*International Environmental Modelling and Software Society (iEMSs)*
*7th Intl. Congress on Env. Modelling and Software, San Diego, CA, USA,*
*Daniel P. Ames, Nigel W.T. Quinn and Andrea E. Rizzoli (Eds.)*
*http://www.iemss.org/society/index.php/iemss-2014-proceedings*

# Linking Bayesian Belief Networks and GIS to assess the Ecosystem Integrity in the Brazilian Amazon

**Peter Verweij**[a], **Margareth Simoes**[b], **Andrei Alves**[c], **Rodrigo Ferraz**[b], **Anouk Cormont**[a]

[a]*Alterra, Wageningen-UR, peter.verweij@wur.nl*
[b]*Embrapa Solos*
[c]*Universidade do Esado do Rio de Janeiro*

**Abstract:** Deforestation and climate change heavily impact the ecosystem of the Amazon rainforest threatening its resilience and the sustainability of many human activities. Land protection may prevent ecosystems and their services to deteriorate from the pressures of agricultural expansion, population growth and wood harvesting. In the Brazilian Amazon land protection occurs in several forms such as environmental conservation, setting biodiversity priority areas and the delineation of indigenous lands. Still, the effects are not clear as understanding of the ecosystems is incomplete and responses to human actions are highly uncertain. Bayesian Belief Networks (BBN) are models that probabilistically represent correlative and causal relationships among variables. BBNs have been successfully applied to natural resource management to address environmental management problems and to assess the impact of alternative management measures. By training the probabilistic relationships using field data, Remote Sensing data and GIS data the BBN can provide information on the ecosystems: the ecosystem integrity and their likely response to climate change or alternative management actions. An increasing number of studies train and apply BBNs with evidence originating from GIS data; a cumbersome and error prone soft-linking method requiring manual conversion of data files between the BBN and GIS software systems. This paper presents the full integration of a BBN software system within an existing GIS based Discussion Support System (DSS) illustrated by the case of the ecosystem integrity of the Brazilian amazon. The full integration speeds up the processing and thereby allows doing multiple runs within a short period of time such as a stakeholder workshop. Each consecutive run is based upon insights from a previous one. Furthermore, the DSS provides the management of different options, visualize spatial summaries and trade-offs between different impact indicators and see regional differences.

**Keywords:** Ecosystem integrity; participatory modelling; Bayesian Belief Network; GIS; DSS.

## 1. INTRODUCTION

### 1.1 Uncertain causalities in assessing Ecosystem Integrity

Ecosystem Integrity (EI) is used as a synonym for intactness, completeness and integration of ecosystems. Sustainable ecosystem management has healthy functioning of ecosystems as goal by integrating human society with the natural environment with benefits for both (Constanza, 2012). The concept of EI is however heavily discussed and a huge amount of definitions are used. Jorgensen and Müller (2000) describe EI as a balanced, integrated adaptive community of organisms having a species composition, diversity, and functional organization comparable to that of a natural habitat of a region. Supporting and maintaining such an ecosystem requires indicators to assess its state, as a way to operationalize the concept of EI. Indicators used to assess EI are e.g. sun energy input (radiation balance, production of biomass) and export (respiration, transpiration), storage capacity of plants, biotic water flow (evapotranspiration), nutrient flows (mineralization) and biotic diversity (species number). Still, the causality of the relations between these indicators and the state of EI are not clear as understanding of the ecosystems is incomplete and responses to human actions are highly uncertain.

### 1.2 Bayesian Belief Networks and GIS

Bayesian Belief Networks (BBN) are models that probabilistically represent correlative and causal relationships among variables. BBNs have been successfully applied to natural resource management to address environmental management problems and to assess the impact of alternative management measures. By training the probabilistic relationships using field data, Remote Sensing data and GIS data the BBN can provide information on the ecosystems: the ecosystem integrity and their likely response to climate change or alternative management actions.
While BBN's are used to study results from deliberative participatory questionnaires linked to GIS-data (e.g. Gret-Regamy et.al., 2013) and in preference elicitation methods with a very little amount of spatial entities (e.g. Haines-Young, 2011) few studies have fully integrated BBNs and GIS and explored the resulting benefits (Stelzenmuller et al., 2010). It is expected that tools integrating BBN and GIS are being developed both by academia and commercial entities (Ames and Anselmo, 2008).
In this study we fully integrate a BBN with GIS, use it to map a large number of spatial entities and have multiple iterations based on results during a single group session using the Ecosystem Integrity for the Brazilian Amazon as case study.

