

Evaluation of environmental decision and information support tools: from adoption to outcome

B.S. McIntosh^{a, b}

^a *International WaterCentre, Brisbane, Australia*

^b *Smart Water Research Centre, Gold Coast, Australia*

Email: b.mcintosh@watercentre.org

Abstract: The value of environmental decision and information support tools (DISTs) and technologies is located in the impacts that the use of such technologies have on the behaviour of individuals and organisations, and consequently on our collective ability to negotiate the difficult interfaces between human social and economic activity, and environmental impact. There are significant, but as yet insufficiently exploited opportunities for improving the value of these tools and technologies through the use of assessment and evaluation practices.

Iterative and evolutionary tool development processes offer the opportunity to deliberately incorporate user assessment information into design. Doing so can improve adoption and use impacts through better tailoring of tool functionality, output and interfaces to user needs. Managing such processes successfully is demanding and we do not possess sufficient, generalised or published practical understanding of what works and what does not in terms of user interaction. However, there is an emerging understanding of how environmental DIST teams should be structured and a wide body of knowledge yet to be imported from the software, product and interaction design communities, so the future is promising.

The evaluation of the impact of using environmental DISTs is more problematic and less well understood, both methodologically and empirically. Our conceptual frameworks for characterising use and identifying where impacts arising from tool use might lie are nascent and require refinement, both to reflect the variety of developmental routes and eventual end-use destinations for those tools, and the complicated chain of influences amongst which the use of DISTs is only one on decision processes, be they individual or organisational in scale and character. Further, our empirical understanding of the range of impact types which might occur at either individual or organisational scales across different stages of the life cycle of an environmental DIST is in need of development both conceptually and in measurement terms.

This paper presents a critical discussion of our conceptual and empirical understanding of how to use evaluation and assessment practices to improve the value of environmental decision and information support tools. Matthews *et al.* (2011)¹ framework for understanding the life cycle of DISTs from to basic science to use within wider societal governance and media processes is used as the basis for identifying points where additional evaluation and assessment might offer benefits. Difficulties in evaluating the ultimate outcomes of using DISTs are subject to critique, and the current lack of an accepted framework for conceptualising organisations and the kinds of impacts that DISTs might have within them discussed. What we know about how to conduct evaluation and assessment at each point is briefly reviewed alongside what we know empirically about tool impacts. The paper ends with some recommendations regarding both the framing and measurement of impact across the DIST lifecycle from adoption to outcome.

Keywords: *Decision support; Information support; Evaluation; Adoption; Impact*

¹ Matthews, K.B., Rivington, M., Blackstock, K., McGrum, G., Buchan, K. and Miller, D.G. (2011). Raising the bar? – The challenges of evaluating the outcomes of environmental modelling and software. *Environmental Modelling and Software* 26:247-257.

1. INTRODUCTION

How can we ensure and know that the model-based tools we produce to provide support to environmental decision formulating and making processes are effective? How can we act to ensure that those tools and the underpinning mathematical, computational and software technologies are used, useful and help us collectively as a species, and more specifically help organisations and agencies with policy and management responsibilities, navigate a successful course around the interfaces between economy, society and the environment?

With the population of the world due to grow from around 7 billion to 9.3 billion by 2050 (United Nations, 2010) it is clear that the need to generate and implement effective policy change and direction is pressing. It is also clear that environmental decision and information support tools (DISTs) (see Diez and McIntosh, 2011 for a definition of the term) from GIS through to sophisticated integrated assessment models and tools are used, and have the potential to positively influence policy and management processes and outcomes (Diez and McIntosh, 2011; Nilsson *et al.*, 2008; van Delden *et al.*, 2011). What is less clear is the extent to which those tools are as effective as they could be, to what extent they are effective at all and what effects they have in the first place (McIntosh *et al.*, accepted).

Our knowledge about these gaps can be filled through the use of assessment and evaluation practices, both to inform and improve DISTs, and to contribute to a more robust understanding of their role in the complex and ever-changing world of environmental policy and management. The objectives of this paper are to:

- Identify a range of points across the development and use of a DIST that could benefit from more focussed use of assessment and evaluation;
- Critically review what we currently know about the kinds of impacts and outcomes that DISTs have;
- Critically discuss ways in which assessment and evaluation might benefit each life cycle point, and the overall impacts and outcomes of DISTs.

