relu
Rural Economy and Land Use Programme

Harnessing the social and natural sciences for sustainable rural development

We inhabit the same world: integrating social and natural science data

Veerle Van den Eynden
RELU Data Support Service
UK Data Archive
www.data-archive.ac.uk/relu

UK Data Archive

Centre for Ecology & Hydrology
NATURAL ENVIRONMENT RESEARCH COUNCIL

E·S·R·C ECONOMIC & SOCIAL RESEARCH COUNCIL

NATURAL ENVIRONMENT RESEARCH COUNCIL

bbsrc

Rural Economy and Land Use programme

- **interdisciplinary research programme - social & natural / environmental sciences**
- **teams of social and natural scientists study contemporary challenges facing rural areas in the UK (food issues, flooding, land and water use, agriculture, plant/animal diseases,….)**
- **cross-research council funding**
- **dedicated data support service**

RELU Data Support Service

- **UK Data Archive + Centre for Ecology and Hydrology**
- **oversee implementation of RELU data policy**
  - **publicly-funded research data = valuable, long term resource**
  - **data generated by RELU projects must be well managed**
  - **data must be made available by researchers for archiving**
- **support RELU researchers: information, training, etc. on data management, data archiving / sharing**

Presentation overview

- **examples of data integrations social-natural science data by RELU projects**
- **integration challenges encountered**
- **can data archives / data centres help?**

Data integration

Examples of research integrating social science and natural science data:

- **Spatial integration (GIS)**
- **Modeling**
- **Databases / spreadsheets**
- **Visualisation / conceptualisation**

Data used by RELU projects

Frequently used existing datasets:

- **Land Cover map (CEH)**
- **Countryside Survey (CEH)**
- **BADC climate data**
- **Biodiversity data**
- **Admin. boundaries (Digimap, OS)**
- **Census data (ONS)**
- **Agricultural Census (ESDS / EDINA)**
- **Farm Business Survey (ESDS)**
- **National Diet and Nutrition Survey**
- **Food and Expenditure Survey**

RELU project:
Integrating Spatial Data on the Rural Economy, Land Use and Biodiversity

Piran White, University of York

Spatial integration of land use data with biodiversity data within a GIS

study trade-off between agric. productivity and biodiversity

- **disaggregate agricultural census data for England at ward level (aggregated from individual farm data) to 1km² level; for integration with bird species distribution (10x10km) and abundance (1x1km)**
- **using genetic algorithms, with CEH Land Cover map (land cover types per km²) as key**
- **genetic algorithms: evaluate various estimations of % land use within a land cover class, evaluate fitness and assign to km² grids**
- **divide ward into grid, use land cover to assign most probable land use distribution in wards**

Integrating Spatial Data on the Rural Economy, Land Use and Biodiversity

Challenges of data integration:

- **differing areas: land use corresponds to area farmed; land cover corresponds to total area**
- **some land use areas suppressed due to disclosure**
- **agricultural census data – discrepancy farm postcode vs farming land**
- **grid agricultural census data exist (AgCensus, EDINA), but not available at km² level (confidentiality)**

http://www.relu.ac.uk/research/projects/White.htm
(Huby et al 2006)

RELU project:
Combining Socio-Economic and Environmental Data for Rural Areas

Meg Huby, University of York

Spatial integration of environmental data and socio-economic data for England, mapped onto SOAs within a GIS

to consider social and environmental perspectives of rural problems

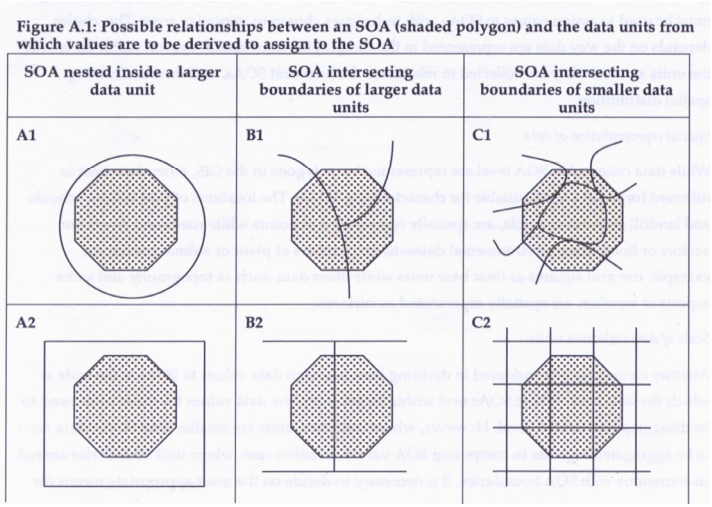
- **base area = Super Output Areas (SOAs):**
  - **2001 Census outputs map directly onto SOAs**
  - **consistent population size areas (400 households)**
  - **fit within ward boundaries / larger administrative areas**
  - **commonly used for govt. social data & small area statistics**
- **variety of environmental data (land use, protected areas, biodiversity, pollution, etc.) and socio-economic data (demography, housing, economic activities, travel, services, deprivation, public health, etc.) are mapped onto the SOAs**

Combining Socio-Economic and Environmental Data for Rural Areas

- combining various data formats: point data (pollution, places of interest), line data (river, road), grid data (species richness, climate)
- different size areas:
  if unit > SOA – disaggregate data value
  if unit <SOA – aggregate data values
- calculations (sum, average,….), weighing by area or population size

