

QUALITY OF MODELLING IN FRUIT RESEARCH AND ORCHARD MANAGEMENT: ISSUES FOR DISCUSSION

W. van der Werf¹, C. Leeuwis² and W.A.H. Rossing¹

¹IOBC-WPRS Study Group Critical Assessment of Modelling Approaches in Integrated Crop Management

²Laboratory of Theoretical Production Ecology, Wageningen Agricultural University, Bornsesteeg 47, 6708 PD Wageningen, the Netherlands

³Group Communication and Innovation Studies, Wageningen Agricultural University, Hollandseweg 1, 6706 KN Wageningen, the Netherlands

1. Introduction

The purpose of this paper is to highlight items for discussion during a one-morning workshop on Critical assessment of modelling approaches in Integrated Crop Management. This paper is based on information provided in nine papers that were submitted for this workshop (Blaise *et al.*, Boshuizen and van der Maas; DeJong; Graf *et al.*; Groot; Hardman *et al.*; Mols and Boers; Szafran *et al.*; van der Werf *et al.*; all in this volume) in response to the questions listed in Rossing *et al.* (this volume). Distinction is made between process-based research models, which are considered in section 2, and decision support systems, which are addressed in section 3.

2. Research models

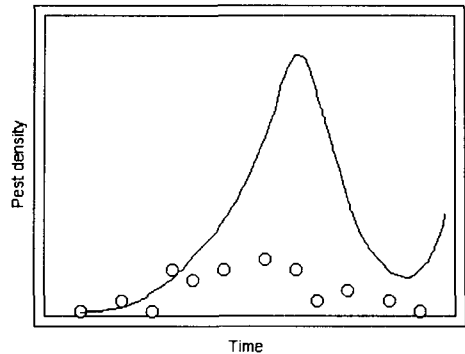
2.1. The process of model validation; when to stop and on which grounds?

The submitted papers indicate that building a research or DSS model helps researchers to systematize their knowledge. It leads to the identification of knowledge gaps and new research questions (e.g. Blaise *et al.*, Hardman *et al.*, DeJong). It seems, however, that one could continue forever to refine and expand a model. Even if the original research questions justified the investment in modelling, one can question whether this holds automatically for each subsequent question that is formulated in the process. If we automatically take up each research question raised by the modelling, we allow models to take over in setting the research agenda. An important issue therefore is how we should go about setting our research agendas, and in particular how we should evaluate research questions that are generated in the process of building a model.

One reason to expand and refine process-based models is the desire of researchers to make models that represent 'all' processes that they know to be relevant in the system. For instance, DeJong (this volume) mentions dry matter allocation to roots as such a refinement. Another reason for model expansion and refinement is the desire to obtain reasonable correspondence to observed data based on observed input functions, mechanistically sensible model structure and quantified relationships, without empirical (site-specific) fudge factors, because such a model could truly be called 'explanatory'. This is a high aim, which appears to be difficult to achieve in complex systems in the field. This is true both for models of multitrophic multi-species plant-pest-predator population interactions (Hardman *et al.*, this volume) and epidemiological processes (Blaise *et al.*, this volume) and for models of crop growth, dry matter accumulation and

allocation, morphogenesis and architecture (DeJong, this volume). An important question, therefore, is at what point should we stop with refining model structure and increasing the number of input functions, in our attempts to make a model describe field experiments better?

Figure 1 - A modeller's horror? Poor correspondence between simulation (drawn line) and field observations (circles). Causes for discrepancy may be manifold, including wrong model code, inappropriate or incomplete model structure, and lack of knowledge about external input functions and initial conditions.



A characteristic pattern in model development is to start off with a basic model and modify it as discrepancies between model predictions and validation trials suggest necessary changes (Fig. 1). The response to discrepancies between simulations and observations is generally threefold:

1. check the code
2. check the conceptual basis of the model for soundness and completeness
3. include more explanatory factors from outside the system in the modelling (Fig. 2).

