Whole-farm models: a review of recent approaches

Michael J Robertson1, David J Pannell2 and Morteza Chalak2

1CSIRO Ecosystems Sciences and CSIRO Sustainable Agriculture Flagship, PMBS, PO Wembley Perth WA 6913
Australia

2Centre for Environmental and Economics and Policy, School of Agricultural and Resource Economics, University of Western Australia, Crawley WA 6009 Australia.

Abstract. There is a wide variety of approaches to whole farm models (WFMs) including purely biophysical models as well as models that combine biological and financial elements. Our study was motivated by the notion that researchers may benefit from guidelines on the choices they must make about modelling approach, when they are interested in moving to the whole-farm scale in order to understand the impacts of management changes on farm businesses. This paper reviews 53 studies published in Agricultural Systems, including the recent development of WFMs for developing country situations. We document current approaches and develop a typology, and describe strengths and weaknesses of various approaches. Models differed in the extent to which they accounted for resource constraints, endogeneity of input levels, spatial heterogeneity, interactions between activities, inter- and intra-annual variability, and risk. Models varied in objective (profit, environment, household food security), audience and whether they used real or representative farms. We found many studies did not provide a new insight into a farming system that has general relevance outside the specifics of the case being examined, did not use sensitivity analysis or validation, and were unclear on the objective of the work and the target audience. In response to this we consider issues around the communication of this type of research, proposing guidelines for the publication of WFM papers that are well documented, clear and useful.

Keywords: whole farm systems modelling, economics, optimisation, simulation, risk, sensitivity analysis, validation.

Introduction

In recent times there has been a proliferation of whole-farm models (WFMs) to address a multitude of questions in agricultural systems. Janssen and van Ittersum’s (2007) review of WFMs concentrated on mechanistic normative models. Normative approaches try to find the optimal solutions and alternatives to the problem of resource management and allocation rather than trying to model the actual behaviour of the farmer (so called “positive” approaches). Mechanistic (as opposed to empirical) approaches use mathematical programming and existing theory and knowledge to encapsulate the processes on farms in an idealised sense. Mechanistic normative models can use constrained optimisation and this is the class of model than Janssen and van Ittersum focussed on. They did not cover non-optimisation models nor did they address issues of spatial explicitness, representativeness, and coupling with biophysical simulation models. They also focussed on industrialised countries and since their review there has been a number of WFMs produced for developing country situations (e.g., Giller et al. 2006; Tittonell et al. 2009). Hence, it is timely to review recent trends.

This review is intended to inform researchers who are currently working with models at the plot or field level and who are interested in moving to the whole-farm scale in order to understand the impacts of management changes on farm businesses or use whole-farm cash flow modelling and who need guidance about how to legitimately extrapolate their findings to other farming systems in adjacent locations. The whole-farm modelling approaches of interest include purely biophysical models as well as models that combine biological and financial elements. There is a wide variety of approaches within this scope, so we believe that researchers may benefit from guidelines on the choices they must make about modelling approach. For example, choices include the extent to which models should represent seasonal variability, price variation for key inputs and outputs, interactions between activities (a particular issue for mixed farming systems), and heterogeneity of soil resources. In making these choices, modellers face a trade off between the benefits of developing a more comprehensive model (e.g., broader relevance, improved model capability, perhaps greater accuracy) and the costs of providing greater comprehensiveness (e.g., greater data needs, greater difficulty in
model construction and solution, delays until model completion, model opaqueness).

We review recently published papers on WFM with the aim of providing insights into these choices. We document current approaches and develop a typology of models. Strengths and weaknesses of various approaches are described. In addition, we consider issues around the communication of this type of research, proposing guidelines for the publication of WFM papers that are well documented, clear and useful.

Methods

Papers using WFM published in the journal Agricultural Systems between 2006 and 2011 were read and classified according to a range of criteria shown in Table 1. All reviewed papers post-date those included in the survey of Janssen and van Ittersum (2007).

In classifying papers, the following categories were used:

