GEOGRAPHIC INFORMATION ISSUES ASSOCIATED WITH SOCIO-ECONOMIC MODELLING FROM NIGHT-TIME LIGHT REMOTE SENSING DATA

C. N. H. Doll [a,*], J. G. Morley [a], J-P. Muller [a]

[a] Dept. of Geomatic Engineering, University College London, Gower Street, London WC1E 6BT, United Kingdom.
(cdoll, jmorley, jpmuller) @ge.ucl.ac.uk

Commission VII, WG VII/6

KEY WORDS: Remote Sensing, GIS, Human Settlement, Urban, Mapping, Modeling, Economy

ABSTRACT:

The ability to model socio-economic parameters from remote sensing data sources is an important element in addressing the data gap that currently exists for linking human and physical systems in Earth system models. The provision of information about such parameters in a gridded format would be highly advantageous for incorporation with other environmental datasets within a Geographical Information System (GIS) environment. This paper outlines some of the geographical information issues encountered in a recent study to analyse and map the relationship between night-time radiance from the Defense Meteorological Satellite Program's Operational Linescan System (DMSP-OLS) and Gross Domestic Product (GDP) for countries in the European Union and the United States. Relationships between total radiance and GDP were constructed at a number of sub-national spatial scales. These relationships were then used to create maps of economic activity for these regions at an enhanced spatial resolution of 5km. The use of calibration statistics from administrative areas of different sizes gave rise to issues such as the Modifiable Area Unit Problem (MAUP) and the Ecological Fallacy. Correlation coefficients of these relationships were found to be highly variable as a result of the MAUP. Analysis of different economic sectors revealed that the overall relationship is a combination of different relationships for the different economic sectors present in any given administrative area. This paper not only intends to highlight how one may use remote sensing data to successfully create a map of economic activity compatible for use within a GIS, but also to demonstrate the complementarity between remote sensing and GIS at every level of analysis.

1. INTRODUCTION

Night-time light imagery can be intuitively perceived to be indicative of a range of socio-economic parameters. Socio-economic data are collected at irregular spatial units. Night-time light imagery potentially offers a means to disaggregate these data to a constant spatial resolution. Parameters such as human population, Gross Domestic Product (GDP), power consumption and even greenhouse gas emissions all have strong correlations with lit-area and/or total radiance (Elvidge et al., 1997; 1999; 2000; Doll et al., 2000). Previous work used country level aggregations of light and GDP to build a global relationship. This paper investigates the characteristics of sub-national relationships and discusses the implications for producing disaggregated maps of these parameters. The main body of results pertaining to both the relationships and maps are presented in an expanded paper submitted to Ecological Economics (Doll et al., 2004).

1.1 Issues surrounding multi-scale analysis of night-time light data.

The use of multi-scale data introduces issues, which ought to be taken into consideration when results are analysed. In particular the Modifiable Areal Unit Problem (MAUP) (Openshaw, 1984) and the ecological fallacy (Cao and Lam, 1997) are two complementary themes which frequently occur in GIS analysis, but are usually overlooked due to the lack of general robust solution. The MAUP consists of two separate but closely related components, which have been observed to cause a variation in statistical results depending on the spatial arrangement of zones or the scale used to analyse the data. These are known as scale and zoning effects. The scale effect occurs when different results are observed from the same data at different spatial resolutions (Wrigley *et al.*, 1996). This arises due to the aggregation of data into larger units (e.g. enumeration districts – wards – counties – regions). In addition to this, different results can also be produced where the scale of analysis remains constant but constituent areas are aggregated in a different way or zone boundaries are altered. These two issues are particularly pertinent in this analysis since the relationships derived from the data will be used as an input to the map production procedure.

A practical application of using the zoning effect can be seen as far back as the early 19th century, where it was observed that voting districts could be divided in such a way so as to concentrate the power of the ruling party. This occasionally gave rise to bizarrely shaped districts epitomised by that of a former governor of Massachusetts, Elbridge Gerry, whose party created a salamander shaped district in 1812. The term gerrymander was coined to describe the elected member from such a district as well as the action of altering district boundaries to gain an electoral advantage.

