



Jul 1st, 12:00 AM

# Raising the Bar – Is evaluating the outcomes of decision and information support tools a bridge too far?

K. B. Matthews

M. Rivington

K. L. Blackstock

K. Buchan

D. G. Miller

Follow this and additional works at: <https://scholarsarchive.byu.edu/iemssconference>

---

Matthews, K. B.; Rivington, M.; Blackstock, K. L.; Buchan, K.; and Miller, D. G., "Raising the Bar – Is evaluating the outcomes of decision and information support tools a bridge too far?" (2008). *International Congress on Environmental Modelling and Software*. 250. <https://scholarsarchive.byu.edu/iemssconference/2008/all/250>

This Event is brought to you for free and open access by the Civil and Environmental Engineering at BYU ScholarsArchive. It has been accepted for inclusion in International Congress on Environmental Modelling and Software by an authorized administrator of BYU ScholarsArchive. For more information, please contact [scholarsarchive@byu.edu](mailto:scholarsarchive@byu.edu), [ellen\\_amatangelo@byu.edu](mailto:ellen_amatangelo@byu.edu).

## **Raising the Bar – Is evaluating the outcomes of decision and information support tools a bridge too far?**

**K.B. Matthews<sup>a</sup>, M. Rivington<sup>a</sup>, K. Blackstock<sup>b</sup>, K. Buchan<sup>a</sup> and D.G. Miller<sup>a</sup>**

<sup>a</sup>*Integrated Land Use Systems Group, Macaulay Institute, Craigiebuckler, Aberdeen, AB158QH, U.K. [k.matthews@macaulay.ac.uk](mailto:k.matthews@macaulay.ac.uk)*

<sup>b</sup>*Socio-Economics Research Group, Macaulay Institute, Craigiebuckler, Aberdeen, AB158QH, U.K.*

**Abstract:** This paper continues a series of reflections on the challenges of developing and deploying decision and information systems (DIST) for environmental management. Our focus is on the additional challenges being posed by funders to evaluate the outcomes (changes in the world beyond the research institute). This is a significant raising of the bar for DIST, particularly when many still struggle to overcome the simple hurdle of being used at all. The paper reflects on the challenge of evaluating outcomes, placing it in the context of conventional analysis of DIST performance (for example validation, software testing and utility assessment). Several particular challenges are identified, the intangibility of many outcomes, the difficulty of receiving credit for post project outcomes, how to disentangle cause and effect for changes that occur, how to decide the relative importance of outcomes and finally a failing to recognise that despite best endeavours research may still be ignored. The paper then presents a simple evaluation process conducted as part of a transdisciplinary research project that uses model based indicators to communicate the consequences of climate change to land managers and generate discussion of likely adaptations to management. The outcomes of the evaluations were that the process provided new information and raised awareness of the specific research being carried out. The project was also successful in changing views on climate change for a majority of attendees; particularly where the existing levels of knowledge were limited. Yet despite the relative success of the evaluation process and the useful lessons learned, designing an evaluation process and interpreting the findings remains a serious challenge. The authors conclude by questioning whether outcome evaluation is a vital requirement for successful DIST development or does it generate expectations that cannot be met?

**Keywords:** DIST, climate change, evaluation, outcomes, transdisciplinary

### **1. INTRODUCTION**

This paper continues the bridging-the-gap metaphor used in previous IEMSS and MODSIM conferences where the focus was on the design-use gap and asks is the evaluation of outcomes a bridge too far?

The desire for the evaluation of outcomes (changes in the world beyond the walls of the research institute) rather than outputs (in the form of peer reviewed articles or software) is increasingly seen within policy-relevant and policy-led strategic and applied research programmes in the U.K. This is partially driven by the perceived failure of research and particularly decision and information support tools (DISTs) to deliver against their “utopian” promises of the 1990’s [Matthews et al., 2007]. The drive for outcome evaluation also reflects funders seeking to assess the cost-effectiveness of their research to provide a justification for future spend as part of the new managerialist paradigm embraced by UK government [Richards and Smith, 2003]. Researchers are locked into this search for evidence through competition for limited research funds. Finally, there is a desire for research to clarify and rationalise the messy confusion of pluralist democratic decision making [Solesbury, 2001] i.e. to provide evidence for which decision should be taken.