1. Constrained resources. The mix and level of production activities in the model is constrained by the availability of resources, such as labour, machinery and finance.
2. Endogeneity of input levels. Levels of inputs, such as fertiliser, labour and feed, are set endogenously to optimise the model’s objective function.
3. Interaction of activities over time. One activity, such as cropping, interacts with another, such as grazing, on the same area of land. For example, nitrogen fixation by a legume pasture may increase the yield of cereal crops in subsequent years.
4. Interaction of activities over space. One activity, such as cropping, interacts with another, such as grazing on another part of the farm. For example, the production of manure by livestock is used as fertiliser on crops.
5. Spatial separation of activities. For example, forage production for hay, grain production, and grazing occur on separate land use units.
6. Spatial heterogeneity. Land use units have different production levels, inputs and profit margins, often as a function of soil type or topography.
7. Dynamics. Within-year dynamics are represented through sub-periods of the year, often to account for variation in feed supply and demand in livestock-related activities. Between-year dynamics are included through representation of inter-annual variability in activities such as crop production. Related to this is the extent to which a WFM focuses on tactical (e.g., within season enterprise and production method selection) versus strategic decisions (e.g., multi-year enterprise selection, flock or herd structural decisions).
8. Real versus representative farms. Models run on real farms may require a high level of detail, while representative-farm models may be a summation of local knowledge or farm surveys and hence an amalgam of synthetic and actual data on farm characteristics.
9. Audience. The aim of the model is to inform the settings of farm policy (e.g., levels of farm subsidies), research prioritisation (e.g., selecting those technologies with largest impact on whole-farm profit), or farm management (e.g., decisions made by farmers and their advisors of a strategic or tactical nature).
10. Objective. These are the key outputs or performance indicators for the model, and in the case of optimisation models are the objective function. Objectives might be financial (e.g., profit maximisation, risk minimisation), social (e.g., objectives related to the farm household), or environmental (e.g., minimise nutrient losses, greenhouse gas emissions, or energy consumption).
11. Variation accounted for. This refers to both variation in production (e.g., as a function of season), prices and costs and the reporting of model output.
12. Timeframe. For non steady-state models this category represents the number of years represented in a single solution of the model.

A total of 53 studies utilising 42 models were analysed. Studies based on smallholder agriculture in less developed countries (LDCs) comprised 21% of the 53 studies. All studies were of the type described by Janssen and van Ittersum (2007) as “mechanistic normative”. Normative approaches try to find the optimal solutions and alternatives to the problem of resource management and allocation rather than trying to model the actual behaviour of the farmer (so called “positive” approaches). Mechanistic (as opposed to empirical) approaches use mathematical programming and existing theory and knowledge to encapsulate the processes on farms in an idealised sense.

We conclude that the selection of one journal for the review of literature will exclude many WFM studies from our survey. The scope of Agricultural Systems ("...an international journal that deals with interactions - among the components of agricultural systems, among hierarchical levels of agricultural
systems, between agricultural and other land use systems, and between agricultural systems and their natural and social environments”) ought not to have excluded any particular type of WFM paper. Ultimately the aim of the study was to arrive at some general insights and principles derived from this cross-section of the literature on whole-farm modelling.

Results

Representation of resource constraints can make important differences to the evaluation of the impact of technologies and management practices. Models differed in terms of whether they accounted for resource constraints, such as labour, machinery and finance. Those models with an economic emphasis (often based on optimisation) evaluated the impacts of resource constraints explicitly (68% of the studies examined, Table 1), while those with a biophysical emphasis (often based on simulation) did not. Omitting such constraints from the analysis may result in the benefits of interventions being over-estimated. Of course, simulation modellers can consider whether any particular management strategy violates a specific resource constraint, but this is a relatively cumbersome process, and is rarely done in practice. Others argue that the implications for resource use can be handled by using outputs of resource use (e.g., labour) as a criterion for comparison of scenarios ex-post (Lisson et al. 2010).

Models differed in the degree to which they accounted for within-year and between-year variation. Models with livestock activities tended to focus on within-year dynamics (28%, Table 1) because of the need to reconcile feed supply and demand on a seasonal basis (Bell et al. 2008a), whereas those with an emphasis on crop production included only between-season variation (8%). Only 8% of studies accounted for neither within- nor between-season variation while 43% of studies accounted for both. The approaches that were used ranged from full-dynamic simulation models (often daily time step) (e.g., Chapman et al. 2008) to those that represented a single year without intra-year dynamics (e.g., Yirdoe et al. 2006).

A total of 13% of models accounted for price variation only, 17% for seasonal variability only and 21% for both (Table 1). When prices were varied it was done so through changing baseline values up and down by arbitrary percentages (e.g., Byrne et al. 2010). None of the reviewed models used an assumed or actual distribution or sequence of prices, although there are examples where this has been done previously (e.g., Kingwell 1994). Many of the reviewed models were based on mean yield or price values, based on a sequence of years but ignoring variation from the mean (e.g., Huhtanen et al. 2011).

A common reason for modelling at the whole-farm level is to examine inter-enterprise interactions and optimal enterprise mix. A number of models were applied to mixed crop-livestock systems (49%), and most treated them as discrete enterprises (74%) operating on separate parts of the farm (Table 1). The notable exception is the MIDAS model (Kingwell and Pannell 1987) based on Australian mixed farming, where there is close temporal and spatial integration of crop and livestock production. This review found eight studies using MIDAS, one applied to mixed crop-livestock issues in Australia. Studies mostly concentrated on the evaluation of new plant options for the farm feedbase (shrubs, perennial pastures, annual pasture, and perennial grain crops) (O’Connell et al. 2006, Bell et al. 2008b, Gibson et al. 2008, Bathgate et al. 2009, Doole and Weetman 2009, Byrne et al. 2010, Monjardino et al. 2010), but also looked at new rotational options (Robertson et al. 2010): Household models in LDCs also linked crops and livestock because of the need to account for livestock utilising crop residues and the use of manure on croplands (e.g., Giller et al. 2006, Lisson et al. 2010, Parsons et al. 2011; Rufino et al. 2011).