The MAUP has only been intermittently studied. It was identified by Gehlke and Biehl (1934 cited in Openshaw, 1984) and was again revisited in 1984 by Stan Openshaw.

* Corresponding author.

Whilst its effect has been well documented, it is also an issue that is frequently overlooked in spatial analysis, since its effect can take many forms. Openshaw (1984) remarked that 'the aggregational variability is not susceptible to a statistical approach since no systematic regularities could be found'. A corollary to the MAUP is the ecological fallacy.

In an analytical environment where results are known to vary depending on the scale or zoning structure, it follows that erroneous inferences can be made from data that are generalised from one scale to another. It has three manifestations depending on the direction of the scale change. The individualistic fallacy occurs when a macro-level relationship is imputed from a micro-level (individual) relationship. Cross-level fallacies describe those inferences made about one sub-population from another at the same scale of analysis. The final fallacy, the ecological fallacy is the opposite of the individualistic fallacy and describes those inferences made about the fine resolution from the coarse resolution (Cao and Lam, 1997). In this sense, the cross-level fallacy arises from the zoning effect, whilst the individual and ecological fallacies are derived form the scale effect. A multi-scale approach is anticipated to facilitate the identification of MAUP effects that exist with this type of analysis and what implications this has with respect to creating an accurate map.

2. DATA SOURCES

Satellite data from the Defense Meteorolical Satellite Programme's Operational Linescan System (DMSP-OLS) were used in conjunction with economic and energy consumption figures for sub-national administrative units.

2.1 DMSP-OLS Radiance Calibrated Data

DMSP-OLS data products are produced by the National Oceanic and Atmospheric Administration's National Geophysical Data Center (NOAA-NGDC) in Boulder, Colorado. Temporal composite imagery is used to build up a global map of night-time radiances over the Earth's surface. The radiance calibration is facilitated by using a variable gain setting on the OLS sensor in order to reduce the effects of saturation from very bright urban centres (Elvidge et al., 1999). In this respect, the radiance calibrated dataset is technically superior to the previous city lights dataset. The dataset refers to imagery captured between March 1996 and February 1997.

2.2 Ancillary Statistics

Co-temporal GDP and energy consumption data was obtained for the United States and 11 European Union (EU) countries. For the US, Gross State Product (GSP) were downloaded from the US Bureau of Economic Analysis (BEA, 2002) and state power consumption data from US Energy Information Administration (EIA, 2003). A more general term for GSP would be the GRP (Gross Regional Product). It is introduced here to define the contribution of a sub-national areal unit such as a generic administrative region or even square pixel areas to the national GDP total. The term GSP will only be used when explicitly referring to a state as opposed to aggregations of states.

European economic data were purchased from Eurostat, the statistical body of the European Commission. Statistical reporting in the European Union is done according to the Nomenclature of Units for Territorial Statistics (NUTS) system. The NUTS is a five-level hierarchical classification based on three regional levels and two local levels. Each member state is divided into a number of NUTS-1 regions which in turn are divided into a number of NUTS-2 regions and so on. There are 78 NUTS-1 regions, 210 NUTS-2 and 1093 NUTS-3 units within the 15 EU countries (Eurostat, 2002). In spite of the desire to create a single, coherent structure of territorial distribution across Europe, the areal extent of NUTS regions of a given level can differ substantially between countries.

The digital boundaries of these regions were downloaded from UNEP's Global Resource Information Database (GRID) data collection in Geneva (UNEP, 2003). The digital dataset of the US state boundaries was obtained from the US Geological Survey's National Atlas (USGS, 2003).

3. SUB-NATIONAL CORRELATION CHARACTERISTICS: THE UNITED STATES

The log-log correlations between socio-economic parameters and night-time light imagery have been demonstrated to be strong at the country level (Doll *et al.*, 2000). Studies of these correlations at the sub-national level are scarce due to the paucity of disaggregated socio-economic data. A pilot study examined the variation in the correlation between light and GDP at two sub-national units for the UK (Doll, 2003). The strength of the correlation was found to be stronger at the county level than the regional level. Comparing countries at the international level also gives a stronger correlation than at the regional level. This effect is believed to be a manifestation of the MAUP.