This context raises questions about the role of researchers and our acquiescence to such agendas. There is, however, a more positive side to the desire to evaluate outcomes. Recent evaluation literature has

highlighted the importance of understanding the relationship between context, process and outcomes [Blackstock et al., 2007; Patton, 1998]. Therefore, a focus on outcomes requires understanding *how* and *under what conditions* information is interpreted and used by stakeholders – moving away from an ‘information deficit’ model to a model of knowledge exchange [Ekboir, 2003]. So as researchers committed to improving standards of environmental management, planning and policy - how can we make sensible steps towards evaluating outcomes of DISTs whilst managing the expectations of funders of research, decision makers and other stakeholders? This paper argues that the evaluation of the outcomes of using DISTs presents conceptual, resource and institutional challenges. Lessons from social research design can help DIST development teams deal with these challenges of evaluating outcomes.

The paper first reflects on some of the alternative interpretations of evaluation and how these may relate to the particular case of evaluating outcomes from DISTs. The paper presents a case study that evaluates the outcomes of a process communicating to land management stakeholders the consequences of climate change for land use using the outputs from agro-meteorological modelling. The paper concludes by trying to draw some general conclusions for the process of using DISTs.

## 2. RELATED RESEARCH

Several reasonable forms of evaluation coexist in the DIST literature: model validation, software testing and utility assessment. These can be usefully considered in terms of their appropriateness of their methods for outcome evaluation judged.

**Model validation** represents the assessment of the performance of a model against an independently collected dataset. Within an experimentalist and/or reductionist paradigm such systems of validation are effective. Validations of environmental systems models when conducted with stakeholders can also significantly enhance the credibility of the DISTs based on them [Carberry et al., 2002]. Validation based evaluations are weaker where key data cannot be falsified or directly measured (the preferences of stakeholders), the systems of interest are very large and/or very complex (issues of pre-analytical choices, scale and equifinality in parameterisation) or where experimentation for validation would raise practical and ethical concerns and is not permitted [Giampietro, 2004]. The validation based evaluation can also fail to address issues of legitimacy and salience. That is the model may perform well in a given context but be overly specialised, too demanding of input information or fail to address key relationships [McNie, 2007] that are of interest to outcome evaluation.

**Software testing** in this context refers to processes of software quality control – either formal or informal. These range from simple debugging to larger structured processes of software testing with hierarchical, recursive breakdown and testing of software components, units and modules [Britton and Doake, 1996]. These can be undertaken by developers (in smaller projects) or by testing and change management teams for larger systems with ongoing development. Such software testing can be assisted by automated testing to ensure repeatability of results or to benchmark systems [Hutchins et al., 2006]. It represents a potentially significant overhead in DIST development. This can be caricatured as a luxury or as non-productive use of staff time by research managers under pressure to deliver more with less. Model testing in the era of: rapid and open source development; the increasing availability of model-bases and model-engines and the adoption of interoperability standards, has to deal with higher level meta-data and meta-model compatibility issues [Ittersum, 2006]. These issues also concern the credibility of DIST outputs as the technical ability to combine elements does not always infer that there is a justification for doing so. Such evaluations also typically have a quality assurance of outputs focus rather than assessing the outcomes of DIST use.

**Utility Assessment.** Both of the cases above have a credibility focus, which is necessary but not sufficient for the evaluation of DISTs. There remains for many DISTs the thorny issue of problem of implementation, or why does my perfectly implemented DIST not get used by anyone [McCown, 2002]? Utility testing has been extensively debated elsewhere [Matthews et al., 2006b] with the conclusion that issues of salience and legitimacy can undermine the utility of DIST's. Firstly there must exist an adequate and honest partnership between developers of DISTs and relevant stakeholders for developing the specification of DIST functionality but more importantly agreement on the process through which a DIST is used for example as part of deliberative inclusive processes. Processes of utility testing may be experimental testing [Matthews et al., 2006a], formal market research survey [Matthews et al., 2007] or based on soft-systems methods of participatory co-development [Matthews et al., 2006c] depending on

the particular circumstances and the resources available. Whatever the particularities utility testing is essential if the DIST is not to become shelfware.