## 2. METHOD

### 2.1 Model development

Initially ecosystem integrity experts and Bayesian Belief modelers set up a conceptual map (Novak, 1991) and converted it into a prototype ecosystem integrity BBN containing solely qualitatively variables.
In order to spatially explicate the model, regional experts and GIS and Remote Sensing specialists were added to the team. During several workshops and tele-conferences the definition of the ecosystem integrity was iteratively adapted. Each adaptation built upon the results of a previous iteration by eliciting feedback from the experts. Present time quantitative indicators, like leaf area index (Watson, 1947) and vegetation cover (Amthor and Baldocchi, 2001), originating from Remote Sensing and qualitative GIS layers, such as land use and protected areas were included to improve the model to better fit the experts' expectations.
Under the assumption that human imposed measures drive pressures that impact the state of the Ecosystem Integrity, several scenarios (van Eupen et al., in prep.) were developed on which basis conservation zones with implicit management were developed and data-intensive models were run: The CLUE model (Veldkamp and Fresco, 1996; Verburg et al, 2010; vanEupen et al., 2014) for projecting land use changes. The scenario maps will be used to replace the current situation equivalents in the model and to determine the likely impacts of these scenarios on the ecosystem integrity.

### 2.2 Technical framework

The BBN modellers used the NETICA software[1] for capturing expert knowledge including the uncertainty therein, training the network with known cases and determining the ecosystem integrity by inference. The spatial data was stored in raster GIS files with a resolution of 1.66 10-2 degrees2 per cell and with a total of 3600 rows x 2800 columns.
QUICKScan is a discussion support software system that links spatial- and statistical data to knowledge rules (Verweij et.al., 2012). Typically knowledge rules define cause effect relations in the form of if..then..else rules and are formulated and applied during a participatory workshop. Knowledge rules capture expert knowledge, layman's knowledge and preferences. QUICKScan shows regional differences and trade-offs, creates summary charts and allows drilling down from calculation results into the decision path.
The QUICKScan software uses the OpenMI model-linking framework (Gijsbers et al, 2005; Knapen et.al., 2013) in which each model is wrapped with a generic interface to trigger the model to execute and exchange data with other models. Janssen et.al. (2009) distinguishes between: conceptual-, methodological- and technical model links. OpenMI is primarily a framework for technically linking models and has limited support for methodological integration by describing the metric, the time- and

[1] https://www.norsys.com

the spatial dimensions of the exchanged data. In QUICKScan knowledge rules and GIS raster maps are wrapped with the generic OpenMI interface.
By wrapping a Netica BBN with the same generic interface the BBN can be included in QUICKScan and linked to any other QUICKScan model components such as raster maps. A user interface assists in mapping input/output variables with metrics to respectively BBN nodes with states.

### 2.3 Approach comparison

When inferring the BBN using input maps from GIS/Remote Sensing each EI state is appointed a probability. In order to get a single EI map two alternative approaches were followed:

- Mapping the highest probable state
- Map a random sample from the state probability distribution. This method maintains the overall probability distribution in the created EI map in comparison to the original BBN probability distribution (Stewart, 2011).

## 3 RESULTS

### 3.1 Software architecture and Graphical User Interface for BBN component

The increasing size and complexity of software force the use of abstraction and to break the system down into separate elements of concern in which each element has its own functional responsibilities. Such a common abstraction of a system, or architecture, manifests early design decisions and software qualities like modularity and extensibility, through which the system can be analysed (Verweij et al., 2010). The QUICKScan software uses a common *layered* architecture in which the system is split up into a number of independent layers, in this case 3 layers: (1) presentation, the *Graphical User Interface*; (2) program logic representing domain logic; and (3) *persistency* for storing session data (see Figure 1).

