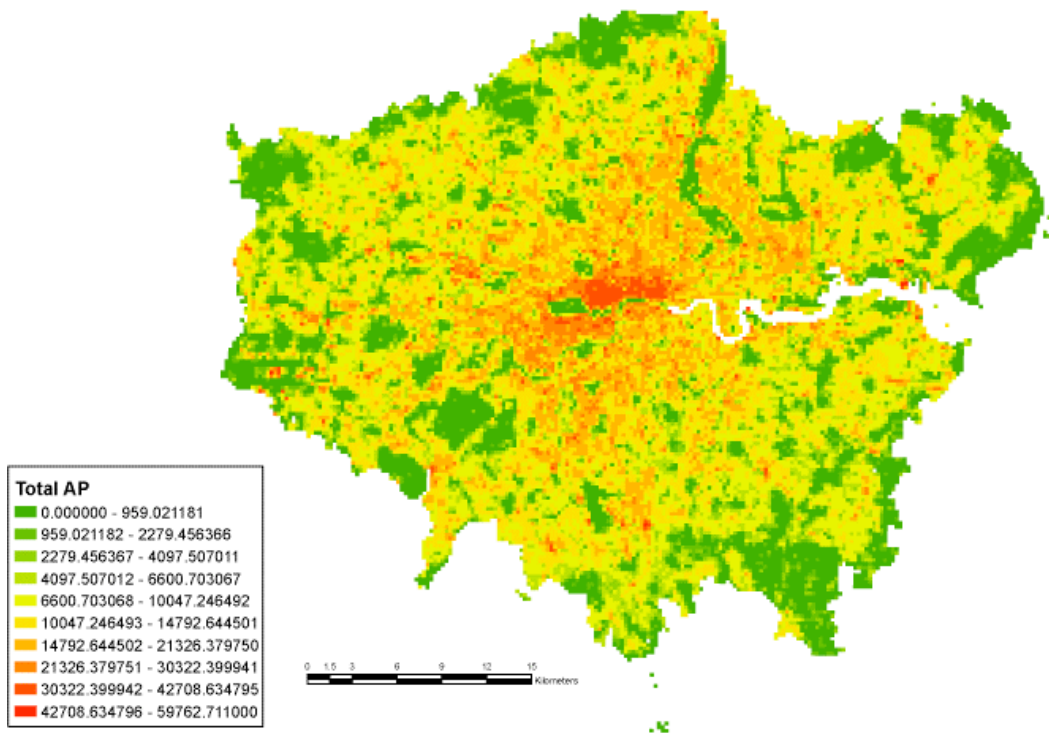


Fig. 12 – A_P for London (m^2) (above)