**Outcome Evaluation**, however, needs to go beyond the utility, validity or structural integrity of a DIST and into assessment of the effects of using that tool within a decision making or management process. Evaluation of outcomes from the strategic rather than routine use of DISTs confronts several taxing issues. First is the intangibility of many outcomes that leads to just measuring what can be measured and calling that success when other outcomes that are more difficult or costly to measure are still very important. For example the value of the social capital bonds between key individuals in a community strengthened through by participation in a DIST development may enhance their adaptive capacity far more than the DIST itself. Secondly, it is also very difficult to attribute any outcomes in complex social systems to any particular interventions [Bellamy et al., 2001]. It is difficult to establish causality in real world research when processes and participants can not be directly replicated or controlled [Robson, 1993]. Attribution of belief that a change is stimulated by an intervention *is* possible, and this can be easier in newer fields where there are fewer interventions to complicate these attributions. Even where outcomes can be distinguished then is any success simply happenstance rather than predicated on the actions of the research team? Thirdly, the long term and cumulative nature of change may also be incompatible with short-term, project-based research and development funding schemes. The project team may have moved on to new research long before the outcomes can be measured [Blackstock et al., 2007]. Funders are also less keen to fund monitoring and evaluation of existing tools rather than being seen as the initiator of a new development. Fourthly even where outcomes can be distinguished in some way then there will be considerable disagreement on the relative importance of individual outcomes, particularly when these outcomes are non-commensurable. In such cases there is too often the resort to crude, *ad hoc*, indicator sets that simply pass on to a third party the difficulty of interpreting of intimately interconnected and highly nuanced systems-level outcomes. And the politics of knowledge suggest that outcomes other than those expected (or even desired) by the DIST funder will be discounted in favour of those that ‘fit’ [Kirk et al., 2007]. Finally there needs to be a recognition of the limits on science generated “expert” evidence within a plurality of expertise derived in different ways [Stilgoe et al., 2006]. Failure to recognise the legitimacy of alternative points of view and engage with deliberative processes can mean that otherwise sound science being ignored as an arrogant attempt at scientism. Assessment of the outcomes of research-based tools also need to recognise that research and research based tools such as DISTs are only one of the sources of influence and probably not one of the more important [Solesbury, 2001]. This is not to discourage DIST developers from contributing to policy and management decision making but expectations must be managed if we are to avoid a further “disappointment” crash as happened with DSS in the late 1990’s.

For the authors there is no preconceived preference for evaluation methods since both quantitative and qualitative approaches can be equally rigorous or specious depending on implementation. Instead, our research team has tried to chart a pragmatic path to evaluation that addresses each of the issues of credibility, legitimacy and salience. Such a mixed-methods based approach to evaluation is presented in the following sections.

### 3. MATERIALS AND METHODS

#### 3.1. Climate data and modelling

The underpinning research for the C4LU project has its origins in our research interest in the consequences of using alternative sources of data meteorological data in the cropping systems components of the Land Allocations Decision Support System (LADSS). Testing in this situation was conventional experimental analysis looking at the impacts of using off-site or modelled on-site data when measurements for particular variables were not available. The is process quantified the uncertainties introduced and the reliability of the DSS results using expert weighted performance metrics combined in a fuzzy logic framework [Rivington et al., 2005; Rivington et al., 2006]

Interest in using for climate change scenario data from UKCIP02 led to a desire to test the effectiveness of the Hadley Centre Regional Climate Model (HadRM3) in replicating historic conditions. Testing these hindcasts allowed the development of simple downscaling methods to correct for representational and other biases within the HadRM3 data for particular sites to increase the confidence in the data for use at specific case study locations. These monthly correction factors were also applied to the future scenarios on the assumption that the representational bias would still be present and other biases were likely to

remain. Here the testing was conventional statistical testing with the outcomes published in peer reviewed journals [Rivington et al., 2007].