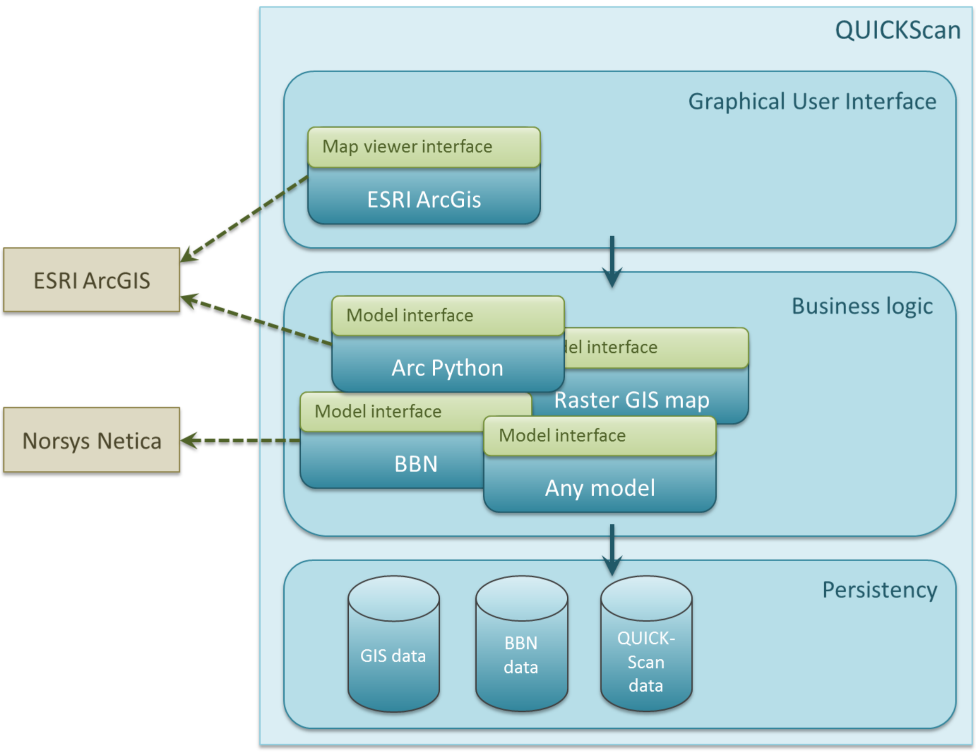



**Figure 1** – Three layered architecture: '*Graphical User Interface*', '*Business logic*' and '*Persistency*' depicting the integration of GIS and Bayesian Belief Networks (BBN) via the standardized '*model interface*'.

The *Graphical User Interface* layer includes the presentation of maps which takes place via the *map viewer interface*, implemented for ESRI ArcGIS, but replaceable by other implementations. The *Business logic* layer contains the definition of a *model interface* for describing and editing the model components, linking them together, triggering calculation and retrieving calculation results in a standardized way. The graphical editor for BBN's is displayed in Figure 2.

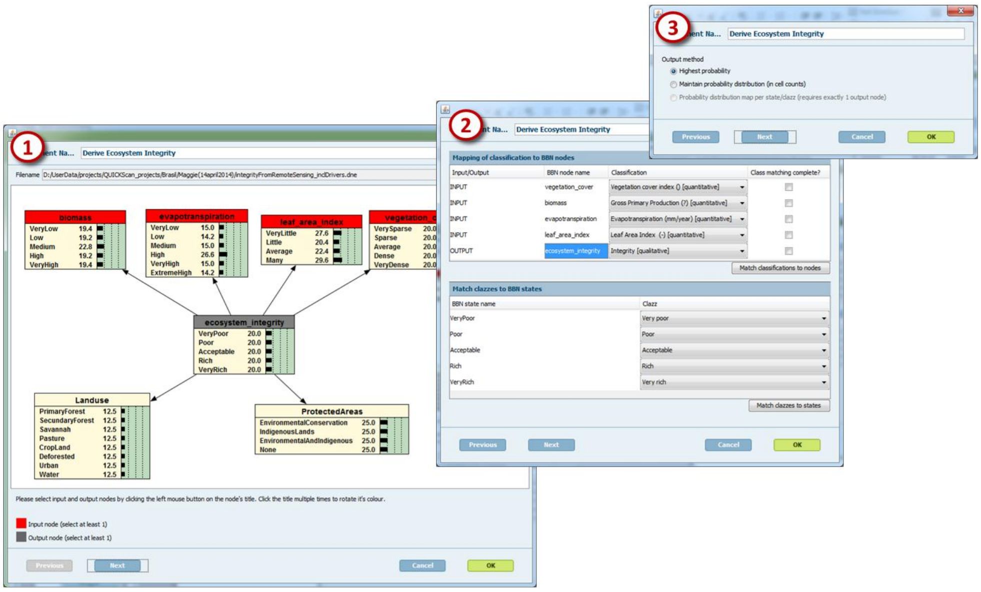



**Figure 2** – Graphical User Interface wizard for linking a BBN in QUICKScan: (1) name the component '*Derive Ecosystem Integrity*', browse to the externally created BBN file and view it, indicate input (displayed in red) and output nodes (displayed in grey); (2) map metrics to nodes and states and; (3) indicate result map calculation method.