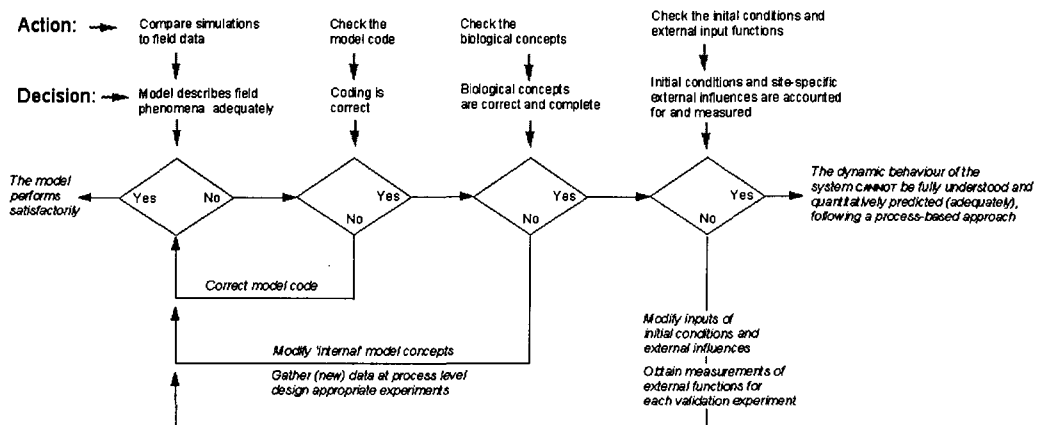


Figure 2 - Flowchart indicating actions and decisions in the cycle of model testing and improvement. The cycle can be said to spiral upward if knowledge and insight in the functioning of the modelled system are accumulated in the process. This learning process may be valuable, whether or not the final result is an adequate simulation of system behaviour.

The response process represents a learning cycle with repeated comparisons between field data and simulation results, to decide upon the effect of model modifications. The case study described by Hardman *et al.* (this volume) illustrates this developmental pattern and highlights the 'to-stop-or-to-detail' dilemma of researchers. Hardman *c.s.* initially constructed a basic age-structured and temperature-dependent predator-prey model, which is conceptually straightforward to build, but noticed that there were important discrepancies between the predictions of that model and what happens in actuality in the field. Considerable complexity and variability in the diet composition of the main predator mite, *Typhlodromus pyri*, was built into the model over a period of 10 years. Beside European red mites, the pest of interest, this predator can use rust mites (*Aculus schlechtendali*), pollen, phyllosphere fungi, and its own offspring as alternate food resources. Such food diversity stabilises predatory mite populations. It enhances an early and effective impact on pest mites and is therefore of considerable relevance to the success of biological control. In the framework of an explanatory model, the consequences of variable food availability on the predation on European red mites are difficult to calculate, due to lack of knowledge of the predator's foraging choices. Work to elucidate the 'rules' for foraging decisions has been undertaken, but it is a major endeavour. Moreover, when such rules are included in the model, extra input functions are needed to specify the time trends of the alternate food sources: rust mites and pollen. In experiments that were executed in the past, such measurements may not have been done; in new experiments, partly aimed at model testing, it may be too much work to collect such extra data. Therefore, a research and modelling program that is solely directed by the desire to get the model 'right' is doomed to get swamped in detailed research questions that might require decades of research work to answer.

Other examples might be chosen. At some point a researcher has to accept that there may be not enough site specific information and knowledge to simulate system dynamics under site-specific conditions, and that it is not practicable and scientifically advantageous to collect such information. The attempt to validate the model will then be open-ended, as the field data are site-specific, whereas the model might apply to an 'average' field, rather than one specific field. This poses a problem in a scientific culture that is based on the principles of 'success' (don't publish failures), 'newness' and 'publish or perish'. Models that do not provide an excellent fit to observed data do not seem to be worth publishing, or worse, authors may feel they have to cover up any discrepancies between simulations and field observations, or lack of independent validation data. Hiding the problems hampers scientific progress, which is critically dependent upon falsification of hypotheses and models. A question is therefore whether authors as well as scientific journals (editors and referees) should be encouraged to publish model falsifications.

There is substantial merit in the ability of process-based models to provide a mechanistically based prediction of system behaviour in response to environment and management. Even if a model cannot be validated, or only in part, due to the mentioned problems, simulation runs ('scenario studies') can be made to investigate options for managing the system under a range of initial and boundary conditions and forcing functions. The mechanistic basis of the model provides an inherent 'explanation' of model outcomes. Studying these explanations may suggest better management alternatives, thus generating new research questions to be investigated experimentally.

Specific questions:

- What are criteria to stop further model refinement?

- Are we satisfied with unvalidated models?
- Should we publish more model falsifications?
- How useful are models in setting the research agenda?