To enable this discussion it will first be necessary to provide a conceptual framework which articulates the meaning of impact and outcome, and links these concepts to the life cycle of a DIST from development through adoption to use. The focus of this paper will be on the use of DISTs by organisations in particular, which the reader should bear in mind.

2. UNDERSTANDING IMPACT AND OUTCOME

How are DISTs developed? How do they relate to the natural, social and computational science which underpins them? How are they used and what kinds of impacts do they have? Can, and how can, such impacts be measured?

Nilsson *et al.* (2008) present evidence demonstrating the use of a range of policy assessment tools across European countries, including some model-based tools that may reasonably be termed DISTs, and argue that it is often difficult to tease out the impact of those tools from the complex web of influences which impact on policy processes over time. In a similar vein, Matthews *et al.* (2011) argue that whilst it is possible to evaluate the outcomes of processes which underpin the delivery of broader social, economic or environmental outcomes, unpicking the contribution of DISTs to the complex web of influences on the achievement of those broader outcomes is significantly more difficult, and may be impossible.

Considering the notions of impact and outcome as they relate to DISTs requires a conceptualisation of how such tools are generated and how they relate to both science and policy. Impact and outcome can only really be defined and discussed in the context of such a conceptual framework. Building on the 'consultancy' model of McCown (2002), Matthews *et al.* (2011) provide a framework for doing so (see Figure 1). The framework itself can be viewed as description of the overall process of knowledge transfer from science to policy, with DISTs being the product of cyclical development processes, which themselves are products of cyclical science knowledge generation processes, both embedded within larger cyclical governance and social processes which determine funding patterns and use opportunities.

In the Matthews *et al.* (2011) framework environmental modelling might play a role in the research cycle, through helping to generate and test theories but such modelling is not intended to be policy-relevant (Oxley *et al.*, 2004). It is not until the development cycle comes into play that modelling technologies become focussed on answering policy or management questions, on informing decision-making and action taking i.e. take the form of DISTs. Matthews *et al.* (2011) identify assessment and evaluation opportunities across this set of inter-linked processes as occurring through peer review (research cycle); through a set of processes

which can inform the development and influence the usefulness of DISTs (validation, reliability, usability, interpretation and relevance); through the periodic evaluation of broader social, economic and environmental outcomes by governmental and non-governmental agencies, and; through subsequent periodic reviews of development and research priorities by those agencies.

