



**The Global land Project Nodal Office on  
Integration and Modelling, Aberdeen, UK**

## **GLP Workshop on**

Data and model integration for coupled models of land use  
change

## **Post Workshop Report**



### **Introduction**

The Global Land Project Nodal Office on Integration and Modelling held its second workshop on July 17<sup>th</sup> – 19<sup>th</sup> 2008 at The Burn House, Glenesk, Aberdeenshire. Integrated models of land systems require a variety of data to drive them. Data can be for different spatial and temporal scales (from the global to the local) and record different aspects of the land system (socio-economic, biophysical etc.). This workshop explored some of the issues that arise when using different data types and sources in integrated models of land systems. Below is a short summary of the presentations and discussions for each session. This is followed by a list of actions and recommendations that came out of the workshop.

## **Workshop Sessions**

### **Session 1 Data sources and issues**

#### *Presentations*

‘From remote sensing of land cover and vegetation to data / model integration’ by Hieko Balzter from the University of Leicester, UK during which Dr Balzter presented aspects of remote sensing information in relation to land cover and land use change.

‘USA agricultural and rural land use data sources’ by Shawn Bucholtz. Mr Bucholtz discussed three categories of agricultural land use/land cover data: survey, administrative and remote sensing, with specific reference to the case in the USA.

#### *Discussion*

In this session two break out groups discussed issues surrounding data sources for coupled models of land use change. The groups talked about data standards and choosing which data you can ‘trust’ and the fact that standardisation is very often voluntary. One group also talked about GLP having a standardized way of reporting in a similar way to the IPCC or UNFCCC. Land function (a concept mentioned several times in Workshop 1) was discussed, framed around hydrological and nutrient cycling but also ecosystem services. This led to a discussion of the way people see landscapes and how this depends on where they are from.

The issue of how much time is needed before change can be detected and the time-steps at which different data are currently recorded were talked about by one group, leading to a recommendation for a study of the time steps currently being used to collate different types of land use data and the reasoning behind choosing this time-step.

Data types were discussed at length by one group. It was agreed that ‘socio-economic’ data is often talked about but that this is not a useful term as it is too broad. The need to separate management decisions from behaviour was made along with the observation that people are often unaware of all of the drivers of their own behaviour. Several times during the discussion the point was made that socio-economic data is context specific and not always spatially specific as agents (people, organisations and governments) can move around. An observation was made that modelling can give information about mechanisms and sometimes predictions of what might happen, but less frequently when it will happen or how.

Remote sensing (RS) data was also discussed. Land cover changes detected by RS are a mixture of real change and uncertainty. Land use intensification is never known which presents a big problem. The idea of GLP producing a database of regional land management information that could inform global analysis was put forward. Both groups independently came up with the recommendation that the GLP should formally link to FAO to join efforts to collate regional land management information. It was also suggested that the GLP focus efforts on putting past land use data into a useable format.

## **Session 2 Using data sets of different scales, merging, scaling and aggregation issues**

### *Presentations*

'Modelling ecological processes at the meso-scale (10-1000km<sup>2</sup>)' by Nicolas Dendoncker, the University of Edinburgh. This presentation explored some of the issues of modelling land use at an intermediate scale, examples of scaling up and down, and why an integrated ABM approach could be suitable for use at an intermediate scale.

'The influence of scale on the analysis of key drivers of land-use/cover change processes' by Tom Veldkamp from Wageningen University. Prof. Veldkamp gave a broad ranging presentation that considered multi-dimensional land use/cover driving forces and how these act and interact.

'Linking Maasai Decision Making to Ecosystem Services, including Data Sources and Scaling Issues' by Randall Boone. Dr Boone talked about an ongoing project that is using agent based modelling to link Maasi decision making to Ecosystem services.

### *Discussion*

The Session 2 discussion considered data processing, scaling and aggregation issues and the way in which ecosystem services are represented in integrated modelling. It was pointed out that different data processing techniques can greatly affect scenario results and that this should be noted and with information being made available to others using the processed data. In terms of scaling, one group pointed out that the global scale is difficult if not impossible to validate/calibrate for decision making models. If working at a large scale the need to understand how levels of aggregation influence linkages was highlighted, as was the need to present multi-scale measures of particular landscape metrics which may inform you of whether or not your landscape metric is sensitive to scale.