### **3.2. Agro-meteorological metrics**

In parallel with the climate data and modelling a project developing agro-meteorological metrics was undertaken with selected stakeholders. This project has been reported elsewhere [Matthews et al., 2008] but in this context the key features were that it tested the utility of existing agro-meteorological metrics as a basis for first for characterising future climate change scenarios and secondarily for stimulating deliberation on possible adaptive strategies that could be adopted by land managers.

The indicator set, the presentation methods and the format of the interactions with stakeholders were all piloted with leaders of the stakeholder organisations and then tested and refined (to a smaller set of key indicators) with stakeholder groups (total participants = 40). A small workshop format (8-16 participants lasting up to 3 hours) has been used with round-table presentation of outcomes from the modelling, deliberation on the significance of the results and the utility of the indicators combined with qualitative information gathering on adaptive strategies and a wider discussion or debate of related issues. Each workshop has thrown up additional refinements to either the process or the elements presented and have seen increasing demands for more sophisticated analyses. The latter result is particularly important since it indicated the research team are gaining skill in communicating the necessarily complex messages and when the means of presentation have been improved stakeholders seek more information.

### **3.3. Evaluating the outcomes**

The team were asked to evaluate the outcomes of the most recent round of (4) workshops in February and March 2008 by the research funders. These are not a usual research funder but a science communication group within the Scottish Government tasked with increasing the interactions between active researchers and the wider publics. Their interests and those of the research team coincided in the communication of climate change consequences but their evaluation of success could not be in terms of peer reviewed papers. We were also anxious not to fall into the trap of judging success by the easily measured outputs, e.g. numbers of people attending. This would compromise our participatory research approaches by simply increasing the number of people at the expense of the quality of interactions, whilst not necessarily improving outcomes (increased awareness and understandings).

The attempt was made to evaluate the outcomes of the workshops at two levels. The first was adding a simple evaluation form to the materials used within the workshop. The second was to record the meetings to allow a subsequent analysis of how participants responded to the information in discussion. The former is presented within this paper while the latter has still to be undertaken (reflecting the balance of effort required for the two approaches).

The survey analysis asked five simple questions with tick box answers (options shown in brackets).

- Q1. Did you know about the Macaulay Institute's research before you attended the workshop? (Y/N)
- Q2. How much did you know about climate change before you attended the workshop? (Options)
- Q3. Has the workshop provided new information on the topic? (Y/N)
- Q4. Have you changed or adapted your views on climate change after the workshop? (Y/N)
- Q5. Where do you work? (Academic Institution, Government, NGO, Private Sector, Other)

These questions sought to assess how effective the deliberative workshop process is 1) to raise awareness, 2) to provide new information and 3) to influence attitudes. The research team did not attempt to measure changes in participants' practices, given that literature on barriers to changing farmers' behaviour [Burton et al., 2007; Kaljonen, 2006] suggests it would be impossible to rigorously assess the degree to which the DIST alone was responsible for change (or the lack thereof). The results of the four workshops are presented below but represent just a first step into this issue for a DIST research team.

## **4. RESULTS AND DISCUSSION**

### **4.1. Utility Analyses**

The overall responses on the utility of the 15 indicators are shown in Figure 1. The previous processes have been successful in largely eliminating indicators that are not seen as relevant and the top two categories (Very and Quite) account for 58% of the responses. A middle category was deliberately not

provided to avoid “neutral responses” but some stakeholders felt the four categories were insufficiently fine grained. An enterprise specific category that allowed delegates to highlight indicators that were important but only for a narrower set of circumstances or activities was suggested.

The breakdown for each of the metrics is summarised in Figure 2, with the counts of the Very and Quite useful responses shown. This highlights the importance of growing season and access periods, but particularly the increasing view that metrics showing the monthly distributions of phenomena are useful (growing days, dry soil days, access days or their combinations). The continuing importance of climate metrics in the survey does not match well with the comments within the later discussions, suggesting the higher score for these indicators may be the result of their appearance first in the workshops (the delegates having nothing to compare the indicators with at that stage). This possible artefact highlights the need for extreme care in designing and interpreting evaluation outcomes.