### 3.2 Ecosystem Integrity metric, BBN and knowledge rule

The experts decided to do an initial mapping of EI using a qualitative metric, from very poor (displayed as dark red), via acceptable to very rich (displayed as dark green). The different alternatives use the same metric to be able to do quantitative comparisons.

Within the BBN the Ecosystem Integrity is based on four derived Remote Sensing products: biomass, evapotranspiration, the leaf area index and the gross primary product. The last one is used as proxy for vegetation cover. Figure 3 shows the classification of the numeric Remote Sensing products into categories as used within the BBN (see Figure 4). The probability tables of the BBN are filled by expert knowledge

| Node | #States | Description of states |
|---|---|---|
| Biomass | 5 | Numeric: 0-1000. <150; 150-300; 300-450; 450-600; >600 |
| Evapotranspiration | 6 | Numeric: 0 – 1800 [mm/year]. <800; 800-1000; 1000-1200; 1200-1400; 1400-1600; 1600-1800; >1800 |
| Leaf area index | 4 | Numeric: 0-100. <17; 17-32; 32-45; >45 |
| Vegetation cover | 5 | Numeric: 0-100 [%]. <20; 20-40; 40-60; 60-80; >80 |
| Land use | 8 | Primary forest, secondary forest, savannah, pasture, crop land, deforested, urban, water |
| Protected areas | 4 | Environmental conservation units, indigenous lands, both env. Consrv. & indigenous |

**Figure 3** – BBN nodes, states and GIS data classification

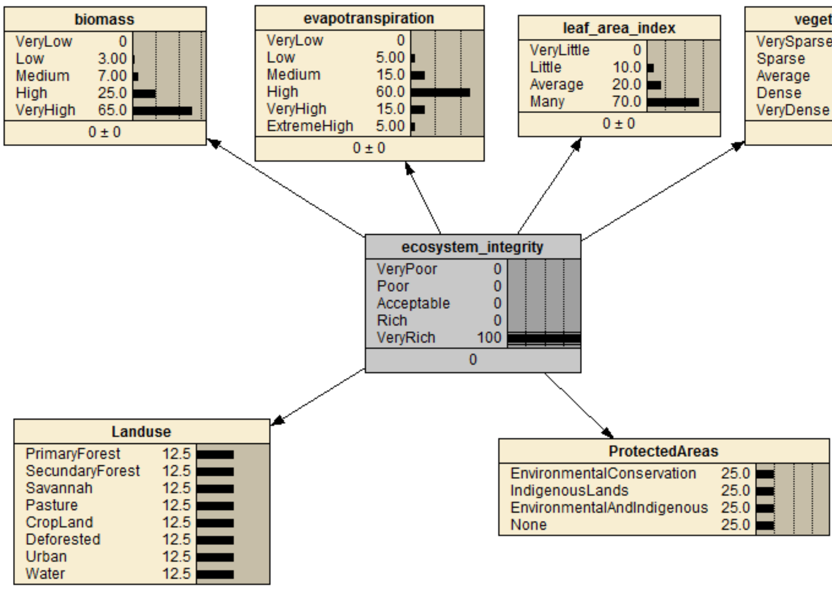



**Figure 4** – Bayesian Belief Network defining the ecosystem integrity on *biomass, evapotranspiration, leaf area index* and *vegetation cover* (top 4 nodes) and on *land use* and *protected areas* (bottom 2 nodes). The bottom nodes implicitly include management

During BBN development the question arose how the EI result map by applying BBN's would differ from using a mechanistic knowledge rule. It was decided to build such a knowledge rule by using the same experts and add it as alternative. The knowledge rule is based on *evapotranspiration* and *vegetation cover* which the experts discussed as being the two most important variables to define the ecosystem integrity. The experts discussed each unique combination and appointed an ecosystem integrity category to it (see Figure 5).

| | Very sparse | Sparse | Average | Dense | Very dense |
|---|---|---|---|---|---|
| Very low | Very poor | Poor | Acceptable | Rich | Very rich |
| Low | Very poor | Poor | Acceptable | Rich | Very rich |
| Medium | Poor | Poor | Acceptable | Rich | Very rich |
| High | Very poor | Poor | Acceptable | Very rich | Very rich |
| Very high | Poor | Acceptable | Acceptable | Rich | Rich |
| Extreme high | Poor | Acceptable | Rich | Rich | Acceptable |