2.2. Relationships between modelling purpose, application and validity requirements

Four research models were presented in the workshop:

- Epidemiology of downy mildew, *Plasmopara viticola*, in grapes in Switzerland (Blaise *et al.*, this volume)
- Growth and yield of peach, *Prunus persica*, in California (DeJong, this volume)
- Population dynamics of European red mite, *Panonychus ulmi*, and its predator *Typhlodromus pyri*, in Eastern North America (Hardman *et al.*, this volume)

Population dynamics of woolly apple aphid, *Eriosoma lanigerum*, and its parasitoid *Aphelinus mali*, on apples in the Netherlands (Mols and Boers, this volume)

Three papers (Blaise *et al.*, DeJong, Hardman *et al.*) emphasize use of the model to better understand system behaviour and to answer broad explorative questions on the effect of management (Table 1). Mols and Boers use a process-based model to answer a very specific research question: does a Canadian strain of *Aphelinus mali*, with lower temperature thresholds than Dutch strains, offer promise for more effective control of woolly apple aphid, due to an earlier impact on the pest in the cool early spring? The model is successful in providing an educated answer, providing a good example of the viability of the scenario approach. The other three models also offer the opportunity to ask such specific questions.

Table 1 - Comparison of four process-based research models with respect to model purpose, desired model attributes and validity requirements.

Model & authors	Model Purpose	Desired Model Attributes	Validation requirements
INSIM Mols and Boers	Comparison of biocontrol efficacy of two parasitoid accessions from different climatic zones	Age-structured, temperature driven model for host-parasitoid interaction in orchard	No data available
VINEMILD Blaise <i>et al.</i>	Understanding mildew epidemics and timing fungicide sprays	Biologically sound, flexible, expandable system model	Epidemiology and damage consistent with general experience
PEACH DeJong	Develop integrated understanding of the annual C-budget of peach and simulate the potential effects of environmental factors, physiological processes and management practices on peach yield and fruit size	Usable for research and teaching	Accuracy of built-in concepts more important than precise numerical predictions
MITESIM Hardman <i>et al.</i>	Understanding mite ecology; Exploring mite management tactics; Guidance in setting research priorities	Biologically sound, flexible, expandable system model	Ideally: valid at orchard level; but realistically representative for 'average' orchards

Specific questions

- Under which circumstances is a simulation model the best tool for asking specific questions, and when are manipulative or controlled experiments a better way?
- What is a better use of process-based simulation models: exploration or prediction?
- Are the validation requirements for a model different when predictive rather than explorative questions are asked?

3. Decision support issues

From an extension point of view, there are four dimensions of 'quality' with respect to DSS:

1. the way in which DSS are connected with learning and decision-making processes;
2. the relevance and validity of the DSS output;
3. the coherence of the built-in 'communication plan';
4. the quality of the DSS development process.

Using these four quality dimensions as section headings, we raise here below various points for debate.

3.1. Connection of DSS with learning and decision-making processes

Is there a need to 'model' learning and decision-making?

As the label 'DSS' suggests, the eventual aim of such computer models is to support human decision-making. It is interesting to note that in virtually all workshop papers 'decision-making' is used as a container concept that is not further refined into different categories. This is remarkable, because human decision-making in itself is a complex process that consists of many stages and sub-processes. Thus, in theory 'decision-support' could take many forms, and be geared towards enhancing very different stages and sub-processes, for example: observation, comparison, problem identification, problem analysis, translation, identification of alternative solutions, evaluation of solutions, experimentation, reflection, etc. Using 'decision-making' as a container concept suggests that developers of DSS wish to support every possible element in this process. The question emerges whether or not this is efficient and realistic. Would it make sense to investigate how decision-making processes evolve, and for which stage or sub-process support is most needed?

Supporting operational decisions or discovery learning?

Most of the DSS which are presented in the papers seem to be geared towards formulating specific advice on particular operational issues, for example on plant protection (Graf *et al.*, this volume) irrigation (Boshuizen and Van der Maas, this volume) and/or the use of agro-chemicals (Groot, this volume). The distribution of tasks here seems to be that the grower provides the data which are deemed relevant by the model, that the model does the reasoning and provides the advice, and that eventually the grower must decide whether or not to follow and/or adapt the advice (Graf *et al.*, this volume). It could be questioned whether or not a grower really learns something from this procedure if -as seems to be the case- the calculation model remains largely a black box. Does his or her understanding increase? Can we really speak of 'decision-support' when farmers take decisions on the basis of models they do not understand? Is there a discrepancy between the understanding gained

by those who develop the DSS, and those who use it? How can DSS be adapted to support 'discovery learning' (i.e. probably the most effective educational strategy) by growers?

What types of problem situations are associated with sustainable and efficient DSS?