Half of the reviewed models represented spatial heterogeneity in land-use units within the farm, including associated heterogeneity in productivity and suitability to different crop and livestock activities (Table 1). Amongst the industrialised agriculture applications, spatial heterogeneity is a notable feature of the MIDAS model where, depending on the version, up to 8 land management units (or soil types) can be specified for a farm. This emphasis in MIDAS is undoubtedly due to the heterogeneous nature of the soil-landscape system found on large farms (>3000 ha) in Western Australia, where MIDAS originated (Kingwell and Pannell 1987). The lack of spatial heterogeneity in other whole-farm models in industrialised settings is probably a function of relatively small farm size. The only other model where spatial issues were a prominent feature was NUANCES (Giller et al. 2006) where human-induced gradients of soil fertility in relation to the farm homestead on smallholder farms in Africa must be accounted for when addressing whole-farm
resource allocation and soil fertility interventions.

Studies varied in terms of whether modelled farms were based on real farms (e.g., Evans et al. 2006) or synthesised representatives of a population (e.g., Crosson et al. 2005). Most studies (75%) used representative farms (Table 1), and farm surveys were used commonly to specify what was representative. Surprisingly, few models varied key characteristics of the representative farm in sensitivity analyses (e.g., O'Connell et al. 2006). Where real farms were used, parameters were often obtained from farm surveys. For example, Claassens et al. (2009) used farm surveys to assess the regional impact of a scenario by aggregating the impacts on each whole-farm up to a district or regional level.

The objective of the models was either expressed as the objective function in optimisation models or as the main reported output in non-optimising models. Unsurprisingly, the objective was predominantly profit (industrialised studies) or food security (smallholders in LDCs), although a significant number also included energy (e.g., Ahlgren et al. 2009), natural resources (e.g., Belhouchette et al. 2011) or environmental measures (e.g., Happe et al. 2011), labour (e.g., Parsons et al. 2011) and 8% of studies accounted for risk (e.g., Mosnier et al. 2009) (Table 1). Household models built to address issues for smallholder farmers in LDCs (21% of studies) all had a model objective of maximising food security. All models built for industrialised agriculture had maximising profit as the objective and a significant number (21%) had profit and some measure of environmental impact, such as energy use, greenhouse gas emissions, soil carbon levels or nutrient losses (e.g., Nousainen et al. 2011). Only one had a "social" objective (maximising labour use, Cittadini 2008). Significantly, only one study had risk reduction as part of the model objective function (Lien et al. 2007).

The modelling studies nominated their target audience or end-user of the model results as prioritisation of research (45%) (e.g., Bathgate et al. 2009, Affholder et al. 2010), government policy that was unrelated to research prioritisation (22%) (e.g., Thornton et al. 2006, Matthews et al. 2006, Happe et al. 2011), farm management guidelines and recommendations (16%) and 18% of studies had more than one target audience (e.g., van Wijk et al. 2009) (Table 1).

Discussion

The insights gleaned from this review lead us to identify three dominant types of WFMs in use. The types are a mix of methodology, scope and target domain. Inevitably there is some overlap.

A typology of whole-farm models

Static optimisation in industrialised agriculture. This approach represents the biological, physical, technical and managerial relationships of a farm and allocates available resources and recommends decisions in order to maximise the objective function of whole-farm profit, subject to resource, environmental and managerial constraints. This class of model was reviewed thoroughly by Janssen and van Ittersum (2007). All of the studies reviewed here used a comparative-static framework, which means that the dynamics of changing from one state to another are not captured. Both intra-year and inter-year dynamics can be represented (e.g., Kingwell et al. 1993), although they are assumed to be the same in every cycle of production. This is not to say that economic optimisation models are incapable of representing a sequence of years (as opposed to a cycle), but all reviewed studies did not do so, presumably for reasons of model simplicity. In the simplified approaches used, year-to-year variations in weather were not explicitly described, and nor were variations in price. However, these models can be run with a range of parameter values to assess the influence of different production levels on the profit-maximising mix of enterprises and the level of farm profit.

Strengths of this approach are that interactions between production activities within a year and between years can considered in enterprise selection, that spatial heterogeneity of resource quality can easily be represented, that resource constraints are represented, and that the optimal integrated package of management practices can be determined. The output includes the marginal values of resources and the level of improvement needed in each practice for it to enter the optimal solution. Examples of this approach include Crosson et al. (2006), Francisco and Ali (2006) and the MIDAS studies listed above. Models are commonly based on representative farms within a defined region. Optimisation approaches are ideal for ex-ante research evaluation (Pannell 1996) because they allow easy assessment of how new practices or technologies are likely to best fit into the farming system. Their main weakness is that
they use the representative farm approach, with the result that the relevance of results to any particular farm is not completely clear. They also require biophysical relationships to be represented in a relatively simplified form.