This section describes the MAUP for correlations involving night-time light imagery and socio-economic parameters. It will also make inferences about why these variations occur, and secondly what it can tell us about the overall relationship between light and the parameter in question. In building an understanding of the smaller scale issues, it is hoped that a further insight will be gained into modelling socio-economic data from night-time lights.

3.1 State and regional correlations

Multi-scale linear relationships between light parameters and Gross State Product (GSP – a state's contribution to national GDP), and Power consumption were tested for the USA. These analyses were carried out at the state level. For each state, the total amount of radiance was calculated and plotted against the corresponding power consumption or GSP figure.

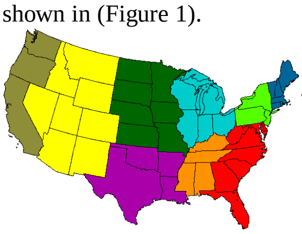
Data are also presented aggregated up to the regional level as defined by the US Census regions (see Figure 1). Examining the relationship between energy consumption, GRP and night-time lights (Table 1) the state level relationship is consistently better than the regional level one. The intra-state relationship is better still for 6 out of 9 regions having a linear R^2 value greater than 0.9. However, intra-regional state correlations reveal that some regions are highly correlated ($R^2 > 0.9$ for the southern states). For all cases, correlations using the total (summed) radiance figures perform better than the lit-area figures.

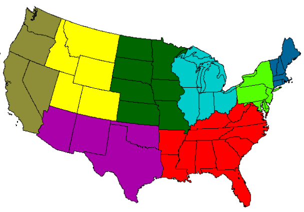
	US Census Divisions	***US States***
Lit-area	0.499 / 0.328	0.699 / 0.563
Log lit-area	0.641/ 0.373	0.729 / 0.467
Total Radiance	0.652 / 0.49	0.844 / 0.75
Log radiance	0.784 / 0.55	0.900 / 0.79

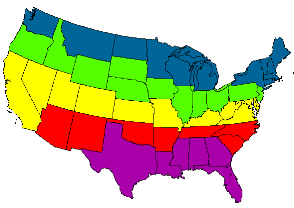
Table 1. Correlation of energy consumption / GRP and night-time lights at state and aggregated US Census Division level.

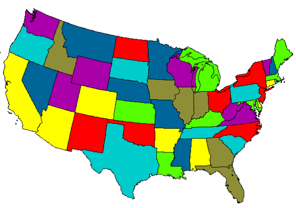
3.2 Regional aggregations

The variation in the correlation of parameters at a given scale depends on how regions are aggregated. In having data at two hierarchical sub-national scales, it is possible to test the national correlation at the smaller scale (regions), by using different aggregations of larger scale units (states). Economic data from the BEA were also presented using their regional classification. Their grouping of states is similar to that of the US Census using traditional aggregations based on states' locations with respect to the various geographical regions (e.g. Plains, Great Lakes, Mountain Division, South Atlantic). In addition to these two classical aggregations, three others were devised. One was a geographical aggregation, which established five regions according to their latitudinal position. The other two consisted of a random aggregation based on seven alphabetical divisions and one, which divided the states after they had been ranked by GSP. The five aggregations are shown in (Figure 1).


US Census Divisions (9)


US BEA (8)


Latitudinal Divisions (5)


Alphabetic zones (7)

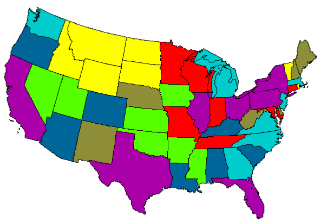

Ranked by Gross State Product (7)

Figure 1. Aggregation of US states into larger zones according to 5 different criteria.