Challenges of data integration:

- SOAs vary in size
- different size areas
- different shape areas
- intersections



Source: Huby, Cinderby, & Owen 2005

Combining Socio-Economic and Environmental Data for Rural Areas

- **varying underlying data distribution:**
  - **uniform over a SOA (e.g. party political controle of local authority)**
  - **irregular variation (e.g. land cover, land use, settlement-related data)**
  - **continuous variation (air pollution)**
- **data integration problems occur equally within disciplines due to different scales (e.g. allign postcode sector data with SOAs)**
- **variations also depend on how data are collected and organised**

http://www.relu.ac.uk/research/projects/Huby.htm

(Huby et al 2006, Huby et al. 2005)

RELU project:
Sustainable and Safe Recycling of Livestock Waste

Dave Chadwick, IGER

Risk assessment modeling of FIO run-off risk at individual farms in a single catchment area, based on social and natural parameters

Risk indicators:

- **E. coli burden**
- **run-off potential (landscape-related environmental factors)**
- **infrastructure characteristics**
- **farmer attitudes / approaches to manure, land and animal management**

Farm-scale risk assessment tool allocating risk class:

risk class = ∑weighing factor x risk indicator

whereby factor weighing is allocated through multidisciplinary specialist consultation

Sustainable and Safe Recycling of Livestock Waste

Data integration:

All social data collected by researchers as primary data at farm level: scale corresponds directly with farm-level natural parameters; confidentiality is no issue

Relative weighing of social vs natural parameters is achieved through consultation with group of social and natural scientists as specialists

http://www.lec.lancs.ac.uk/cswm/foodchains/po.php
http://www.relu.ac.uk/research/projects/Chadwick.htm

RELU project:
Implications of a Nutrition Driven Food Policy for the Countryside

Bruce Traill, University of Reading

Modeling national nutrient / food demand based on household consumption and health

Using data from Family Expenditure Survey, National Food Survey, Expenditure and Food Survey, National Diet and Nutrition Survey

Land Use Allocation Modeling (LUAM) for England and Wales

Model considers England and Wales as a single farm, models likely changes in agricultural practices, and the environmental impacts associated with them, due to changes in policies or market conditions

Uses macro-level input data (national statistics) from Farm Business Survey, Agricultural Census, CEH Land Classes

RELU project:
Implications of a Nutrition Driven Food Policy for the Countryside

Challenges of data integration:

LUAM model: data are used at macro level, thus all data are available from national statistics; no scale issues

Nutrient modeling: survey data planned for use are collected at household or individual level (different samples), but:

- **there is no common variable to merge / integrate data**
- **data made available at macro-scale (e.g. GOR level) to avoid disclosure, which is not detailed enough**
- **data could be integrated through postcodes, but data are not available at this level**

http://www.relu.ac.uk/research/projects/Traill.htm
(Jones & Tranter 2007, Arnoult & Tiffin 2007)

Challenges of data integration

- different data scales - finest resolution is lost during integration
- micro-scale data often confidential and not available
- agricultural data / statistics linked to farm site rather than actual field site
- different shape areas:
  - social science data ~ administrative / political areas; subject to temporal change
  - natural science data ~ ecological zones, grids
- dealing with data area intersections and allocating data to areas
- lack of common variables for integration
- challenges equal within as between disciplines

If primary data are collected, many integration challenges are avoided: scale and area can be chosen, etc.

Can data archives assist data integration?

- **common variables to enable linking**
- **geospatial information for place-related social science data**
- **organise and archive data in smallest units (microscale), thus enabling aggregations at varying size, shape, level**
- **enable access to micro-scale data – 'approved' researchers**
- **does confidentiality force repetition of data gathering?**
- **metadata mapping**
  - **social science data – DDI**
  - **natural science data – ISO 19115, Dublin Core, DIF**
- **metadata details: methods, sampling, units, …**

and….

Can data archives assist data integration?

Suggestions?

References

Huby, M., S. Cinderby, A. M. Crowe, S. Gillings, C. J. McClean, D. Moran, A. Owen and P. C. L. White. 2006. The Association of Natural, Social and Economic Factors with Bird Species Richness in Rural England. Journal of Agricultural Economics 57(2):295-312.

Huby, M., S. Cinderby, A. M. Crowe, S. Gillings, C. J. McClean, D. Moran, A. Owen and P. C. L. White.2006. The Association of Natural, Social and Economic Factors with Bird Species Richness in Rural England. Journal of Agricultural Economics 57(2):295-312.

Huby, M., Cinderby, S. & Owen , A. 2005. Social and Environmental Conditions in Rural Areas. Report to accompany the SECRA dataset produce under the Rural Economy and land Use programme. Dept of Social Policy and Stockholm Environment Institute, University of York. [www.sei.se/relu]

P. J. Jones and R. B. Tranter. 2007. Modelling the impact of different policy scenarios on farm business management, land use and rural employment. Workpackage No. 5, Report No. 02 http://www.relu.rdg.ac.uk/Working%20Papers%20and%20Reports/Doc13WP5Rep02LUAMPolicyScenarios.pdf

M.H. Arnoult & R. Tiffin. 2007 Econometric Models of Nutrient Demand. Workpackage No. 1, Report No. 02. http://www.relu.rdg.ac.uk/Working%20Papers%20and%20Reports/Doc07WP1Rep02EconModels.pdf