**Household models in developing world agriculture** These have the objective function of improving food security and need to account for unique features of smallholder farms such as: household food requirements, the contribution of off-farm income, the ownership of livestock as a significant capital and cultural "asset", the utilisation of communal lands, and the economic implications of home consumption of farm production versus market purchase (Thornton and Herrero 2001, Giller et al. 2011). Notable models are IMPACT (Herrero et al. 2007), NUANCES (van Wijk 2009), the model of Parsons et al. (2011), and IAT (Lisson et al. 2010). Spatial heterogeneity is often accounted for (e.g., soil fertility gradients and variable land types). As economic performance of smallholder households is highly dependent upon resource constraints, accounting for resource endowments of farmers (draft animal power, cash to purchase inputs, labour for farm activities) is an important feature of these models, often informed by farmer surveys (e.g., Tittonell et al. 2009). Both optimisation (Yiridoe et al. 2006, Giller et al. 2006) and non-optimisation (Thornton et al. 2006; Lisson et al. 2010) approaches have been used and short (Lisson et al. 2010) and long-term (Giller et al. 2006) effects are accounted for. A variety of approaches are used to generate the technical coefficients. Some household models used process-based simulation models to generate appropriate input–output coefficients which are then fed to the farm model (Lisson et al. 2010), some link simulation models together (Parsons et al. 2011), while others used summary functions derived from more detailed models (Chikowo et al. 2008).

The strengths of these household models is that they account for resource endowment of smallholder farmers and the economic implications of home consumption of farm production versus market purchase, which are both unique features of agricultural households in LDCs. These features are also associated with a significant weakness of such models: because there is large heterogeneity in resource endowment and degree of reliance on off-farm income, in addition to biophysical variation (farm size, soil type mix, mix of activities) this poses a particular problem for configuring a representative farm. It is also difficult to represent non-financial or food security objectives such as the keeping of livestock for cultural purposes. On top of these weaknesses is the considerable challenge of collecting quality data to parameterise such models in data-sparse environments.

**Biophysical simulation** In this class of model, resources (e.g., labour, expenditure on inputs) are supplied exogenously, often informed by surveys of real farms and hence no constraints are imposed apart from those imbedded in the biophysical functions. More detailed specification of management options is possible, and they often do account for year-to-year variability in weather. They do not seem to have been applied to spatially heterogeneous situations, perhaps due to the increased complexity of the computational task, although there is no theoretical reason why they could not be. Good examples of this class of model are Guimarães et al. (2006), who examined the impacts of variation in production performance of linked dairy-beef production systems on farm financial performance, and Chapman et al. (2008) who examined the impact of feedback composition on dairy farm production and economic performance. Moore et al. (2011) used simulation to inform a budgetary approach to the analysis of water-use efficiency at the whole-farm scale.

The strengths of this approach are that it supports comprehensive and detailed representation of biophysical processes and allows seasonal variability to be dealt with relatively easily (Donnelly et al. 2002). The main drawbacks of simulation approaches are that resource constraints on production activities (e.g., labour, machinery, finance) are not imposed and that optimisation is not used, with the result that the strategies examined may not be the most relevant to farmers; and that it is relatively difficult to represent spatial heterogeneity of resource quality, particularly soil types, across the farm.

**Emerging approaches** In our survey of the literature we detected three emerging approaches to whole-farm modelling that are worthy of highlighting. In the first approach, static optimisation and dynamic simulation are used jointly to overcome limitations with the use of either approach alone. The static optimisation models can be used to define farm configuration subject to resource constraints, and this "optimal" configuration is then used to set the realistic boundary
within which the biophysical model can be used to deal with variability (e.g., stocking rates, rotations, input levels). Whitbread et al. (2010) provides an example of this approach with smallholders in Africa. They show that this "soft coupling" of the dynamic with static optimisation modelling methods offers promise in being able to address issues of uncertainty in farm systems within a realistic resource-constrained scenario. It combines the best elements of each by (i) using the sophisticated modelling capacity that is available through platforms such as APSIM (Keating et al. 2003) to realistically represent the key drivers of the farming system (Chikowo et al. 2006), and (ii) use the household survey approaches to enhance the understanding of resource constraints and flows (Herrero et al. 2007; Tittonell et al. 2009).

The second approach involves the use of WFM’s with farm surveys to assess regional-scale issues, such as the potential scale of adoption of a new farming practice. This approach avoids the limitations inherent in a single-model configuration by being representative of a wider population of farms or farmers. For example, Claessens et al. (2009) used farm survey data and a simple farm model to assess ex ante the economic viability of adopting dual-purpose sweet potato in one district in western Kenya. In a simple ‘representative farm’ analysis, typical or average values of costs and returns would be used to evaluate whether the alternative practice (in this case dual-purpose sweet potato) was likely to be profitable. The significant limitation of that type of analysis is that it is difficult to generalize the analysis to represent the population, particularly where there is a substantial degree of heterogeneity in the farming conditions or context. The methodology of Claessens et al. (2009) was designed to take spatial heterogeneity into account by using the available data to represent the bio-physical and economic heterogeneity in the population.