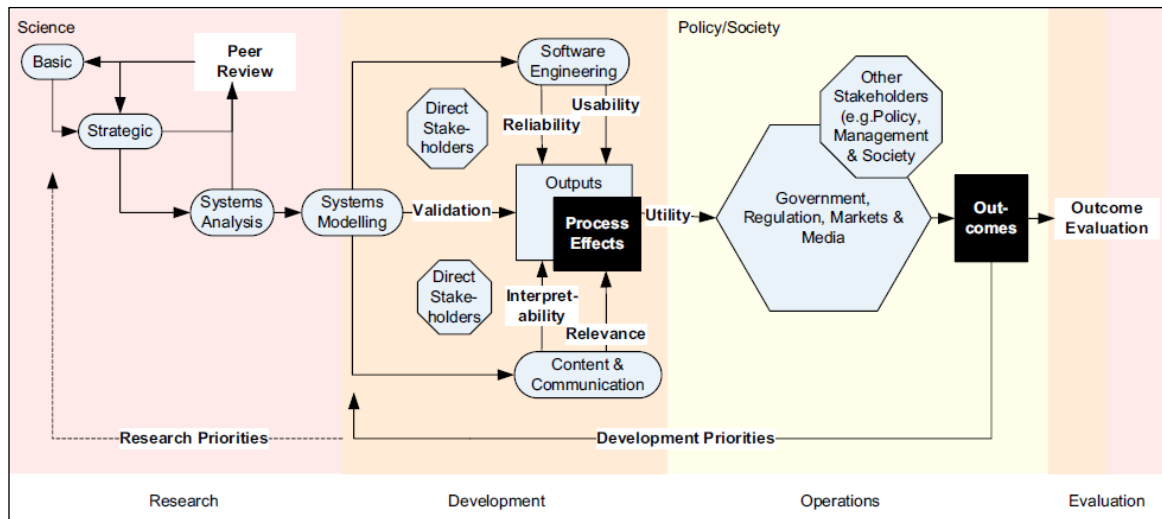


Figure 1. A framework for linking decision and information support tools (see the development cycle) to science (see research cycle) and societal use (see operations cycle) (from Matthews et al., 2011)

We won't consider peer review evaluation processes within the research cycle any further here, nor processes of research or development priority evaluation. Instead we will focus on assessment and evaluation processes within the development cycle, and outcome evaluation processes.

Matthews *et al.* (2011) argue that the main class of impacts and outcomes which in principle it is possible to evaluate are those which are labelled 'process effects' – the set of impacts and outcomes which arise as a consequence of developing and using DISTs individually or in organisations. Outcomes in the sense of improvements to social well-being, economic success or environmental change cannot be evaluated in the same way for they are less easily measurable, and where they are may be subject to contestation. Further, being able to attribute the causes of any such outcome to the use of one or a set of DISTs is likely to be difficult or impossible because of the time lagged nature of action effects on larger scale social, economic or environmental systems, and because of the sheer complexity of the web of causal influences on those outcomes.

Using an example of a communicating climate change project, Matthews et al. (2011) demonstrate that outcomes are not necessarily related to process effects (or outputs) from DIST use. They found that a range of outcomes were identified such as participants changing their attitudes towards climate change and enhanced discussion and awareness of adaptation strategies, but that those outcomes were not necessarily easily relatable to the measurable process effects due to the complexity of the social processes involved. That is, participants who rated the utility of the DIST outputs they employed highly were not necessarily those who changed their attitudes and behaviours.

In one sense Matthews et al. (2011) are correct, and the framework they developed is useful as the only extant (to the author's best knowledge) description of the relationship of science - DIST development and use – broader outcomes. However, one might expect that outcomes in the sense of broader scale social, economic or environmental change should only ever be evaluated periodically, and with regards the use of a broader suite of policy instruments, tools and processes. There is an issue of scale here – one should not really expect single DISTs to be 'game changers', but one might reasonably expect that more widespread use of DISTs or DIST types (e.g. GIS) should create some form of broader but measurable outcome. For example, one might expect that the use of GIS to better map the spatial location of endangered species or habitats should play a key role in arresting the further decline of those species or habitats. Such outcomes are measurable in principle over a period of 5-10 years or more.

Matthews et al. (2011) conclude with a note of caution that the danger in an evaluation agenda is that the expectation is placed on researchers or DIST developers to be able to convincingly and measurably demonstrate outcome based impact from their work, but that this cannot typically be achieved. In response to this position one might reasonably then argue for placing less emphasis on trying to understand the

relationships between DIST use and outcome and instead focus on better understanding the relationships between DIST use and ‘process effects’ – the kinds of impacts which occur with regards to individuals and organisations using DISTs. Doing so would reposition broader outcomes as being a function of the behaviours of many people or organisations, or many behaviours of the same people or organisations, and therefore only indirectly as a consequence of DIST use. Instead assessment and evaluation activities should then focus on improving our understanding of the relationships between DISTs and individual or organisational behaviour as the intermediate step. Doing so does not avoid the difficulties in assessing or evaluating outcomes, rather it recognises that outcomes are complex aggregates of a whole range of individual and organisational behaviours and should only be assessed or evaluated as such.

To focus on assessing and evaluating individuals and organisations means that we need to refine our understanding of how ‘process effects’ are generated within organisations, and how desirable process effects might be promoted through particular DIST development practices. We will start with a brief examination of how assessment and evaluation of organisations can be used to inform the development and enhance the adoption of DISTs before proceeding on to the more problematic issue of assessing organisational change and impact in relation to DIST use.