The limitations of modeling were noted, with models being seen as being useful for understanding system dynamics and being used to gain insight and learn about the system. This led onto a discussion of the potential misuse of model results and the responsibilities that come with communicating model results to policy makers. If an Ecosystem Services approach is taken, issues of scaling and integration need to be recognized with this. Choice of scaling needs to be explicit. The fact that models generally perform better at coarser scales was noted, as was the fact that interdisciplinary modeling inevitably has a disciplinary bias when scale is chosen. While up-scaling was seen as relatively easy, downscaling was seen as more problematic and the need to document the method used was highlighted.

### **Session 3 Collecting data for modelling**

#### *Presentations*

‘Data challenges in creating LURNZ, an econometrically based model of Land Use in Rural New Zealand’ by Prof. Suzi Kerr, Motu, New Zealand. Prof Kerr talked about some of the challenges that her group faced when collating data for this particular land use model.

‘Integrating Urban Land use Dynamics and Precipitation for Flood Analysis in Lagos’ by Dr Odunuga Shakirudeen. This presentation looked at a case study in Nigeria and some of the issues that were faced by the case study in terms of data collation for land use modelling.

‘History Database of the Global Environment (HYDE)’ By Dr Kees Klein Goldewijk. In contrast to the two previous presentations, this presentation talked about a project which is trying to assess historical land use for the last 12,000 years from a variety of data sources.

#### *Discussion*

Group 1 raised the importance of census data being collated (an example was given in Suzi Kerr’s talk of the New Zealand Government stopping census data collection for a time which led to data gaps). Ensuring that those responsible for collecting census data know that it is being used and how valuable it is was seen as crucial. The need to get researchers to think in terms of other people potentially using their data in the future was raised with the suggestion again made by both groups that processed data should be accompanied by proper documentation. Group 2 talked about the issue of getting data from other organisations, such as FAO and recommended that a formal link be established between the GLP and FAO and similar organisations, especially as there is a common link through the UN. Group 2 also pointed out the importance of continuity in the way data is collected and that the structure of many data holding bodies needs to be improved.

Issue of paying for data ‘piecemeal’ were discussed, with recognition that many times researchers pay for the same data several times over. It was also recognised that much data exists in a hard format and needs to be put into electronic format. It was suggested that the GLP discover and make recommendations for priority hard copy datasets to put into electronic format. It was also suggested that more projects are funded that have an accessible database as a final product. However it was also noted that the availability of data online aids the misuse of data. Ideally research should start with the question not the data that happens to be available.

### **Session 4 & 5 ‘Data infrastructures and architecture’ and ‘Ethical issues and data protection, environmental data and socio-economic data’**

#### *Presentations*

Session 4 had two presentations associated with it, the second of which covered issue from both Session 4 and Session 5.

‘Land use data: Integrating socio-economic and ecological data within the HANPP framework’ by Karlheinz Erb, Klagenfurt University, Austria. HANPP or

Human Appropriated Net Primary Production provides an indicator of land system change. HANPP through the consumption of food, paper, wood and fibre alters the composition of the atmosphere, levels of biodiversity, energy flows within food webs and the provision of ecosystem services.

'Data infrastructures and architecture + Ethical data issues' By Richard Aspinall, Chief Executive of The Macaulay Institute, UK. In this presentation Prof. Aspinall talked about data infrastructures and architecture in context of the GLP, e.g. as a means of providing data for modelling and analysis of land change dynamics and consequences of change in land systems relevant to application needs of different users.

### *Discussion*

In general both groups found that many issues to do with data infrastructure and architecture had already been discussed in the previous sessions. Therefore the discussion focused mainly on ethical data issues. These included ways of protecting data, through licenses and confidentiality agreements and how these work once data is freely available on the web. One group discussed the possibility that data could be 'given' to the GLP for it to manage for community use. Both groups mentioned that scientists are already bound by certain ethical codes and professional codes of conduct.