A third emerging approach is that using multi-criteria approaches, although we found little use of it in our survey. Janssen and van Ittersum (2007) reported that nine out of 48 studies in their survey used multi-criteria approaches, and argued the case for such modelling approaches where decision makers are faced with multiple, often conflicting, objectives of which profit maximisation is only one. It is noteworthy that, in our survey of 53 studies, only one (Francisco and Ali 2006) took a multi-criteria decision making approach. However, a number of studies that had profit maximisation as their core objective also recorded levels of other variables (e.g., environmental) that might be of concern to decision makers.

Model “validation”

Janssen and van Ittersum (2007) argued that a WFM and its outcomes should closely match reality and they found in their survey of studies that half carried out some sort of comparison of modelled with actual farming practices. The fact that “validation” is not carried out more frequently is not surprising considering the lack of datasets to calibrate and validate WFM’s (Thornton and Herrero, 2001). Model “validation” of representative-farm models poses additional challenges, given the variation that occurs in the real world. Such a comparison will only ever be approximate because of the myriad of uncontrolled factors operating in the survey dataset that will not be accounted for in the model configuration. Another issue is deciding which model outputs to compare with actual data. A model may closely match reality for one measure but differ greatly for other measures.

A more intuitive approach is the use of “sensibility” testing, which has been applied in crop simulation modelling in recent times (Holzworth et al. 2011; Huth and Holzworth 2005). In this approach model output is tested against expert opinion and common sense and the crediblility of the model evaluated against its ability to mimic reality. While this is a subjective process, it is potentially more robust than any objective process that is subject to the flaws listed above. A key feature of sensibility testing is sensitivity analysis of key technical coefficients, which allows a test of the robustness of the model (Pannell 1997).

The audience for WFM’s

Our survey indicated that the overwhelming justifications for whole-farm modelling remain research prioritisation and informing policy, these being the aims for 45% and 21% of reviewed papers, respectively (Table 1). A less common audience has been farmers (17%, as discussed below).

The experience of the MIDAS model in Western Australia, as described by Pannell (1996), is instructive for developers of other WFM’s who aspire to informing research prioritisation. Since 1984, MIDAS has had a strong influence on agricultural researchers and research managers in the Western Australian Department of Agriculture and
Food and in several national research centres in Australia (the Centre for Legumes in Mediterranean Agriculture, the Cooperative Research Centre for Plant-Based Management of Dryland Salinity, and the Future Farm Industries Cooperative Research Centre). The model has (a) brought together researchers (of various disciplines) and extension agents who otherwise would interact little; (b) allowed scientists and extension agents to assess the economic significance of particular biological or physical information; (c) influenced the thinking of researchers and extension agents about the whole-farm system; and (d) highlighted a large number of data deficiencies and allowed prioritization of research to overcome them. Pannell (1996) lists strategies and practices which may be beneficial to others undertaking similar modelling efforts.

Apart from the Western Australia experience to the best of our knowledge there have been no other cases where major impacts of WFM on research prioritisation have been documented. Sterk et al. (2006), in a reflection of the use of whole-farm design models with decision makers in six case studies, list ingredients that favoured success: an emphasis on learning rather than providing answers, a desire for change on the part of participants, and local relevance of the models. The role of researchers involved in the process was seen as critical and need to be involved with decision makers from the initial problem definition stage. The intended audience should have a guiding role determining the capabilities and outputs of the model, rather than being used by modellers to attempt to find niches for an existing model that was developed according to the judgements and preferences of researchers. These findings mirror those found by the APSRU group in Australia when exploring the role for computerised decision support for cropping farmers (Carberry et al. 2002). The experiences of Sterk et al. (2006), Carberry et al. (2002) and Pannell (1996) emphasise the primacy of the human dimension of model design, use and communication rather than the technical capabilities of such models. There is a need for more studies reporting on attempts to engage decision makers with the models (be they farmers, researchers or policy makers) in order to advance our understanding of the most effective way to generate and communicate model output as an input to decision making.