3. INCORPORATING ORGANISATIONAL ASSESSMENT INTO DIST DEVELOPMENT

There is clear empirical evidence from across the information systems literature that involving users in development is more likely to lead to positive adoption and outcomes from use (Diez and McIntosh, 2009). Such involvement should be viewed as a form of assessment and evaluation for it involves gathering and analysing information on users’ needs, preferences and desires and reflecting those appropriately within the design of a DIST. Such involvement, particularly if geared towards understanding the usability of DISTs, can be viewed as trial assessments of the extent to which desirable ‘process effects’ are generated.

A growing body of professional practice is emerging amongst the environmental modelling and software community about how interactions with users (in the sense of whole organisations and single individuals within those organisations) should be structured to produce more useful and more used tools. Pioneering work over the past decade by commercial companies and academics (van Delden et al., 2011) has shown that a number of different roles are required within an environmental DIST development project for success. Figure 2 depicts these roles and some of their relationships.

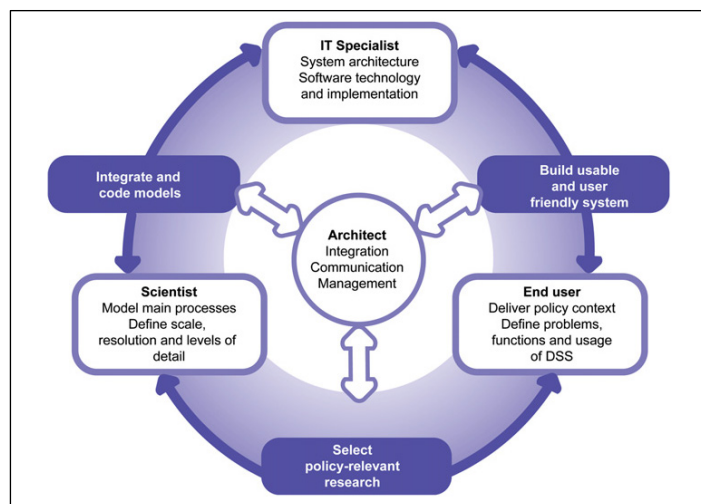


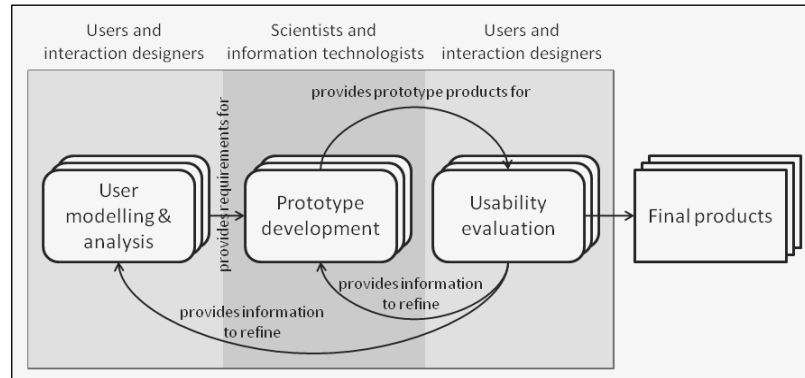
Figure 2. Major roles, responsibilities and interactions during the development of environmental DISTs (from van Delden et al., 2011)

Critical to success is a clear distinction between the end user (person or persons representing an organisation), the scientist (person or persons providing domain specific scientific knowledge into the models), the IT specialist (person or persons providing the skills to code and turn DIST designs into software with appropriate interfacing) and the architect (the person able to engage with every role and ensure overall co-ordination). Between these roles a number of assessment relationships may exist. The IT specialist(s) must assess user needs in terms of software functionality and interfacing, and assess the usability of the tool once built. The scientist must be able to assess the scientific and policy question needs of the end user, to assess the suitability of scientific models to answer those questions, and be able to evaluate the quality of the software implemented by the IT specialist. Central to the whole process is the ability of the architect to assess and evaluate what each party is doing, and to help ensure that there are no mis-understandings or mis-communications.