Two of the aggregation methods produce 'regions' which are composed of non-contiguous states. Examining how different aggregations affect the correlation of total radiance with the GRP, it appears that a wide range of results can be obtained depending on how one assembles the regions.

In addition to comparing the r-squared value at this scale, the intra-regional correlation was also computed to see if any discernable patterns were present. The intra-regional correlation here refers to the average correlation of states within a given zone. Figure 2 shows firstly that a regional correlation of different strengths (0.4 –0.95) can be obtained depending on how regions are arranged. Secondly, that the intra-regional correlation (i.e. those states which when summed form one point on the regional level scatter plot) declines as the regional level correlation increases.

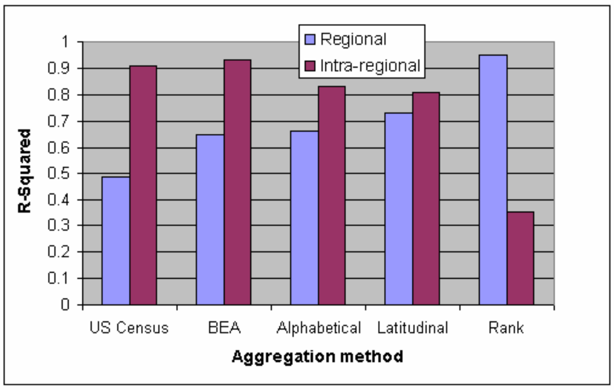


Figure 2. Comparison of regional and (mean) intra-regional correlation coefficients (of total radiance vs. GDP by state) for different aggregation methods.

These results suggest that the US Census divisions seem to be unsuitable for building a national scale correlation, but more appropriate for intra-regional analysis. By shifting one or two states here and there to build the BEA regions, the correlations improve in both measures, though these geographical divisions are generally unsuitable. Even something as random as an alphabetical classification provides a better regional correlation result than the two traditional geographically regional zones. By using a geographical criterion to aggregate states, the latitudinal divisions provide a regional and intra-regional correlation that is most similar to each other. However, ranking states based on their GSP gives the best regional correlation, but the worst intra-regional correlation, despite the component states being of roughly equal economic magnitude to each other. The same pattern of results was also observed for the power consumption data.

4. DERIVED RELATIONSHIPS AND OUTLYING POINTS

One further point to examine is what effect different aggregations have on the magnitude of relationships associated with these correlations. Figure 3 shows the trajectory of the relationship for each aggregation method. Also plotted (bold in red) is the relationship derived from state level radiance-GRP plot. Two points on the plot are of particular interest. Firstly, there is a point around 1E+07 radiance units (point A), where the all-state relationship intersects that of the US Census and BEA relationships. Secondly, further along around 1.75E+07 radiance units (point B) is a point where all the relationships converge *except* the all-state relationship. Despite their divergence at high radiance values, all models would produce similar results if the input values were between these points. The results show that the US Census and BEA relationships are markedly different from those of the other three aggregation schemes. More importantly, it also shows that the overall state-level relationship is most similar to these simulated aggregation schemes and not to those of the conventional US regions. This has important implications for extending this technique to areas that have only one level of sub-national data.

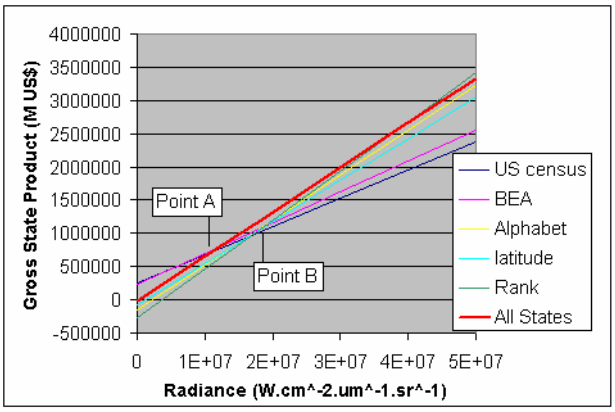


Figure 3. Comparison of different relationships derived from the five aggregation schemes displayed in Figure 1.