Farm managers were the main intended audience for only 17% of studies in our survey. For reasons well-discussed by McCown et al. (2006a, b), we believe that complex WFM will remain largely irrelevant to directly informing management decisions of individual farmers. The quote of Dillon (1979, p.11), reproduced by McCown et al. (2006a), describes five reasons why this is so:

"There is no conceptual difficulty in formulating static production economics in terms of a utility-maximising criterion, nor in conceptualising its logic for non-physical processes. The difficulty lies in application. First, data are not available to be able to specify the relevant production processes (both physical and non-physical) to any significant degree – particularly if we recognise the uniqueness of individual farms. Second, the farm system is dynamic, not static, both in the broad as a purposive organisation in a changing environment and also through the pervasive role of biological time-dependent growth processes in its technical subsystem. Third, even if data were available to specify the required production processes adequately, the task of analysis even under perfect information would be both too complex and too costly for either farmers or computer-aided professionals. "Non-optimising" modes of behaviour have to be used. Fourth, the problem of uncertainty has to be handled. Again this is pervasive in agriculture due to the stochastic vagaries of climate and markets especially, but also because of uncertainty about technology, policy and people. While technologies have been suggested to handle such uncertainty, their cost on anything approaching an individual farm basis makes them impractical. Fifth, even if farmers faced the same judgements about the probabilities they faced, they would still have different preferences and so need different prescriptions for utility maximisation across their individual multiple goals."

An interesting attempt to engage with smallholders in Indonesia using a WFM is described by Lisson et al. (2010, p.491):

"...rather than employing an automated optimization strategy, a creep budgeting approach was selected, involving re-specifying various input and
output variables in a systematic manner to explore the system response to these changes. That is, the decision-maker 'creeps' around the economic response surface in a systematic fashion to examine whether there is a shift towards or away from a more optimal solution. In this way, the use of 'what-if' questions provides smallholders, researchers and extension specialists with important insights into how the economic position of the farm-household system will respond to different activities."

In the approach of Lisson et al. (2010), resource constraints relevant to the farm household, such as the supply of and demand for labour, are accounted for exogenously, and probably subjectively, by the farmer. Malcolm (1990) pointed out the irony that relatively simple modelling methods like this may be more suitable than complex models for capturing the full complexities of real decision making.

**Dealing with risk in WFM**

Risk, defined as an unpredictability of the economic consequences of actions, is widely seen as an issue of critical importance to farmers, and hence conventional wisdom states that it ought to be included in WFM. In our survey, 53% of studies accounted for price, weather or both forms of variability. Risk is seen to matter because most farmers are "risk averse" (Binswanger 1980; Bardsley and Harris 1987; Abadi Ghadim and Pannell 2003) – they are prepared to sacrifice some expected income in order to reduce their exposure to risk. Our survey indicated that few WFM studies (22% of our sample) represent both main types of risk (price and yield) and even fewer (8%) explicitly represent the risk attitudes of farmers. Those that did include risk attitudes either represented risk aversion based on standard expected utility theory (e.g., Mosnier et al. 2009) or included risk minimisation as part of the objective function (e.g., Francisco and Ali 2006).

Given the above observations, the question arises as to whether WFM that fail to include risk and/or risk attitudes are fatally flawed. There are several factors that determine the net benefits of including risk and/or risk attitudes in these models: (a) the importance and relevance of risk to the issues that the model is being used to address; (b) the risk attitudes of the relevant population of farmers; (c) whether the purpose of the model is to predict farmer behaviour or to advise farmers; (d) the sensitivity of results to their inclusion; and (e) the costs of including them. In relation to risk attitudes, Pannell et al. (2000) argued that inclusion of risk aversion in models used to advise farmers is likely to provide minimal benefits, at least in developed countries. Because most farmers have low to moderate degrees of risk aversion (issue b), and because of the phenomenon of flat payoff functions (Pannell 2005), the benefits from improving the quality of recommendations by including risk aversion (an aspect of issue d) are often remarkably small while the costs in terms of model development, use and interpretation (issue e) are non-trivial. The benefits may be greater in developing countries where risk aversion is likely to be stronger.

Where the aim is to predict farmer responses, rather than advise farmers themselves, the benefits of including risk aversion may be greater, as farmer responses to risk can be significant even where the benefits in utility terms from doing so are not great. On the other hand, modelling experience shows that the sensitivity of results to inclusion of risk and risk aversion in whole-farm models is often less than the sensitivity to the inclusion of accurate biophysical relationships or accurate parameter values, such as expected prices and yields (Pannell et al. 2000).

Pannell et al. (2000) and McCown et al. (2006a) both concluded that the most important aspect of risk to be modelled is not farmers’ aversion to risk, but rather their short-term tactical responses to variation in weather and prices. There can be major benefits from responding to particularly favourable or unfavourable weather or prices, through pursuit of opportunities or avoidance of losses. The growing number of WFM that consider the dynamic and tactical features of farming (e.g., Kingwell et al. 1993; Donnelly et al. 2002; Mosnier et al. 2009) could be seen as a sign of the growing recognition of the value of WFM in analysing such features of farming systems.

If risk attitudes are important enough to include, then what matters for whole-farm decision making is aggregate risk at the whole-farm level. Hence it seems likely to be ill advised to include only price or yield risk and not both, and yet 32% of surveyed studies did that.