The investments that often go along with the development and maintenance of a DSS are considerable. Thus, one would think that it is important that DSS can be used for a prolonged time in order to obtain return on investment. This raises a number of questions with regard to the types of problems that justify DSS development. It was mentioned already that many of the DSS presented in the workshop address operational issues. It seems that the types of problems tackled (irrigation, plant protection, etc.) occur frequently enough (either within or among farming enterprises) in order to justify investment. However, a question which remains is the following: do these DSS continue to generate new answers to similar situations over time, or can their outcomes be summarised in simple rules of thumb which effectively make the models predictable (and therefore redundant) in a short while (i.e. do subtle differences in input generate large differences in output)?

3.2. The relevance and validity of the DSS output

Can and should we make a better problem definition?

Blaise *et al.* (this volume) note that 'In retrospect, a better analysis of the problem and a more precise definition of the objectives, with clear milestones to be reached would have increased the efficiency of the exercise'. This remark seems relevant to several other research and DSS models as well; in other cases too it can be observed that in the course of the development process the underlying research questions, objectives and activities change. A question that emerges is whether this phenomenon is an expression of necessary flexibility and learning, or the result of inadequate preparation and problem analysis.

In this respect it can be questioned *whose* problem models are dealing with: farmers' problems, researchers' problems, policy makers' problems? This question relates in part to the issue of agenda setting raised earlier, and also to the issue of targeting that is raised further on.

How to deal with validity problems?

Not surprisingly, a recurring theme within several contributions is the issue of validation. Even if many models are validated for particular purposes and within particular environmental conditions, it seems almost inevitable -especially for DSS that are used in the field- that models are used in situations where their agro-ecological validity is questionable. This may hold even more for the validity of the advice that is being formulated on the basis of agro-ecological models. If one takes seriously that there exists strategic diversity in farming (see e.g. Van der Ploeg, 1990; Leeuwis, 1993), it may very well be that grower A needs a different solution to a given problem than farmer B. How can we tackle these types of problems? Should we try to solve them within our models, or should we make organisational arrangements in order to deal with this (e.g. one could imagine that that validity is assessed in mutual debate, rather than in refining a model)?

3.3. The coherence of the built-in 'communication plan'

From an extension point of view, DSS are a means of communication between those who develop it, and those who use it. Thus, one could argue that implicit to a DSS is (or

should be) a particular 'communication plan'. Normatively speaking, a 'good' communication plan can be characterised as a plan in which a coherent balance exists between the following elements (see Van den Ban and Hawkins, 1996):

- goal
- message
- target audience
- media of communication
- organisation

As we have already discussed several issues that are related to 'goals' and 'message' in the preceding sections, we will raise a few issues in relation to the remaining three elements.

Why are prospective users a moving target?

Groot and Boshuizen & Van der Maas suggest that in the process of developing DSS, the target audience tends to shift from growers or grower study clubs to extension agents. Among communication planners this would be considered a very tricky thing, as one would expect that extension agents have different needs, problems and questions than growers, and operate in a rather different media environment. The underlying question here seems to be whether or not to select the target audience on the basis of the medium we use (i.e. computer models), or to select a medium on the basis of the target audience we want to reach.

Added value vis-à-vis other means of communication?

If one considers DSS a medium for communication, it becomes clear that DSS are only one out of several media that might be used to convey a particular message. Hence, it is important to think critically about the added value of this particular medium. From a viewpoint of communication science, the key advantage of DSS is that they incorporate both characteristics of mass media and interpersonal media. They allow for interaction between the software and the user so that -at least in theory- they combine a large coverage with a certain degree of message specificity. Other added values that are mentioned in the workshop papers include the increased speed of calculation (Groot, this volume), communication speed (Graf *et al.*, this volume) and the fact that models are less fragile and laborious than soil moisture measurement instruments (Boshuizen and Van der Maas, this volume). In the latter case it is interesting to note that the eventual introduction of less fragile and laborious tools has apparently not resulted in an abandoning of the idea to develop a DSS. In any case, it is perhaps relevant to think critically about the added values (but also the shortcomings) that we expect from DSS. Moreover, it is relevant to ask whether or not the added value expected really addresses a problem that has been observed. For example, is lack of communication speed really a problem in the current scab warning system (using phone and fax) described by Graf *et al.* (this volume), or are we really talking about the timeliness of monitoring? And if so, is the Internet really a solution?

What about the organisational design?