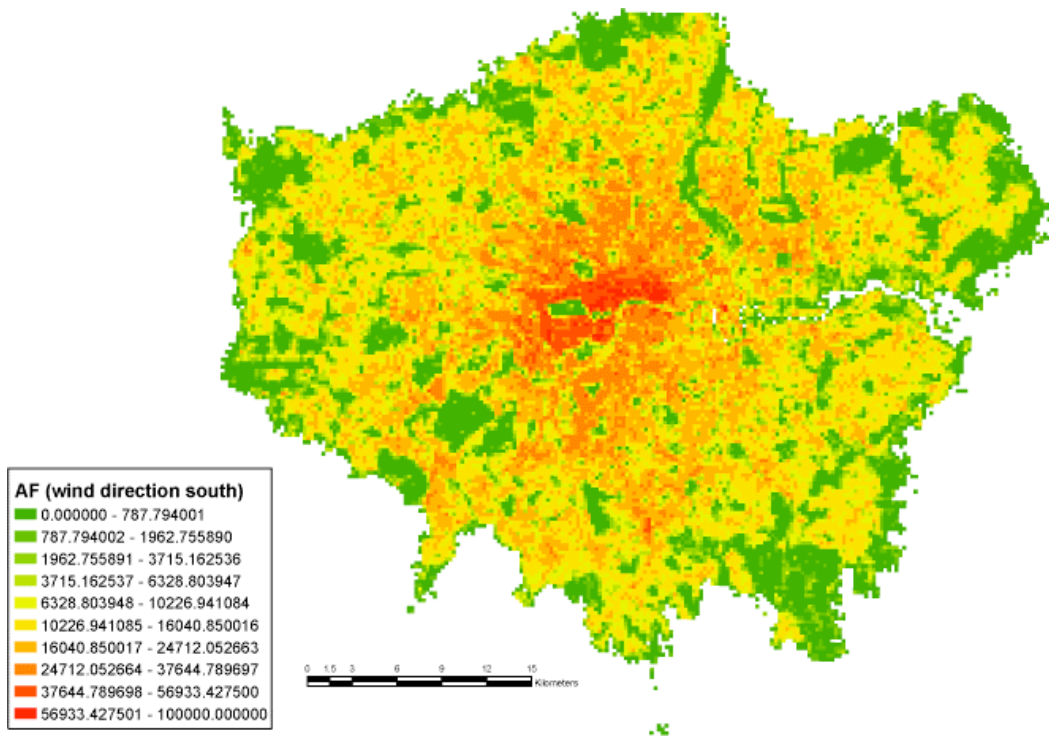


Fig. 13 – A_F for London (m^2) for a southerly wind direction (above)

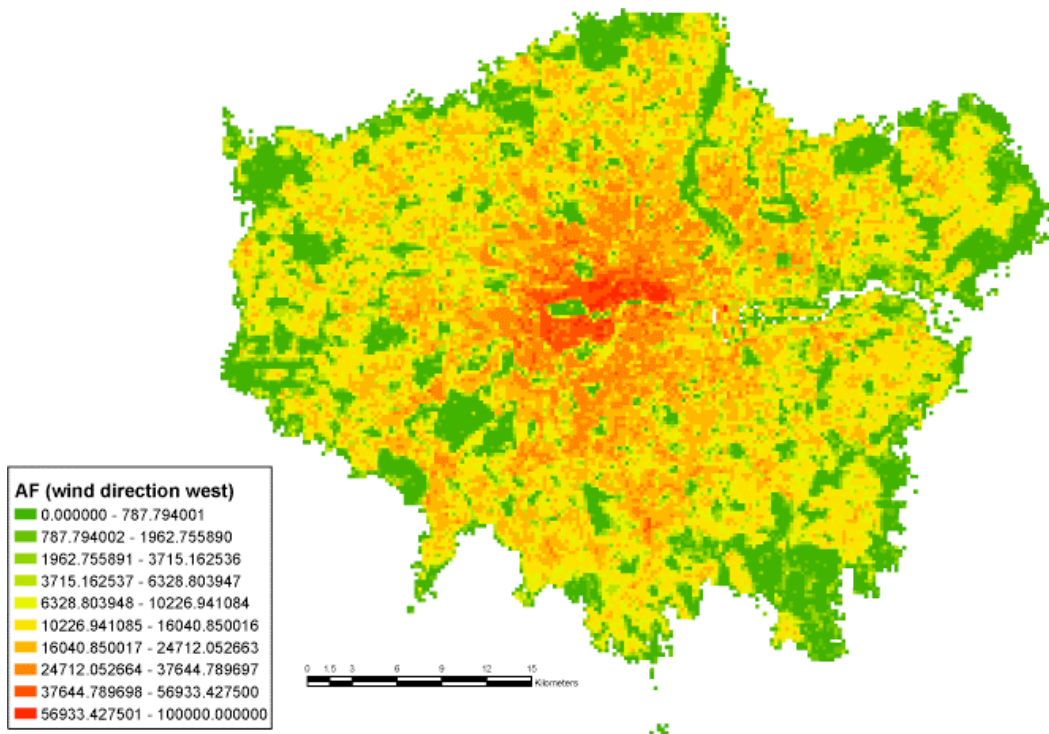


Fig. 14 – A_F for London (m^2) for a westerly wind direction (above)