**Towards criteria for publishing papers on WFM**

The proliferation of WFM (42 models in 6 years of Agricultural Systems) raises the issue of whether criteria for judging the
scientific merit of WFM studies are well established. Many of the studies reviewed for this paper were case studies with limited applicability outside the specific situation concerned. A similar phenomenon in crop simulation modelling prompted Sinclair and Seilgman in 2000 to publish a paper in Field Crops Research on criteria for acceptance of scientific papers on crop modelling. They argued that even though manuscripts on crop modelling may describe modelling efforts of practical perspective with local interest, they may not necessarily present an analysis of general, scientific interest.

We propose that six criteria are used to judge the merits of manuscripts reporting on studies developing or using whole-farm models.

Firstly, the study should provide a new insight into a farming system that has general relevance outside the specifics of the case being examined and / or propose a new technique for modelling that improves upon deficiencies in currently-used methods. In the case of general relevance there needs to be clarity regarding the magnitude, causation and broadness of the applicability of findings. Readers need to know how important the findings are. (e.g., What percentage change in farm profit occurs? How food shortage frequency is affected?) Readers need to know why the WFM findings are novel and the main factors contributing to the findings. Are the findings applicable to other farm types or farming regions? For some issues it may be worth publishing findings that convey "old truths" rather than "novel findings" or "new insights", particularly where "bandwagon" issues are emerging and where there are claims of a new paradigm. An example of a study that meets the criteria of relevance and new technique is that of Kingwell et al. (1993). In their study a discrete stochastic programming model of dryland wheat-sheep farms in Western Australia was described and used to identify tactical adjustments to climate and to calculate the value of those tactical adjustments. The study both described a novel approach to the question and generated novel insights to the value of accounting for tactical adjustments in WFM.

Secondly, studies should include a clear statement of the objectives of the work and a use for the results. If the objective of the work is to develop a better modelling technique, then the audience is modeller developers. If objective is to apply a model to address a whole-farm issue, then the various possible audiences interested in research prioritisation, policy evaluation, or guidelines for farm management should be highlighted. The case for treatment of the issue at the level of a whole- farm or household (as opposed to a field or region) should also be justified.

Thirdly, studies should nearly always include some component of sensitivity analysis. Our survey indicated a notable lack of sensitivity analysis in many studies. Sensitivity analysis (particularly to price, production levels, and farm configuration) should preferably be based on some empirical evidence of the range of feasible values to be varied. This could be obtained from farm surveys, historical records of prices, costs and production values, or simulated values of production in response to changes in biophysical settings (e.g., season, soil type, management). Where information is available on price probabilities then this should also be incorporated, e.g., rather than say high and low prices it may also be worthwhile to indicate that the "low" price is the 25th percentile price and the "high" price is the 75th percentile price. An example of the use of sensitivity analysis from our literature survey is the study of Byrne et al. (2010), who used a linear programming approach (MIDAS) to evaluate the factors influencing potential scale of adoption of a perennial pasture in a mixed crop-livestock farming system. In sensitivity analysis they varied commodity prices (standard, low and high values), relative productivity of the perennial pasture, mix of suitable soil types on representative farms, and type of livestock enterprise (meat or wool-producing sheep).

Fourthly, a criterion for any paper on WFM should be that there is some degree of "validation" against actual farm practice or outcomes (as suggested by Jannsen and van Ittersum 2007). As noted above there is a conspicuous lack of attempts to validate model outputs. The exact form which validation takes needs to be flexible and dependent upon data availability and the number of model variables that can be evaluated. As argued above we see merit in some combination of subjective ("sensibility testing") and objective methods (comparisons with farm surveys, etc). One example of an attempt to "validate" model output was the study of Robertson et al. (2010). They compared surveyed areas on farm of broadleaf break crops (canola, pulses) with economically-optimal areas predicted by a WFM. In the cases where
there was agreement between surveyed and predicted areas this provided "validation", while those cases where there was a discrepancy generated a fertile discussion about possible explanations. A fifth criterion is the need for an explicit statement of what inputs are exogenous or endogenous to the model. If different farmer types are being modelled, the assumptions of resource endowments must be clearly stated. This is a particularly important feature for WFM in LDCs, where resource endowment has a large influence on the impact of interventions on household outcomes. The study of Tilton et al. (2009) is a good example of how to couple typologies of farmers varying in resource endowment with a WFM to evaluate the impact of various interventions to intensify the farming system. Finally, a criterion must be documentation of the source of technical coefficients. Jannsen and van Ittersum (2007) noted this was missing from many of the studies they reviewed, as did we. Many journals allow the publication of online supplementary material, so the constraint of space for publication of voluminous tables of model coefficients is now irrelevant. The study of Parsons et al. (2011) is a good example of how this can be done. It provides a model description supplemented by tables of model coefficients, a number of which are derived from expert opinion.

We are conscious that if our six criteria are followed by journal editors this could easily lead to the practical issue of accommodating all these tasks (especially documenting coefficients, sensitivity analysis and validation) within an editorial word limit. There will be a need to achieve the required balance between informing readers of novel findings whilst ensuring these findings stem from a clearly and fulsomely detailed WFM. One solution is that technical or on-line appendices should be a more common feature of WFM studies. Perhaps editors allow more studies with multi-part papers.