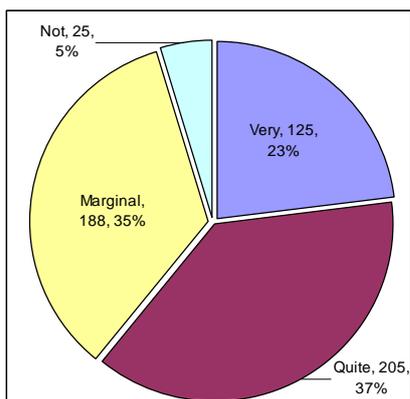


Figure 1 - Utility of the C4LU Indicators

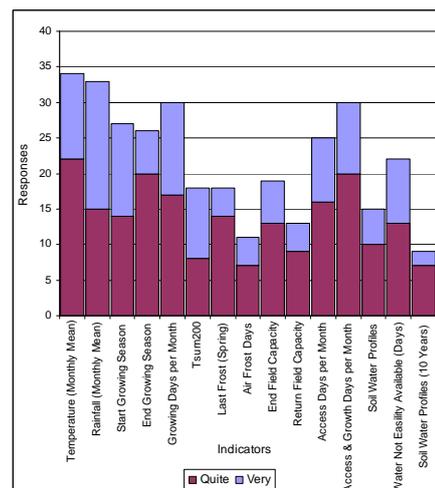


Figure 2 - Responses per Indicator

Figure 2 also shows the preference for the compound indicator access during the growing season. This indicator and the way in which it is presented was the outcome of earlier rounds of workshops. The benefits of processes that can encompass aspects of learning are thus evident. However, such flexibility and process evolution can conflict with the formalisms of conventional science where replication is required. Such methodological approaches may inadvertently constrain the ability to achieve the outcomes sought (increased awareness, knowledge and adaptation) by stifling adaptive participatory processes.

#### 4.2. Outcome Evaluations

The responses from the C4LU workshops are summarised in Figure 3. These show that the participants were predominantly from the private sector with smaller numbers from the government, NGO and academic sectors. Of the private sector participants all were actively involved in land management decision making either in farming, forestry, fishing/hunting or in related activities such as tourism and the agri-food supply chain. In many cases the participants ran diversified businesses and thus had a broad view of potential climate change impacts. There was a significant representation of agri-environmental interests so the discourse was not dominated by productivism, though food security issues were frequently raised.

In terms of awareness of the Institute’s research the workshops were marginally effective (5/40), but this reflects the Institute’s previous extension role. The participants are aware that research is carried out but the topics of research (even at a headline level such as climate change) are poorly understood or are those of earlier programmes. The workshops thus represent a significant opportunity to do awareness raising for particular projects and the wider programmes of research, and this is vital if there is to be continued support for public-good funded research.

In terms of expertise all participants had at least a little knowledge (perhaps reflecting the efforts in the mass media and trade journals to inform on the issues of climate change). There were more participants classifying themselves as having a fair knowledge (and the lack of any classifying themselves as experts

may be due to modesty – considering how well informed many of the participants were). There was a fair degree of variation in the level of reported knowledge – with the Golspie case study (Figure 4) reporting lower levels of knowledge compared with for example Aviemore (Figure 5). This perhaps reflects the backgrounds and interests of the land managers in Golspie where conventional mixed-farming systems dominate rather than the Aviemore land managers who were strongly diversified enterprises and with a strong agri-environmental representation.