How user needs should be captured, and how tool usability and usefulness should be assessed and evaluated is less well developed within the environmental modelling and software literature. Methods for doing so are however significantly more well developed within the interaction design literature, which is related both to

information systems and more generally to product design. For example, user needs can be elicited and used to inform DIST design through a combination of user modelling to capture the underlying ambitions, needs and desires of individual users (see Cooper et al., 2007) with work modelling (see Holtzblatt et al., 2005) to capture the ways in which DISTs should interface with and improve users work flow patterns. Usability analysis can be employed to then test DIST attributes such as learnability, understandability etc. (see Tullis and Albert, 2008), the results of which could then be incorporated into DIST prototype refinement. Figure 3 shows a proposed scheme for employing these interaction design methods with the kinds of DIST roles distinguished by van Delden et al. (2011).

Figure 3. Proposed user centred DIST development approach showing key activities (rounded rectangles), activity inputs / outputs (labeled arcs) and role responsibilities (shaded areas and text at top indicate who does what)



4. PROCESS EFFECTS AND ORGANISATIONS

Measuring the impacts, or as Matthews et al. (2011) label them, ‘process effects’ of DIST use on organisations depends on one’s view of what constitutes an organisation – the attributes one might be interested in or expect to change will be determined in part by how one views organisations. The framework provided by Matthews et al. (2011) does not provide sufficient conceptual structure regarding organisational attributes to provide a basis for determining how to measure DIST use impacts. Something else is required.

Quoting Checkland and Scholes (1990), Reeve and Petch (1999) raise some questions that illustrate the basis for a view of organisations which might be labelled ‘social interactionism’. The questions they raise challenge the kinds of impacts and effects one might expect to see as a consequence of DIST use:

“In the 1960s the adoption of the standard assumption from management science that organisations could be treated as if they were instrumentalities, goal-seeking machines, seemed not unreasonable. But in the 1980s such an assumption seemed increasingly dubious. Why not treat organisations as if they were not goal-seeking machines but discourses, cultures, political battle-grounds, quasi-families, or communications and task networks?”

Answering these questions, Checkland and Scholes (1998) articulate the Process for Organisation Meanings (POM) model to conceptualise what organisations are in terms of a set of inter-linked social processes, and how those processes relate to using information systems (to be treated here as synonymous with DISTs). Figure 4 shows the POM model.

Interpreting the POM model one might expect to see organisational impacts of DIST use taking the form of behavioural changes amongst individual actors or groups of actors within organisations. For example one might look to see changes in individual understandings of the basis of or consequences of action; to see changes in intention to act, or; to see changes in agreement or disagreement over issues. Some of these kinds of individual and group behavioural changes have been observed empirically in relation to group system dynamics model building processes in organisations (Rouwette et al., 2002). Related work in participatory modelling has begun to report some empirical detail (Voinov and Bousquet, 2010), but the context is not organisational and so not necessarily transferable in insight.

Rouwette et al. (2002) undertook a significant meta-analysis of 107 organisational group model-building processes to understand the extent to which a range of different social outcomes were achieved, and were achieved positively. The outcomes they assessed included, at the individual level, ‘insight generated’; at the group level, ‘commitment’ (the intention to implement changes suggested by the modelling exercise), ‘communication improvements’, ‘consensus generated’, ‘shared language created’, and; at the organisational level, ‘system changes’ and ‘positive results’. The relationships between desirable changes and the characteristics of the group model building process were difficult to tease out with really only ‘commitment’ and ‘system changes’ seeming to be lower under one configuration.

Diez and McIntosh (2011) undertook a multiple organisation study of the drivers for, constraints on and impacts of DIST use in a particular field of environmental policy and management – desertification. The results of this study were able to attribute different kinds of impact to different kinds of DIST type from remote sensing through GIS to simulation models and DSS. A wide range of impacts of DIST use were reported by employees of the organisations interviewed, with the biggest range of impacts reported for the DIST types most widely used – GIS and remote sensing. Impacts ranged from those related to increases in efficiency and effectiveness of tasks and entire organisations, through improved participation and communication to cost impacts including the need to hire new staff, to re-train staff and to invest in IT. There were no clear relationships to DIST type, leading one to either conclude, as with Rouvette et al. (2002) did, that more and better data collection is needed, or perhaps towards the conclusions of Matthews et al. (2011) that the complexities of the social process are significant and likely to make the identification of easy patterns of cause and effect between DISTs, use and impacts (whether process effects or outcomes) difficult at the very least.