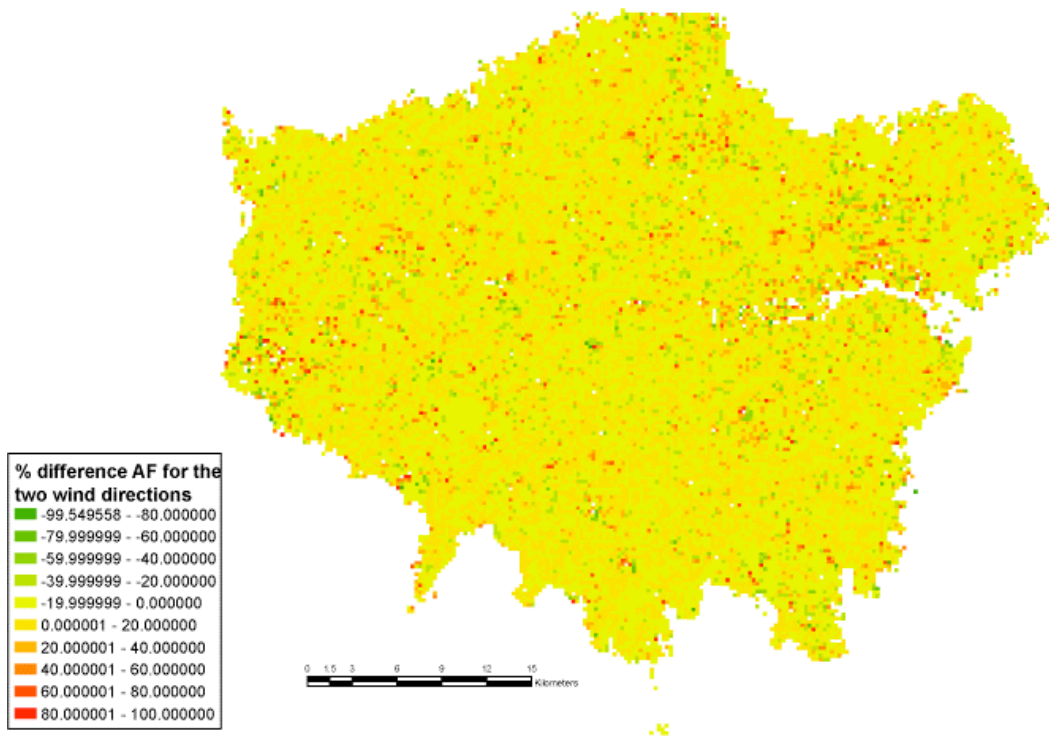
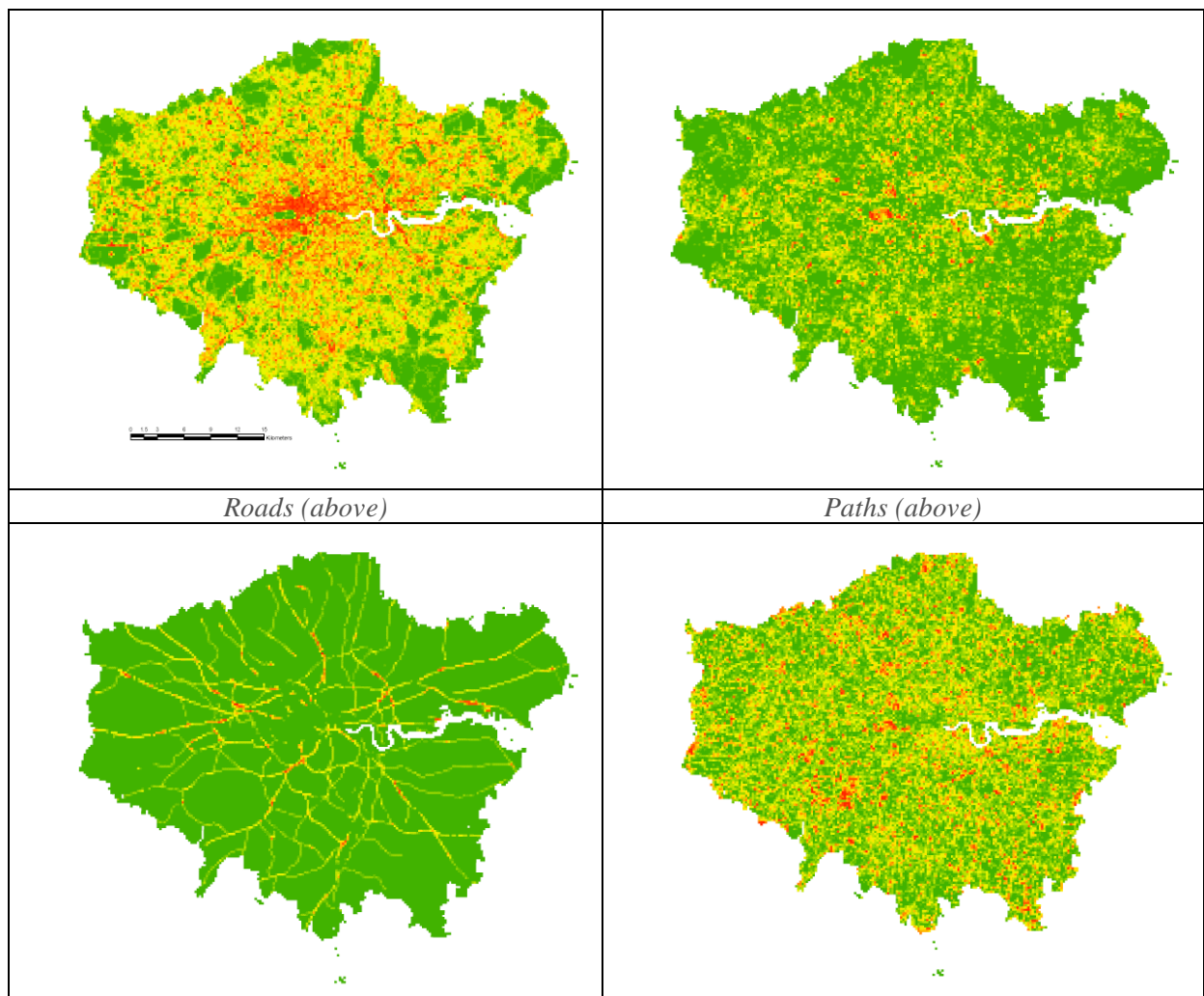