In our review of 53 papers, it is our subjective assessment that the greatest deficiencies that authors should address first is that of sensitivity analysis, so that results can be understood within the context of variation in model parameterisation and assumptions. A modest degree of "validation" and some documentation of key coefficients would improve many of the studies we reviewed.

Conclusions
The 53 studies from six years demonstrate the diversity of approaches being used to address whole-farm production, environmental and economic issues and the emergence of a focus on smallholder households in the developing world. Models with a more economic emphasises accounted for constrained resources, while those with a biophysical orientation focussed on dynamics of processes. The further development of approaches to couple dynamic biophysical and static optimisation models would provide opportunities to address issues of dynamics and risk in farming systems under conditions of realistic resource constraints. The focus on most studies is still policy guidance and research prioritisation, with few studies attempting to influence farm management. More studies are needed to explore effective means to engage with decision makers using WFM. This is likely to de-emphasise technical issues and place more focus on problem definition, declaration of assumptions and communication of results along with their uncertainty.

The establishment and acceptance of criteria for publications on whole-farm models should strengthen the quality and transparency of studies and encourage the development and application of modelling studies that will lead to a larger number of generalisable insights. A notable deficiency in many studies was lack of a clear objective or audience for the work. Better documentation of coefficients, validation of model output, and sensitivity analysis around prices, seasonal conditions and farm configuration would strengthen many studies.

Acknowledgements
We are grateful to Prof. Ross Kingwell for comments on an early version of the paper. The work was funded by CSIRO and the Centre for Environmental economics and Policy, School of Agricultural and Resource Economics, University of Western Australia.

References
Affholder F, Jourdain D, Quang DD, Tuong TP and Morize M 2010, "Constraints to farmers" adoption of direct-seeding mulch-based cropping systems: A farm scale modeling approach applied to the


Donnelly JR, Freer M, Salmon EM, Moore AD, Simpson RJ, Dove H and Bolger TP 2002, 'Evolution of the GRAZPLAN decision support tools and adoption by the grazing industry in temperate Australia', *Agricultural Systems*, 74: 115–139.


Herrero ME, Gonzalez-Estrada E, Thornton PK, Quiro C, Walthaka MM, Ruiz R and Hoogenboom


Kingwell RS, Pannell DJ and Robinson SD 1993, 'Tactical responses to seasonal conditions in whole-farm planning in Western Australia', Agricultural Economics, 8: 211-226.


Pannell DJ 2006 'Flat-earth economics: The far-reaching consequences of flat payoff functions in economic decision making', Review of Agricultural Economics, 28: 553-566.


## Appendix

Table 1 Criteria used to classify papers on whole-farm modelling published in *Agricultural Systems* 2006-11 and percent of papers in category (n=53).

<table>
<thead>
<tr>
<th></th>
<th>No</th>
<th>Yes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constrained resources?</td>
<td>69%</td>
<td>31%</td>
</tr>
<tr>
<td>Are input levels endogenous?</td>
<td>65%</td>
<td>35%</td>
</tr>
<tr>
<td>Do different activities on the farm interact over time?</td>
<td>51%</td>
<td>49%</td>
</tr>
<tr>
<td>Do different activities on the farm interact across space?</td>
<td>49%</td>
<td>51%</td>
</tr>
<tr>
<td>Are activities spatially separated?</td>
<td>75%</td>
<td>25%</td>
</tr>
<tr>
<td>Is there spatially-specific heterogeneity?</td>
<td>51%</td>
<td>49%</td>
</tr>
<tr>
<td>Are risk attitudes accounted for?</td>
<td>92%</td>
<td>8%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Temporal dynamics</th>
<th>Intra-year</th>
<th>Inter-year</th>
<th>Both</th>
<th>No</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>29%</td>
<td>18%</td>
<td>45%</td>
<td>8%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Based on real or representative farms?</th>
<th>Real</th>
<th>Representative</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>24%</td>
<td>76%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Primary audience</th>
<th>Government policy</th>
<th>Research prioritisation</th>
<th>Farm management</th>
<th>More than one</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>22%</td>
<td>45%</td>
<td>16%</td>
<td>18%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Objective</th>
<th>Profit/food security</th>
<th>Environment</th>
<th>Both profit and environment</th>
<th>Other</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>57%</td>
<td>16%</td>
<td>22%</td>
<td>1%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Variation accounted for</th>
<th>Season</th>
<th>Price</th>
<th>Both</th>
<th>No</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>18%</td>
<td>14%</td>
<td>22%</td>
<td>47%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Time frame</th>
<th>Short (&lt;=3 years)</th>
<th>Medium (Between 3 and 10 years)</th>
<th>Long (&gt; 10 years)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>37%</td>
<td>27%</td>
<td>35%</td>
</tr>
</tbody>
</table>