Taking these points into consideration it is suggested that the US has a number of regional sub-economies, hence the high intra-regional correlation coefficients of the US Census and BEA divisions. However, the regions themselves vary greatly and bear little relationship to each other. This is not to say that there is no general nationwide correlation. The reason why the intra-regional correlation is so poor for the ranked states may relate back to the nature of the individual regional economies. Assembling regions by ranking states puts the most economically productive states (California, Texas, Florida, Illinois, Ohio, Pennsylvania, and New York) together. However, they apparently bear little resemblance to each other. This trend continues for each group of states. BEA (and US Census) regions by contrast usually consist of one dominant state and a number of less economically productive satellite states. Working down the list of states, each BEA region generally has a good mix of members from each group of (ranked) states. The exceptions are the Plains and Mideast regions which contain Ohio and Illinois, and New York and Pennsylvania respectively, and conversely the Rocky Mountain region, which has no member from the most economically active group of states. However, when viewed in totality their component states fit a single nationwide model albeit with two outliers (New York and California). Many countries anlaysed also exhibited one or two regions which did not fit the general model for the rest of the country.

Spain provides an interesting example of how an outlying point can affect its parent regions. The relationships derived at each NUTS level were found to vary more than any other country. The gradient becomes progressively shallower as larger areas are considered in the analysis ranging through $y = 0.069x$, $0.059x$ and $0.056x$ for NUTS levels-3, 2 and 1 respectively. These values are for relationships constrained to run through the origin and include the outlying points, whose effect becomes less influential as the spatial units become larger in size. The generally good correlation in Figure 4 masks much of the detail. Despite the standard deviation of the GRP/radiance quotients being relatively low compared to other countries analysed the individual NUTS-3 zones have been combined in such a way as to result in vastly different relationships at different NUTS levels and confuse the base-level correlation.

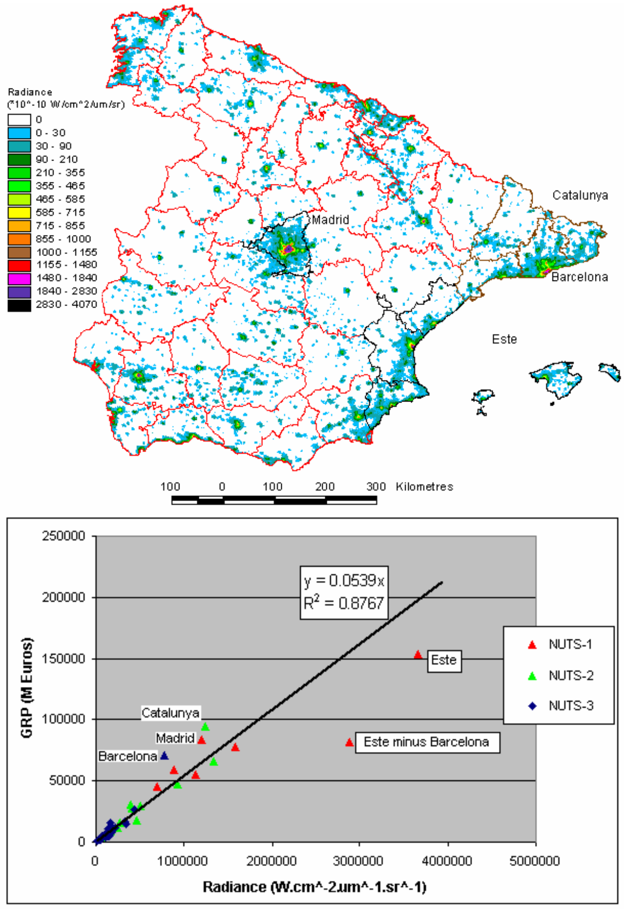


Figure 4. Night-time lights of Spain with its NUTS-3 boundaries and total radiance – GRP scatterplot. The relationship is based on NUTS-3 points only, excluding Barcelona and Madrid. The NUTS-1 regions of Madrid and Este are outlined in black on the map. Cataluña lies within the region of Este and is outlined in brown ($n=40$).