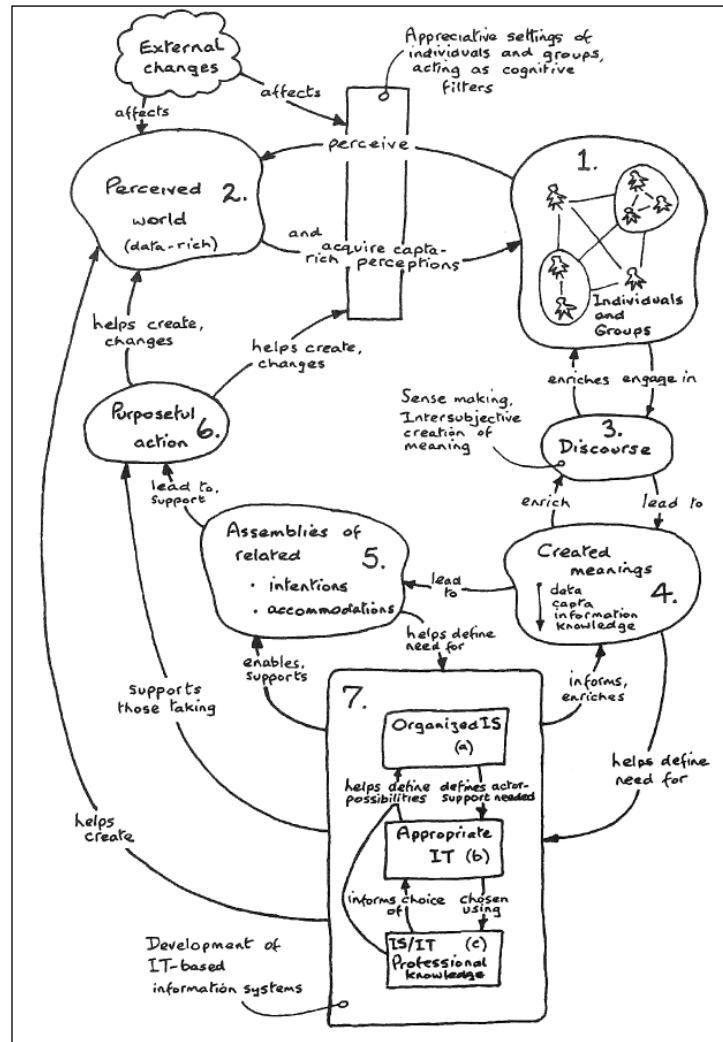


Figure 4. Process for Organisation Meanings (POM) model (from Checkland and Holwell 1998)

5. CONCLUSIONS

The social, economic and environmental challenges of the 21st century require the effective transfer and use of scientific knowledge into policy. DISTs offer one route for this transfer. Ensuring that DISTs offer real value as a transfer mechanism will require improved design practices and an improved understanding of the relationships between DIST use and changes in organisational behaviours and attributes (like efficiency and effectiveness), and an improved understanding of the relationships between those behaviours and attributes and broader social, economic and environmental ambitions. Assessment and evaluation practices offer the potential to help realise this agenda.

Improved use of interaction design methods for capturing user needs and assessing DIST usability offer the potential to improve tool adoption rates and eventual utility. The extent to which improved adoption will result in desirable organisational changes is unclear however. Available published evidence is scant and there is no clear consensus on the kinds of changes which should be prioritised as desirable, nor how to measure them or best promote them in tool development. In short, more conceptual and empirical work is required before we really understand the relationships between DIST characteristics, use processes and impacts within and across organisations, and how best to manage these relationships.