**Figure 5** – Knowledge rule defining ecosystem integrity based on *evapotranspiration* categories (displayed on the row headers) to *vegetation cover* categories (displayed on the column headers)

### 3.3 Ecosystem Integrity maps

Three resulting Ecosystem Integrity maps based on the three alternative approaches (see Figure 6) were created. The areal distribution per EI category is quite similar in all approaches. Especially in the alternatives '*sample from probability distribution*' and '*knowledge rule*' the areal distribution is almost exactly the same. However, the spatial pattern is different. Where '*knowledge rule*' tends to cluster categories, '*sample from probability distribution*' is shows a more scattered image.

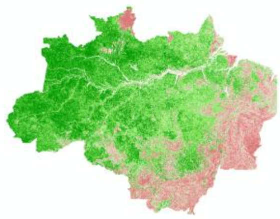
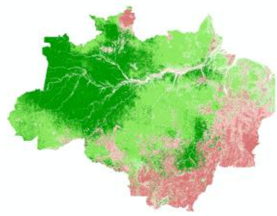
| | **BBN – Highest probability** | **BBN – sample from probability distribution** | **Knowledge rule** |
|---|---|---|---|
| |  |  |  |
| Very poor | 30 % | 30 % | 30 % |
| Poor | 37 % | 34 % | 33 % |
| Acceptable | 8 % | 11 % | 11 % |
| Rich | 9 % | 12 % | 12 % |
| Very rich | 16 % | 12 % | 13 % |

**Figure 6** – Ecosystem integrity maps of each alternative with occupied area per ecosystem integrity category

### 3.4 Data preparation and discussion support

QUICKScan requires GIS data to use the same projection, spatial extent and spatial resolution. Harmonizing data was done using ESRI ArcGIS[2] in a preparatory phase.

The ability to do multiple iterations of modeling, calculation and result analyses via a variety of result visualizations triggered the exploration of different modeling approaches. However, when the group process took place at different locations via tele-conferencing software (i.e. Brazil and the Netherlands) the inter-activeness was hampered by technical causing us to revert into theoretical conceptual and methodological discussions.

## 4. DISCUSSION AND CONCLUSIONS

### 4.1 Technical activities versus content discussions

Without technical integration of BBN and GIS, data had to be manipulated manually by the following sequence: export various GIS data files, manual transformation of the export format into BBN acceptable format, manual import into BBN, run the BBN, export the BBN data, transform the data into GIS import format, import the data into GIS after which it could be displayed for interpretation. A time consuming and error prone process leaving less time for content discussions and with a risk of losing participant interest in the process. Still, data preparation –and especially scale harmonisation- is a time consuming activity. It is now excluded from the iterative modelling activity.

[2] http://www.esri.com/software/arcgis

### 4.2 Interpretability of alternative Ecosystem Integrity maps

The EI map from '*BBN highest probability*' does not correctly represent the underlying probability distribution. The alternative '*BBN sample from probability distribution*' does, but will create a different map every time the BBN is inferred to create the map due to its randomized nature of drawing a sample. This might be a problem when discussing the maps with (local) stakeholders and decision makers. The alternative '*knowledge rule*' does not have these problems as it ignores probabilities. Although statistically the resulting EI distribution is almost identical to the second alternative, the spatial distribution of the EI categories is very different.

### 4.3 BBN – GIS – OpenMI integration

Wrapping a BBN within a standardized modelling framework makes all the functionalities of that framework available to the BBN. In this case the framework provides visualisations for maps, (regional) summary graphs, the notions of '*alternative*' and '*indicator'* to compare between alternatives with a set of comparison visualisations and connectivity with other framework models, like the described knowledge rules or other OpenMI models (e.g. Knapen et al., 2013). In general, going beyond a framework is expensive as maintenance can become time consuming. The framework advantages should therefor outweigh possible restrictions.
Very recently Norsys made GeoNetica available which supports the use of raster GIS data similar to the approach described here. It does not have the aforementioned framework advantages.