The NUTS-3 area of Barcelona is far more radiant than other NUTS-3 regions and is observed to be an outlier. The capital, Madrid is a NUTS-1 region and is about one third brighter than Barcelona. Barcelona is part of Cataluña, which is itself part of the Este NUTS-1 region. The other regions of Este have anomalously low GRP for the total radiance in the region. Barcelona dominates this region to such an extent that not only does it pull its NUTS-2 point away from the trendline but it has also pulled its NUTS-1 point *towards* it. In fact, it is only due to the presence of Barcelona in the region that its NUTS-1 point is anywhere near the regression line. Its position without Barcelona's contribution is also shown in Figure 4. In this case, due to the magnitude of these modifiable areal unit effects, it is most prudent to just use the original NUTS-3 relationship and treat Barcelona and Madrid as separate outliers. This is the relationship displayed on the graph in

Figure 4. This is the most extreme example of an outlying point's effect being carried throughout the aggregation process encountered for any of the countries analysed and highlights how points can shift in the plot space according to whether an economically overactive area is aggregated with points of average economic activity or below average activity.

4.1 Characteristics of Outlying points

For a few of the countries tested in the analysis of European countries, there is a number of points which deviated from the derived regression line. These points invariably represent heavily urbanised areas, usually a capital city or its hinterland. The scatterplot shown in Figure 4 is a visualisation of how the total amount of radiance in a NUTS region is related to its GDP output on a national scale. It is a more sophisticated measure than a simple lit-area versus GDP plot since pixels of equal areas can have very different radiances. It was hoped that this feature might help to reduce the magnitude of outlying regions but it appears that this is still not sufficient for some countries.

11 EU countries were analysed in a similar manner to that of Spain. Certain countries exhibit extreme outliers, where points on the graph lies well above the main trendline. They are not the most radiant administrative areas but they are the most economically productive. France and Denmark both had these types of scatterplots. There is a number of similarities between the French and Danish outliers. Firstly both Paris and Copenhagen city centres are themselves NUTS regions. They are of the order of 100 km^2, when the average NUTS-3 region is 5700km^2 in France and 2800 km^2 in Denmark. The Ile-de-France region around Paris is also a NUTS-1 region which is a fraction of the size of its sister NUTS-1 regions in France. It is highly urbanised and has a high proportion of higher radiance pixels. The Ile-de-France region is still an obvious outlier even after aggregation. However, if it is further combined with the much larger annular Bassin Parisien region, it begins to align with the other NUTS-1 regions. If a single scale independent relationship exists however, then it shouldn't matter what spatial units are employed to construct the relationship.

The difference in the relative size of NUTS regions is an important issue. Contrasting Spain's NUTS-3 boundaries with that of France or Denmark, Spanish areas are reasonably large and uniform in size – even for Madrid and Barcelona. It begs the question whether a correlation is only as good as the spatial units it is based on. Would Barcelona's NUTS-3 point be an outlier of the magnitude of Paris or Brussels, if it had a small administrative boundary in the centre of the city? If so, then the disaggregation will underestimate the centre of the city and overestimate the surrounding area by virtue of not having an appropriate spatial unit to describe its radiance relationship to economic activity. The same can be said for Milan or Rome, which also have large NUTS-3 areas when compared to the concentrated zones of Paris or Copenhagen. Whilst the problem can be "aggregated away", the application of a 'macro-relationship' to create a disaggregated map would be a misrepresentation of the finer points of the data.

Examining the relationship between total radiance and different sectors of the economy (Figure 5), it is clear to see that different sectors of the make vastly differing contributions to an area. In this example from the Italy, it can be seen that the services sector has a far higher gradient with total radiance than for agriculture.