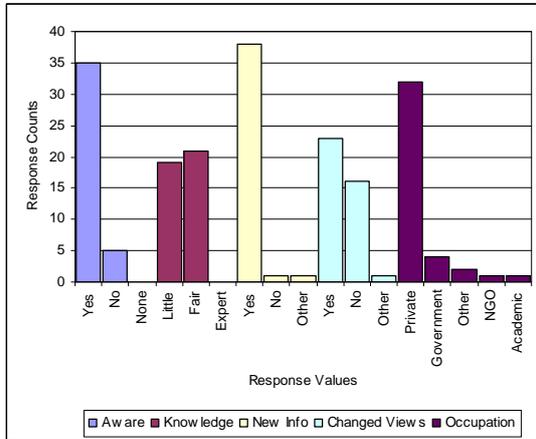


Figure 3 – Overall C4LU outcomes

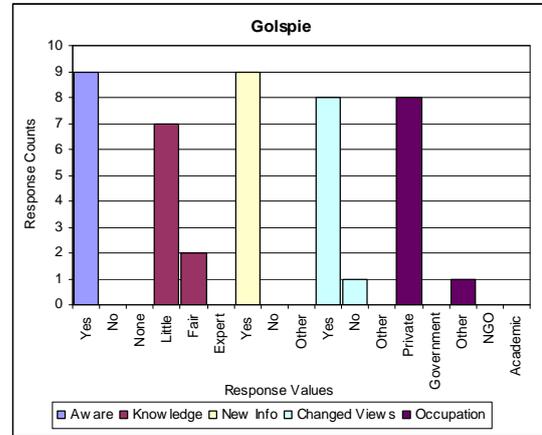


Figure 4 - Golspie C4LU responses

In almost all cases (38/40) the workshops were successful in providing new information. The two cases where this was not the case were government and academic colleagues who were participating in the workshops as participants but also providing expert interpretations of the implications of the metrics for particular sectors (e.g. forestry). The new information evaluation is a key baseline for the evaluation since this is the conventional role for scientific research – failure to be providing new information would raise serious questions of whether the research was relevant to the decision makers at all. The usefulness of the new information is assessed as outlined in Section 4.1 (above).

The acid test for the workshops was in changing or adapting views on climate change. In a majority of cases (23/40) the participants had changed their views on climate change. (Further analysis of the discussion may also indicate if those who have not changed their view already felt adapting to climate change was important). This was interesting as previous research had found that climate change was not seen by land managers as a priority for action or one for which they had much adaptive capacity [Scottish Executive, 2006]. The message of the workshops was that mitigation efforts will be important but that change in climate will have to be planned for and adapted to since a certain degree of change has already been entrained by previous emissions. The authors think this message was successfully delivered through the collective evaluation of the indicators than conventional research presentations.

The survey, however, only sought to evaluate if peoples views had changed, not *how* they had changed. This reflected a desire to keep the form as simple as possible and not to alienate participants by use of a complex questionnaire. The questionnaire is backed up by audio-recording of the workshop, which can be qualitatively analysed to bring out these more nuanced questions. The multi-method approach avoided a more explicit focus on changes in attitude within the workshop, which could be quite confronting in a group based process.

By combining the utility and evaluation datasets the authors attempted to develop a “profile” of the participants to explore if there was a strong link between the reported utility of the indicators and whether workshop participants has altered their views. Figure 6 shows each participant’s judgement of the utility of each indicator, where the participants are sorted into those who did or did not change their views. There is no clear difference between the two profiles of the change and no-change groups. This suggests that individual actors have a complex set of values, interests and existing knowledge that strongly modify their responses to any research based intervention. It is interesting to note that the case-study with the greatest rate of change in view was Golspie (Figure 4) where the level of knowledge was lowest. This perhaps reflects the previous view (Section 2) that it is easiest to make the biggest impact in areas (in this case geographic but also by extension for subjects) where the knowledge base is less well developed and where perhaps actors have less entrenched views. The inference from this finding is perhaps that as researchers we need to be engaging with key actors and decision makers as early as possible rather than

trying to reshape a debate that has already “solidified” such that research outcomes are ignored rather than incorporated. Conversely, the failure of a DIST to change practices may be due to entrenched views not a poor tool.