Fig. 15 –% difference between A_F (westerly) and A_F (southerly) for London (m^2)



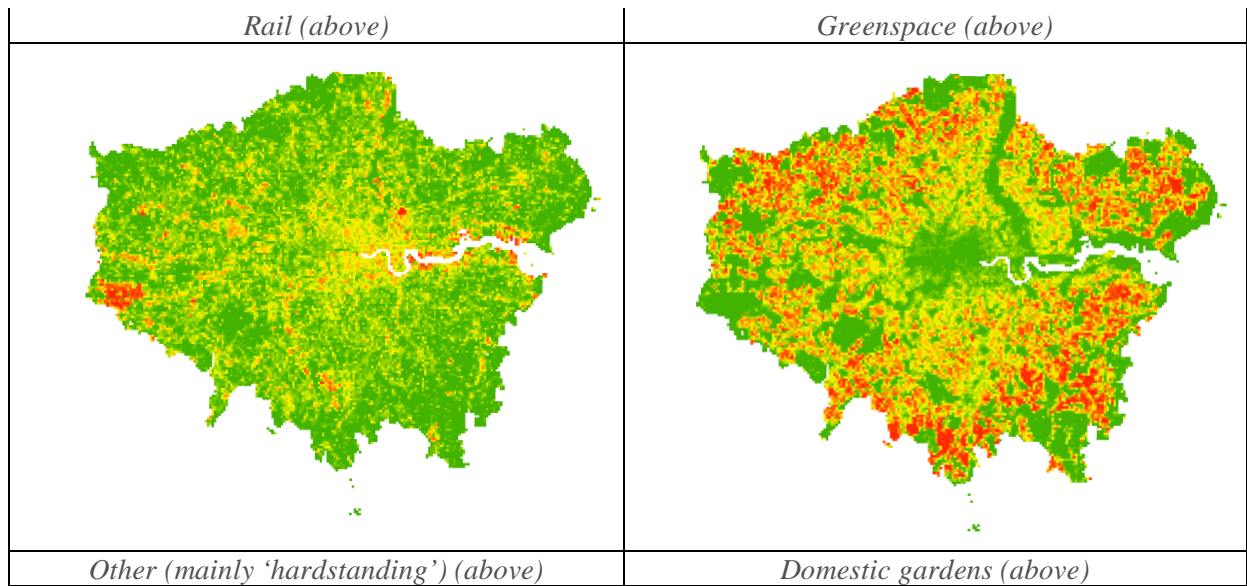


Fig. 16 – Total area (m^2) of the land-use contained by A_T (green values are low, red values are high)