Group 2 talked about the problems of using land use information at the individual household level and how this can conflict with privacy laws. Conflicts between the UK Freedom of Information Act and the Data Protection Act were discussed and it was noted that Google Earth has superseded much national legislation and raises data protection issues (e.g. number of cars on your driveway). Group 2 discussed the use of model output by policy makers and considered if the GLP as a whole needed to consider how output might be used before projects are endorsed. The recommendation was put forward that ethical issues surrounding data collation and use of model output be mentioned in the GLP project endorsement process.

### **Actions/Recommendations**

The following actions and recommendations came out of the discussion sessions at the workshop.

#### **Session 1.**

1.1 GLP produce a database of examples of land use and land management parameters for specific regions/areas tapping into regional expertise. This could then be drawn on by those carrying out global assessments. As a starting point this could be carried out for grazing lands (with the idea of extending it to croplands, other land uses as a next step).

1.2 Put existing data into a useable format. Prioritise which 'existing data' should be formatted first.

1.3 'Socio-economic data' is a very broad term. The GLP and integrated land science community should break this term down in a consistent way. Also GLP should encourage recognition of the role of 'non-mappable' social information

1.4 Link land use intensification information to remote sensing images

1.5 GLP partner with FAO efforts on data collection/formatting. Take advantage of the common IHDP/IGBP UN root

1.6 Activity to find the time intervals being used for land use/land cover data. Information collated on the time intervals being used and why with the idea of finding an ideal or suitable range

1.7 Come up with a standardised way for GLP synthesis activities to report, in the same way that the IPCC and UNFCCC do.

## **Session 2.**

2.1 Understand of levels of aggregation influence processes and linkages in integrated models. Use the same data set with different levels of aggregation and see how this affects results (has been done in some instances already)

## **Session 3.**

3.1 GLP highlights the importance/usefulness of census data and communicates this to those who collect it

3.2 GLP through endorsed projects encourage researchers to think long term about how their data might be used after their own investigation is finished. E.g. data is available with clear documentation on how it has been 'cleaned up' or processed – GLP protocol for dealing with data? Encourage data producers to archive their data in its rawest form for use when new techniques are developed.

3.3 More individual country datasets are made available

3.4 How people are using classifications/definitions are given in project documentation

3.5 A Wiki for regional historical land use information

3.6 GLP points out the value of census data being geo-coded

3.7 Recommend more activities where the database itself is a product

3.8 Keep uncertainty on the agenda

3.9 Recognise in GLP that there are so called empirical and theoretical based models but we cannot necessarily distinguish between the two. A dialogue should be encouraged on the theory behind them to discover similarities and overlaps and

promote better understanding between different modelling groups coming from different disciplinary angles.

3.10 GLP to 'campaign' to administrations on the importance of continuity of land use statistics datasets. Also quality of data – harmonisation and standardisation.

3.11 Publishing: referees need to make sure there is enough description of the data collection and processing. Electronic resources made this better.

3.12 Should start with what the question is and then go to data sources and methods of answering, rather than the other way round. Pass this wisdom on! GLP publish excellent example of this on website/newsletter.

3.13 Training courses, e.g. Modluc? European commission framework 7 Marie Curie program – funding.

3.4 Encourage collaboration between researchers from developed and developing countries.

3.5 Encourage research and data generations in regions that have less data/information.

#### **Session 4.**

4.1 Support global databases through international networked contribution to their development and testing

4.2 Investigate the relations between temporal interval of land cover data and detection of change

4.3 Investigate the scaling relationships in model results. Can these be detected through use of multi-resolution versions of data (and the effects of generalization)

4.4 Examination of differences in global responses to price/policy changes through national comparisons?

4.5 GLP Project Endorsement procedure to include awareness of ethical issues and codes of practice

4.6 Database of models and code starts with endorsed projects

4.7 Amend the endorsement procedure to an awareness of potential ethical issues

4.8 Researchers already work to codes of practice, have this recognised by the GLP