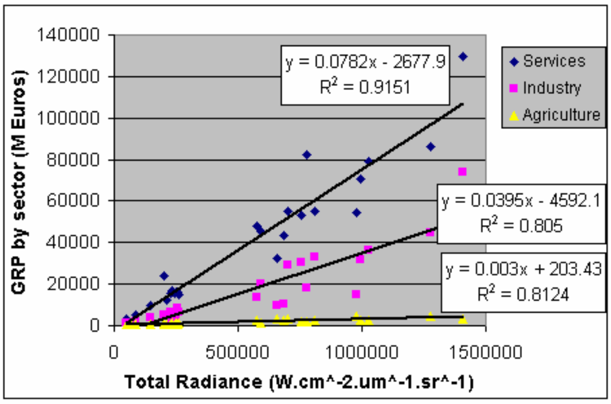


Figure 5. Total Radiance versus GRP by economic sector for the Italy at NUTS-2 level.

A single point in the total radiance-GRP relationship shown for Spain in Figure 4 is the combination of each set of points for the three economic sectors. Most regions have an even mix of each type of economic sector within a NUTS region, however very small urban NUTS areas such as Paris or Copenhagen are exclusively urban and the service sector relationship dominates the economic make-up of the area generating an outlying point in the scatterplot.

Since outlying points in the scatterplot are smaller in size as well as highly urbanised NUTS regions, two additional metrics may be identified as being potentially helpful to discriminate these areas. Firstly, the mean radiance of a region could be a useful feature to discriminate those areas, which, while having equal total radiance, have different GRP. Secondly, the proportion of area lit in a NUTS region can give an indication of how urbanised that region is and therefore an inference can be made as to which sector of the economy is contributing to its GDP. The fact that different economic sectors have different GRP per unit radiance values suggests that a priori knowledge of these areas (from a land use map for example) could be used as weighting functions for a region.

5. MAPPING RESULTS FROM RADIANCE-CALIBRATED DATA

What implications do these observations have for the practical mapping of economic activity? One key consideration is how far it is prudent to extend the relationship beyond the known data range to finer spatial resolutions (The Ecological Fallacy).

Reconciling what the map aimed to display and the technical specifications of the DMSP-OLS sensor affected the choice of spatial resolution for the map. Whilst the relationship between total radiance and GRP has been demonstrated to be consistent from the NUTS-3 level up to NUTS-1, an unknown factor is how far the relationship can be extrapolated in the other direction (NUTS-3 to finer scales) before it becomes invalid. In addition to this unquantified micro-scale relationship, there is a desire to maintain an aggregate nature to the GDP figures in the map. Figure 5 shows that different relationships exist between the different economic sectors and total night-time radiance. The overall derived relationship such as that shown in Figure 4 is an aggregate of these sector combinations. Keeping the map resolution sufficiently coarse avoids issues of sector specific GDP representations at finer scales where the aggregate relationship may not be valid. Although the night-time data has a nominal resolution of 1km, this is resampled from 2.7km raw data. This also supports a decision to increase the spatial resolution of the map by reducing any noise effects present in the fine resolution data.

Bearing these points in mind, the output resolution was set at 5km as this was reasoned to provide both a detailed map at the continental scale whilst being coarse enough to generalise economic activity to a level which shows enough detail in large cities without compromising small towns. In addition to this is the advantage of making the mapping process less computationally intensive. Outlying areas identified for certain countries were excluded from the main relationship of that country. However, values were attributed to the radiance cells in these areas by allocating the remaining value from the national total to these areas thus ensuring that all of the cells in a country sum to the published total. The map will over estimate and under estimate certain areas. This is an inevitable consequence of the method. Nonetheless it is encouraging to see that very little adjustment to the value of outlier(s) was required to constrain the map to a country total. Full details of the results and maps are presented in Doll *et al*. (2004)

6. CONCLUSION

Although, the method used to construct the economic activity map from a given relationship has the inevitable consequence of over- and underestimation, alternative aggregations of the finest-scaled zones data did not appear to give substantially different results compared to the administrative NUTS-1 aggregation from the original data. Estimating country level GDP in the EU was found to be accurate to within 5% for most countries and within 7% for all. For a given country-level percentage error, the method used for the treatment of outliers will yield an increasing amount of error between its predicted and observed value as the outlier forms a decreasing percentage of total GDP for that country.