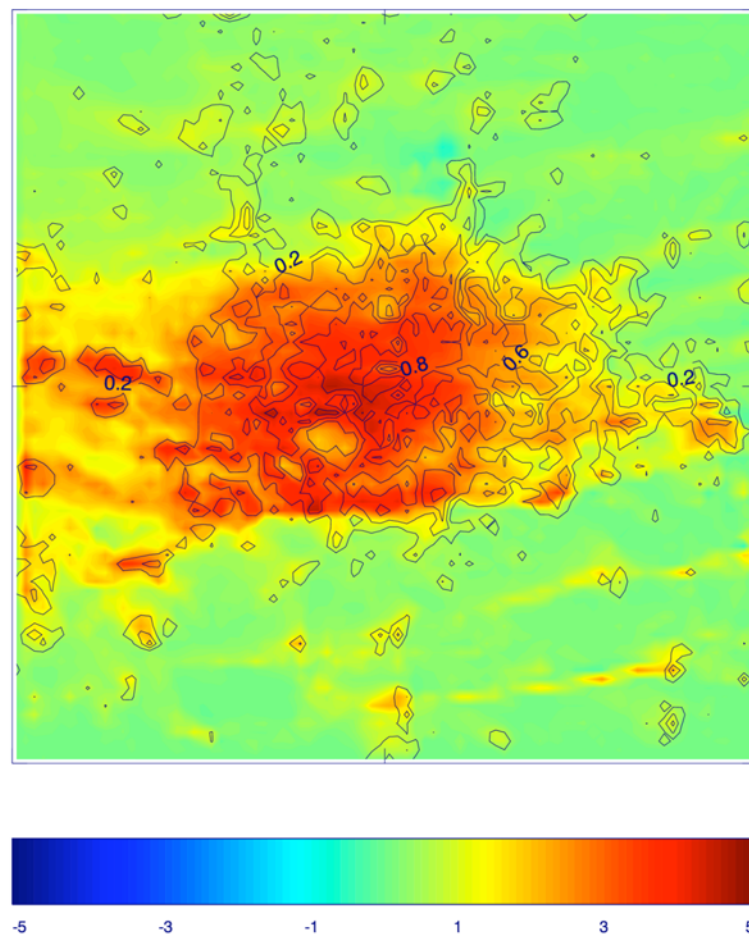


Fig. 17 – A plot of London from the NWP model, showing temperature differences between a model run using standard Met Office data and a model run using the urban morphology data (AF, AP and AT as described above).

The view is approximately 80x80km. The model run was 'fed' by recorded historical data which provides the boundary conditions. This was data from 8th May 2008. The time of this screenshot is 23:00 which not only shows the effect that the urban morphology can have upon the weather prediction, but also neatly illustrates the urban heat island effect, which is often most pronounced at night. Isolines depict urban land-use fractions per grid box. The horizontal resolution is 1km and boundary values were taken from a 4 km simulation.

Conclusion

The author has worked with 3D GIS for more than a decade now, but in the vast majority of cases, the 3D model is used purely for visual purposes. Often the most impressive models are the ones with the least versatility when it comes to carrying out spatial analysis on the model geometry. Despite this, we are seeing more and more cases where it makes sense to 'couple' different types of spatial models in order to tackle environmental issues. This work has shown how a 3D model can be taken and developed to create input data for a numerical weather prediction model, in order to model the interaction between meteorology and the urban environment. The methods are not always straightforward, and some techniques have not been tried before at these spatial scales. This is a challenge to all of us who build and work with 3D city models – they can't just be 'dumb and pretty', but must be developed to allow detailed spatial analysis to be carried out to the highest level of accuracy possible, and to be easily integrated into future modelling of the urban environment. Only then will the true value of these 3D urban models be fully realised. Hopefully the LUCID work to date has shown that this can be done, and may encourage further applications of 3D GIS to assist planners of the built environment, meteorologists, climatologists, epidemiologists, economists and others in their efforts to predict and adapt to future changes in the climate.

Acknowledgements

The author would like to thank the LUCID project for funding this work and to thank Phil Steadman, Mike Davies, Mike Batty for their inputs. Data from the numerical weather prediction model appears thanks to Sylvia Bohnenstengel and the UK Meteorological Office. All Ordnance Survey data in the images is © Crown Copyright/database right 2009. An Ordnance Survey/EDINA supplied service.

References

Batty, M., Carvalho, R., Hudson-Smith, A., Milton, R., Smith, D., Steadman, P., (2007), "Scaling and Allometry in the Building Geometries of Greater London". CASA working paper 126 (http://www.casa.ucl.ac.uk/working_papers/paper126.pdf)

ESRI (1995), Understanding GIS, The Arc/Info Method sections 2-9 to 2-10

Evans, S., Hudson-Smith, A., et Batty, M. (2005) « 3-D GIS : Virtual London and beyond », *Cybergeo*, Sélection des meilleurs articles de SAGEO 2005, article 359, mis en ligne le 27 octobre 2006, modifié le 04 juillet 2007. URL : <http://www.cybergeo.eu/index2871.html>

Grimmond, C.S.B. and Oke, T.R., (1999), "Aerodynamic properties of urban areas derived from analysis of surface form." *Journal of Applied Meteorology*, **38** (9), 1262 – 1292.

Theobald, D., (2001) Understanding Topology and Shapefiles in ArcUser June 2001 URL: <http://www.esri.com/news/arcuser/0401/topo.html>