Some papers describe in some detail the way the DSS has been designed in software-technical terms. However, for a DSS to be effective all sorts of organisational arrangements are needed in order to guarantee the use and maintenance of the DSS. Some papers describe how they have tried to minimise the problems of regular maintenance (e.g. Groot, this volume; Boshuizen and Van der Maas, this volume). Little information is however provided

on other organisational issues, such as the need for farmer and extension worker training, support services, financial arrangements concerning maintenance and use, organisational and/or disciplinary frictions during development processes, etc. Yet, one could argue that a viable 'organisational' design is of crucial importance (see also Leeuwis, 1993). Should we pay more attention to these issues? What are important lessons in this respect?

3.4. The quality of the DSS development process

Should we use research models as a basis for DSS development?

Blaise *et al.* and Hardman *et al.* (this volume) indicate that in the longer term the idea is to use these models as a basis for DSS development. First the emphasis is on increasing understanding, and then on application. At the same time, it transpires that such a shift in purpose may require important changes in the model, e.g. the inclusion of economic modes of reasoning in the case described by Blaise *et al.*. The implication here seems to be that growers may pose rather different questions than researchers, and that their considerations in taking decisions may cover a much wider (in the sense of multi-disciplinary) range than is covered by the 'mandate' of the researcher. Likewise, to develop a DSS for fruit thinning in peach, Szafran *et al.* (this volume) use statistical techniques which describe crop physiological responses observed in field trials under a variety of conditions, rather than modelling the physiological processes mechanistically.

A related issue is that during the 'research phase' the complexity of the models may increase continuously (due to the constant stream of newly emerging questions), whereas one can wonder if such complexity is required and/or helpful in case of a DSS. Van der Werf *et al.* (this volume) indicate that in the design process of sampling and monitoring methods, simple descriptive models are used rather than complex explanatory models. The question emerges whether or not it is a wise strategy to sequence DSS development as a follow-up on process-based models.

Do we need more user-participation in DSS design?

While reading the workshop papers, one gets the impression that many of the models are developed in relative isolation within the research community. An exception seems to be the case of IRRY, which has been developed using an interactive prototyping approach (Boshuizen and Van der Maas, this volume). Especially if the idea is to support growers, one could argue that it is essential to communicate intensively with prospective users in the development process, in order to identify needs and problems. Research has shown that information needs tend to be dynamic, and may very well alter and/or become more specific during a prolonged period of software-development and use (Leeuwis, 1993). Hence, discussing the pros and cons of user-participation may be an issue for debate.

References

- Blaise, Ph., Dietrich R. and Gessler C. Vinemild: an application-oriented model of *Plasmopara viticola* epidemics on *Vitis vinifera*. This volume.
- Boshuizen A.J. and Van der Maas M.P. IRRY: a Decision Support System for the water supply in orchards. This volume.
- DeJong T.M. PEACH: Peach crop yield and tree growth simulation model for research and education. This volume.

- Graf B., Höhn H., Siegfried W., Höpli H.U. and Holliger E. Recommendations for an efficient plant protection programme in Swiss apple orchards: Current state and future development of a Decision Support System. This volume.
- Groot M.J. A Decision Support System for economic and ecological calculations for fruit crops. This volume.
- Hardman J.M., van der Werf W. and Nyrop J.P. Modelling mite dynamics on apple trees in Eastern North America. This volume.
- Leeuwis C., 1993. Of computer, myths and modelling. The social construction of diversity, knowledge, information and communication technologies in Dutch horticulture and agricultural extension. Wageningen Studies in Sociology Nr. 36, Pudoc, Wageningen, 468 pp.
- Mols P.J.M. and Boers J.M. Characteristics of a better adapted strain of the parasitoid *Aphelinus mali* (Hald) for the control of woolly apple aphid *Eriosoma lanigerum* (Hausmann) in the Netherlands; a simulation study with a Dutch and a Canadian strain of *A. mali*. This volume.
- Rossing W.A.H., Leeuwis C. and van der Werf W. Quality of Modelling in Fruit Research and Orchard Management; Introduction to the Workshop. This volume.
- Szafran E., Zilkah S. and Kizner Z. Modelling peach response to chemical thinning. This volume.
- Van den Ban A.W. and Hawkins H.S., 1996. Agricultural Extension. Blackwell Science, Oxford, 294 pp.
- Van der Ploeg J.D., 1990. Labor, markets, and agricultural production. Westview Press, Boulder, 313 pp.
- Van der Werf W., Nyrop J.P., Binns M.R. and Kovach J. Computer-methodology for developing pest sampling and monitoring programs. This volume.