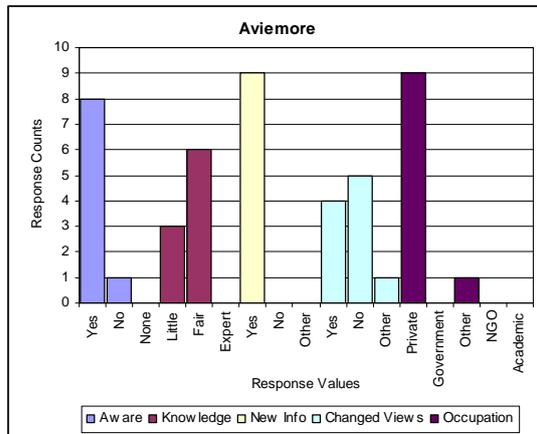


Figure 5 - Aviemore C4LU responses

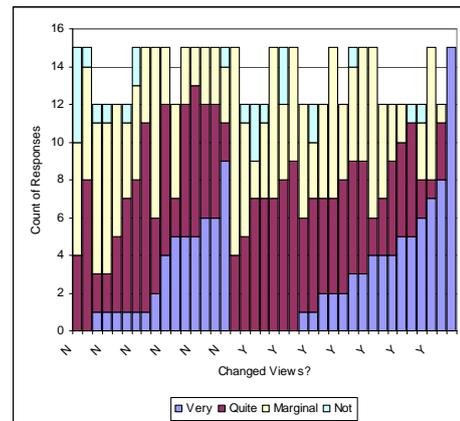


Figure 6 - Profile of utility of indicators responses

## 5. CONCLUSIONS

In this paper the authors have identified a further challenge for DIST developers: the evaluation of the outcomes of the use of their tools. The authors have reflected on the justification for such requirements and on the limited capacity of most DIST research team to deliver rigorous evaluations rather than easily measured yet largely meaningless output metrics. The authors have presented an example of evaluation used as part of a transdisciplinary research project communicating the climate change consequence for land managers and seeking to simulate deliberation on possible adaptive responses.

While it has been possible to generate simple but informative headline evaluations of the C4LU process outcomes, it is clear that the design and interpretation of such evaluation processes requires great care. These data open up discussions about potential changes in attitudes but illustrate the outcomes are hard to assess without exploring the participant’s individual context and the influence of the deliberative process. While the research team developing the C4LU project are experienced in both model development and managing social learning processes, the additional requirements for evaluating the outcomes of the project still suggests this is a gap that may be too broad to bridge. Transdisciplinary research is already demanding of skills and resources within constrained budgets, yet this further demand is often introduced without recognition of the additional resource implications.

The evaluation of DIST outcomes is a challenge that needs to be met since evaluation and reflection in effect closes the loop that starts from initiation, and continues through design, build, and testing. How best to organise the evaluation of DIST? Should the teams developing them do the evaluation, since they are most familiar with the tools and can most quickly respond to feedback with improvements? Should the DISTs be evaluated by external teams with the skills to undertake the task? Is there the need for partnership, but if so how do we resolve the power relationships between a technologically skilled development team and evaluators with a different skill set and priorities? In part this requires engaging funders in a discussion about the meaning of evaluation metrics and debating what are the outcomes that DISTs can reasonably be expected to achieve. In the authors view a considered evaluation of the outcomes of DISTs could be the key to realising their potential in environmental management. Yet outcome evaluation may prove a bridge too far if we fail to control inflated or utopian expectations of science based tools and interventions?

## REFERENCES

Bellamy, J. A., Walker, D. H., McDonald, G. T., and Syme, G. J., A systems approach to the evaluation of natural resource management initiatives, *Journal of Environmental Management* 63(4), 407-423, 2001.

Blackstock, K. L., Kelly, G. J., and Horsey, B. L., Developing and applying a framework to evaluate participatory research for sustainability, *Ecological Economics* 60(4), 726-742, 2007.