### 4.4 Looking ahead

This paper describes intermediate results from working with BBN and GIS to define the current and the possible future EI in the Brazilian Amazon. Next steps are:

- Training of the BBN relation between EI and drivers: land use and protection status to be able to infer likely EI changes from future land use and protection delineation;
- Identifying method to visualize the EI together with its probabilities within a single map;
- Include the same Remote Sensing products in the knowledge rule as in the BBN to improve spatial distribution of EI categories
- On site workshop to gain insight from stakeholders and decision makers in dialog with experts. This might be a deliberative participatory approach, but possibly enhanced to include participatory modelling.

## ACKNOWLEDGEMENTS

The authors acknowledge the valuable contributions of many of our EU FP7 ROBIN colleagues. Specifically we would like to thank Octavio Perez Maqueo and Miguel Equihua that developed the initial- and basis for our Ecosystem Integrity BBN and helped with constructive feedback.

## REFERENCES

Ames, D.P., Anselmo, A., 2008, Bayesian Network Integration with GIS, In: Encyclopedia of GIS, Springer 2008, S.Shekhar, H., Xion (Eds.), pp. 39-45

Amthor, J.S. and Baldocchi, D.D., 2001, Terrestrial Higher Plant Respiration and Net Primary Production. Terrestrial Global Productivity, Academic Press, 33-59

Constanza, R., 2012, Ecosystem health and ecological engineering, Ecological engineering, 45, 24-29

vanEupen, M., Cormont, A., in prep., Joint land use and vegetation scenarios: Joint land use and vegetation scenarios covering observation sites as well as regional-wide application under future climate conditions, D2.2.2, EU FP7 ROBIN – role of biodiversity in climate change mitigation, towards better climate mitigation and biodiversity protection in Latin America

Gijsbers, P., Brinkman, R., Gregersen, J., Hummel, S., Westen, S., 2005, OpenMI: Open Modeling Interface. The OpenMI document series: part C The org.OpenMI.standard interface specification (version1.0).

Gret-Regamy, A., Brunner, S., Altwegg, J., bebi, P., 2013, Facing uncertanity in ecosystem services-based resource management, Journal of environmental management, 127, 145-154

Haines-Young, R., 2011, Exploring ecosystem service issues across diverse knowledge domains using Bayesian Belief Networks, Progress in physical geography, 35, 681-699

Janssen, S., Athanasiadis, I., Bezlepkina, I., Knapen, R., Li, H., Perez Dominguez, I., Emilio Rizzoli, A., van Ittersum, M., Linking models for assessing agricultural land use change, Computers and electronics in agriculture, 76(2), 148-160

Jorgensen, S., Müller, F., 2000, Handbook of ecosystem theories and management, CRC press, 600p.

Knapen, R., Janssen, S., Roosenschoon, O., Verweij, P., de Winter, W., Uiterwijk, M., Wien, J, 2013, Evaluating OpenMI as a model integration platform across disciplines, Environmental modelling & software, 39, 274-282

Novak, J.D., 1991, Clarify with concept maps: a tool for students and teachers alike, The science teacher, 58(7), 45-49

Stelzenmuller, V., Lee, J., Garnacho, E., Rogers, S., 2010, Assessment of a Bayesian Belief Network-GIS framework as a practical tool to support marine planning, Marine pollution bulletin, 60, 1743-1754

Stewart, W., 2011, Probability, Markov Chains, Queues, and Simulation: The Mathematical Basis of Performance Modeling. Princeton University Press. p. 105. ISBN 978-1-4008-3281-1.

Veldkamp, A., Fresco, Lo, 1996, CLUE-CR: an integrated multi-scale model to simulate land use change scenarios in Costa Rica, Ecological Modeling, 91

Verburg, P.H., vanBerkel, D., van Doorn, A., van Eupen, M., Heiligenberg, H., 2010, Trajectories of land use change in Europe: a model based exploration of rural futures, Landscape Ecology, 25(2), 217-232

Verweij, P., Knapen, R., de Winter, W., Wien, J., te Roller, J., Sieber, S., Jansen, J., 2010, An IT perspective on integrated environmental modelling: the SIAT case, Ecological modelling, 221, 2167-2176

Verweij, P., Winograd, M., Perez-Soba, M., Knapen, R., van Randen, Y., 2012, QUICKScan: a pragmatic approach to decision support, In: International Environmental Modelling and Software Society (iEMSs), R. Seppelt, A.A. Voinov, S. Lange, D. Bankamp (Eds.),

Watson, D.J., 1947. Comparative physiological studies on the growth of field crops: I. Variation in net assimilation rate and leaf area between species and varieties and within and between years. Annals of Botany, 11: 41-76