ACKNOWLEDGEMENTS

The preparation and presentation of this paper would not have been possible without the financial assistance of the conference organisers and of the International WaterCentre. Intellectually this paper would not have occurred had it not been for the many fruitful discussions held with colleagues over the past five to ten years, including but not limited to, Paul Jeffrey, Roger Seaton, Hedwig van Delden, Esther Diez, Keith Matthews, Susan Cuddy, Jennifer Ticehurst, Stefan Sieber, Carlo Giupponi, Jaroslav Mysiak, Bernard Hahn, Tim Oxley, Nick Winder, Nigel Quinn, Martin van Ittersum and Sondoss El-Sawah. Many thanks. I am indebted.

REFERENCES

- Checkland, P., and Holwell, S. (1998). *Information, Systems, and Information Systems: Making Sense of the Field*. John Wiley and Sons, Chichester.
- Checkland, P., and Scholes, J. (1990). *Soft Systems Methodology in Action*. John Wiley and Sons, Chichester.
- Cooper, A. Reimann, R., and Cronin, D. (2007). *About Face 3 : The Essentials of Interaction Design*. John Wiley and Sons Inc., Chichester.
- Diez, E., and McIntosh, B.S. (2009). A review of the factors which influence the use and usefulness of information systems. *Environmental Modelling and Software* 24:588-602.
- Diez, E., and McIntosh, B.S. (2011). Organisational drivers for, constraints on and impacts of decision and information support tool use in desertification policy and management. *Environmental Modelling and Software* 26:317-327.
- Holtzblatt, K., Burns Wendell, J., and Wood, S. (2005). *Rapid Contextual Design, A How-To Guide to Key Techniques for User-Centred Design*. Morgan Kaufman, San Francisco.
- Matthews, K.B., Rivington, M., Blackstock, K., McGrum, G., Buchan, K., and Miller, D.G. (2011). Raising the bar? – The challenges of evaluating the outcomes of environmental modelling and software. *Environmental Modelling and Software* 26:247-257.
- McCown, R.L. (2002). Changing systems for supporting farmers' decision making: problems, paradigms and prospects. *Agricultural Systems* 74:179-220.
- McIntosh, B.S., Ascough II, J.C., Twery, M., Chew, J., Elmahdi, A., Haase, D., Harou, J., Hepting, D., Cuddy, S., Jakeman, A.J., Chen, S., Kassahun, A., Lautenbach, S., Matthews, K., Merritt, W., Quinn, N.W.T., Rodriguez-Rodan, I., Sieber, S., Stavenga, M., Sulis, A., Ticehurst, J., Volk, M., Wrobel, M., van Delden, H., and El-Sawah, S. (accepted), Environmental Decision Support Systems (EDSS) development – challenges and best practices. *Environmental Modelling and Software*
- Nilsson, M., Jordan, A., Turnpenny, J., Hertin, J., Nykvist, B., and Russel, D. (2008). The use and non-use of policy appraisal tools in public policy making: an analysis of three European countries and the European Union. *Policy Science* 41:335-355.
- Oxley, T., McIntosh, B.S., Winder, N., Mulligan, M., Engelen, G. (2004). Integrated modelling and decision-support tools: a Mediterranean example. *Environmental Modelling & Software* 19 (11) 120-125.
- Reeve, D., and Petch, J. (1999). *GIS Organisations and People, A Socio-Technical Approach*. Taylor and Francis, London.
- Rouwette, E.A.J.A., Vennix, J.A.M., and van Mullekom, T. (2002). Group model building effectiveness: a review of assessment studies. *System Dynamics Review* 18:5-45.
- Tullis, T., and Albert, B. (2008). *Measuring the User Experience: Collecting, Analyzing and Presenting Usability Metrics*. Morgan Kaufmann, San Francisco.
- United Nations (2010). *World Population Prospects, The 2010 Revision*. <http://esa.un.org/unpd/wpp/>
- van Delden, H., Seppelt, R., White, R., and Jakeman, T. (2011). A methodology for the design and development of integrated models for policy support. *Environmental Modelling and Software* 26: 266-279.
- Voinov, A. and Bousquet, F. (2010). Modelling with stakeholders. *Environmental Modelling and Software* 25:1268-1281.