- Britton, C. and Doake, J., *Software System Development, A Gentle Introduction*, McGraw-Hill Publishing Company, London, 1996.
- Burton, R. J. F., Dwyer, J., Blackstock, K. L., Ingram, J., Brown, K. M., Mills, J., Schwarz, G., Matthews, K. B., and Slee, B., *Influencing positive environmental behaviour among farmers and landowners - a literature review for DEFRA*, Report, Macaulay Institute, 2007.
- Carberry, P. S., Hochman, Z., McCown, R. L., Dalgliesh, N. P., Foale, M. A., Poulton, J. N. G. H., Hargreaves, D. M. G., Cawthray, S., and Hillcoat, N., The FARMSCAPE approach to decision support: farmers', advisers, researchers' monitoring, simulation, communication and performance evaluation, *Agricultural Systems* 74, 141-177, 2002.
- Ekboir, J., Why impact analysis should not be used for research evaluation and what the alternatives are, *Agricultural Systems* 78, 166-184, 2003.
- Giampietro, M., *Multi-scale integrated analysis of agroecosystems*, CRC Press, Boca Raton, Florida., 2004.
- Hutchins, M. G., Urama, K., Penning, E., Icke, J., Dilks, C., Bakken, T., Perrin, C., Saloranta, T., Candela, L., and Kamari, J., The model evaluation tool: guidance for applying benchmark criteria for models to be used in river basin management., *Archiv fur Hydrobiologie - Large Rivers Supplement* 2006.
- Ittersum, M. K. v., Integrated assessment of agriculture and environmental policies: towards a computerised framework for the EU (SEAMLESS-IF), In: Voinov, A, (ed.), *3rd Biennial Meeting - Summit on Environmental Modelling and Software*, 2006.
- Kaljonen, M., Co-construction of agency and environmental management. The case of agri-environmental policy implementation at Finnish farms, *Journal of Rural Studies* 22(2), 205-216, 2006.
- Kirk, E. A., Reeves, A. D., and Blackstock, K. L., Understanding environmental regulation: disciplinary interactions and path dependency, *Environment and Planning C*. 25, 250-268, 2007.
- Matthews, K. B., Buchan, K., Sibbald, A. R., and Craw, S., Combining deliberative and computer-based methods for multi-objective land-use planning, *Agricultural Systems* 87, 18-37, 2006a.
- Matthews, K. B., Hutchins, M., and Hill, G., Bridging the design-use gap for DSS in environmental policy and practice., In: Voinov, A, (ed.), *Summit on Environmental Modelling and Software - 3rd Biennial meeting of the International Environmental Modelling and Software Society (iEMSs)*, 2006b.
- Matthews, K. B., Rivington, M., Buchan, K., Miller, D. G., and Bellocchi, G., Characterising and communicating the agro-meteorological implications of climate change scenarios to land management stakeholders., *Climate Research* 2008.
- Matthews, K. B., Schwarz, G., Buchan, K., Rivington, M., and Miller, D. G., Wither agricultural DSS?, *Computers and Electronics in Agriculture* 61, 149-159, 2007.
- Matthews, K. B., Wright, I. A., Buchan, K., and Davies, D. A., Assessing the options for upland livestock systems under CAP reform: developing and applying a livestock systems model within whole-farm systems analysis, *Agricultural Systems* 90, 32-61, 2006c.
- McCown, R. L., Changing systems for supporting farmers' decisions: problems, paradigms and prospects, *Agricultural Systems* 74(1), 179-220, 2002.
- McNie, E. C., Reconciling the supply of scientific information with user demands: an analysis of the problem and a review of the literature., *Environmental Science and Policy* 10, 17-38, 2007.
- Patton, M. Q., *Utilization-focussed evaluation*, Sage, Thousand Oaks, CA, 1998.
- Richards, D. and Smith, M. J., *Governance and Public Policy in the UK*, Oxford University Press, 312 pp., Oxford, 2003.
- Rivington, M., Bellocchi, G., Matthews, K. B., and Buchan, K., Evaluation of three model estimations of solar radiation at 24 UK stations, *Agricultural and Forest Meteorology* 132, 228-243, 2005.
- Rivington, M., Matthews, K. B., and Bellocchi, G., Evaluating uncertainty introduced to process-based simulation model estimates by alternative sources of meteorological data, *Agricultural Systems* 88, 451-471, 2006.
- Rivington, M., Miller, D. G., Matthews, K. B., Russell, G., and Buchan, K., Evaluating Regional Climate Model estimates against site-specific observed data in the UK., *Climatic Change* 2007.
- Robson, C., *Real World Research: A Resource for Social Scientists and Practitioner-researchers*, Blackwell Publishers, Oxford, 1993.
- Scottish Executive, *A Forward Strategy for Scottish Agriculture: Next Steps*, Report, Natural Scotland, 32pp., 2006.
- Solesbury, W., *Evidence Based policy: Whence it came and where it's going*, Report, UK Centre for Evidence Based Policy and Practice, 2001.
- Stilgoe, J., Irwin, A., and Jones, K., *The received wisdom. Opening up expert advice*, DEMOS, London, 2006.