This is a highly encouraging result when considering that the map is based on just one variable and re-affirms night-time light data as a promising tool for mapping socio-economic parameters. Economic activity mapping could also benefit from the inclusion of other variables as has been shown with respect to human population mapping from the Landscan project. The use of such data would facilitate the construction of ‘poverty maps’ where GDP/capita can be spatially mapped thereby highlighting areas of social inequality. An example of this for Guatemala can be found in Sutton et al. (2004). Indeed, the application of these datasets are arguably of most use to the developing world. The importance of other variables such as land-use and transportation networks is likely to become more acute as the spatial resolution of the map increases.

REFERENCES

BEA, 2002. *Bureau of Economic Analysis: Gross State Product data*. Available from: http://www.bea.doc.gov/bea/regional/gsp/

Cao, C. and Lam, N, S-N., 1997. Understanding the Scale and Resolution Effects in Remote Sensing and GIS. *In:* Quattrochi, D.A. and Goodchild, M.F. (Eds.) *Scale in Remote Sensing and GIS.* CRC Lewis: Boca Raton, FL, USA, 57-72.

Doll, C.N.H., Muller, J-P. and Elvidge, C.D., 2000. Night-time imagery as a tool for global mapping of socio-economic parameters and greenhouse gas emissions. *Ambio*, 29 (3), 157-162.

Doll, C.N.H., 2003. Estimating non-population parameters from night-time satellite imagery. *In*: Mesev, V. (Ed.) *Remotely Sensed Cities.* Taylor & Francis: London, UK, 335-354.

Doll, C.N.H., Muller, J-P. and Morley, J.G., 2004. Mapping regional economic activity from night-time light satellite imagery. Submitted to *Ecological Economics.*

EIA, 2003. *Energy Information Administration: Multi-state data.* Available from: http://www.eia.doe.gov/emeu/states/multi_states.html

Elvidge, C.D., Baugh, K.E., Kihn, E.A., Kroehl, H.W., Davis, E.R. and Davis, C.W., 1997. Relation between satellite observed visible-near infrared emissions, population, economic activity and electric power consumption. *International Journal of Remote Sensing*, 18 (6), 1373-1379.

Elvidge, C.D., Baugh, K.E., Dietz, J.B., Bland, T., Sutton, P.C. and Kroehl, H.W., 1999. Radiance calibration of DMSP-OLS low-light imaging data of human settlements. *Remote Sensing of Enviroment*, 68 (1), 77-88.

Elvidge, C.D., Imhoff, M.L. and Sutton, P.C., 2000, Relation between fossil fuel trace gas emissions and satellite observations of nocturnal lighting. *International Archives of Photogrammetry and Remote Sensing*, GITC bv: Amsterdam, The Netherlands, 33 (B7), 397-401.

Eurostat, 2002a, *Main characteristics of the NUTS.* Available from: http://europa.eu.int/comm/eurostat/ramon/nuts/mainchar_regions_en.html

Openshaw, S., 1984. The modifiable areal unit problem. *Concepts and Techniques in Modern Geography 38,* GeoBooks: Norwich, UK.

Sutton, P.C., Taylor, M. and Elvidge, C.D. 2004. Urban populations in the developing and developed countries - the use of night time data. *In*: Jurgens, C and Rashed, T (Eds.) *Remote Sensing of Urban and Suburban areas.* Kluwer Acadamical. *In Press.*

UNEP, 2003a. *UNEP-GRID GNV158 - Administrative Regions/Boundaries of Europe.* Available from: http://www.grid.unep.ch/data/grid/gnv158.php

USGS, 2003. *State Boundaries of the United States*. Available from: http://nationalatlas.gov/statesm.html

Wrigley, N., Holt, T., Steel, D. and Tranmer, M., 1996. Analysing, modelling and resolving the ecological fallacy. *In:* Longley, P. and Batty, M. (Eds.). *Spatial Analysis: Modelling in a GIS Environment.* New York, USA: Wiley, 25-41.