Auditing predictive models: a case study in crop growth

Promotor: dr. ir. G. van Straten
hoogleraar in de meet-, regel- en systeemtechniek
Co-promotor: dr. M.J.W. Jansen
biometricus bij het Centrum voor Biometrie Wageningen, CPRO-DLO

K. Metselaar

Auditing predictive models: a case study in crop growth

Proefschrift
ter verkrijging van de graad van doctor
op gezag van de rector magnificus
van de Landbouwuniversiteit Wageningen,
dr. C.M. Karssen,
in het openbaar te verdedigen
op vrijdag 12 februari 1999
des namiddags te 16.00 uur in de Aula

Nederlandse vertaling titel:
Rekenschap afleggen over voorspellende modellen: een voorbeeldstudie over gewasgroei.

Metselaar, Klaas, 1999.

Auditing predictive models: a case study in crop growth. / K. Metselaar.
Thesis Wageningen Agricultural University. -With ref. -With summary in Dutch.
ISBN 90-5485-977-6

NN08201, 2570

Stellingen

1. De discussie over modelkwaliteit is niet tot een bevredigend einde te voeren, tenzij de kwaliteit van de beschrijving van de modelkwaliteit buiten de discussie gehouden kan worden.

2. Een relatie- of Forrester diagram is een model van een simulatiemodel: het gebruik is daarom af te raden.

3. In de tijd discontinue differentie- en differentiaalvergelijkingen, zoals gebruikt in gewasgroeimodellen en stroomgebiedsmodellen, verdienen meer aandacht van toegepast wiskundigen.

4. Er is geen alternatief voor literatuurstudies van parameteronzekerheid, zolang onderzoekers geen formele methoden gebruiken om hun ervaringskennis over parameteronzekerheid vast te leggen.

5. In rassenproeven zou de bloeidatum van snijmaïs standaard waargenomen moeten worden, gelet op de grote bijdrage van deze parameter aan de onzekerheid van de gesimuleerde gewasopbrengst.
   Dit proefschrift

6. Voor een verantwoorde keuze van parameters in calibratie moet het model hiërarchisch gestructureerd zijn, of moeten de onzekerheden van de parameters bekend zijn.
   Dit proefschrift

7. Validatie is een noodzaak.

8. Wie bij de ontwikkeling van een model verzuimt het model goed te beschrijven, de modeleigenschappen te analyseren, de onzekerheden van de parameters te inventariseren en een calibratieprocedure te ontwikkelen, heeft daar in een later stadium de tijd en het geld niet meer voor.

9. De zorgen over de empirische basis van modelonderzoek zoals die uit stellingen in Wageningse proefschriften naar voren komen zijn terecht en worden nog versterkt door het feit dat er binnen het WUR geen formele infrastructuur voor gegevensdocumentatie, -beheer en -hergebruik is, en de informele infrastructuur door de manier van onderzoeksevaluatie (en onderzoeksfinanciëring), personeelsbeleid, fusies en verhuizingen ondergraven wordt of verdwijnt.

10. Bij het beoordelen van publicatielijsten zou het aanbeveling verdienen de impactfactor van een tijdschrift waarin gepubliceerd wordt meer naar beneden bij te stellen naarmate de abonnementsprijs hoger is.

11. Het is vruchtbaar om het ontwerpen van gewassen met behulp van simulatiemodellen (zogenaamde ideotypering) als een optimalisatieprobleem te zien.

12. Het verschil tussen high external input agriculture en low external input agriculture bestaat ook uit het verschil tussen cultuurgewas en cultuur rond het gewas.

13. Informatiemoeheid heeft meer te maken met een overaanbod van kwalitatief slechte informatie en daaruit voortkomende zoek- en selectieproblemen dan met een werkelijke verzadiging van de consument.

14. Het spreekwoord "Vogeltjes die vroeg fluiten zijn voor de poes" moet vanuit een evolutionair oogpunt worden gerelativeerd. In het Engels (the early bird catches the worm) wordt juist het evolutionaire voordeel benadrukt.

Stellingen behorend bij het proefschrift van K. Metselaar
Auditing predictive models: a case study in crop growth
Wageningen, 12 februari 1999.

Abstract

Metselaar, K., 1999. Auditing predictive models: a case study in crop growth. PhD-thesis, Wageningen Agricultural University, Wageningen, The Netherlands, 265 p., English and Dutch summaries.

Methods were developed to assess and quantify the predictive quality of simulation models, with the intent to contribute to evaluation of model studies by non-scientists. In a case study, two models of different complexity, LINTUL and SUCROS87, were used to predict yield of forage maize under Dutch meteorological conditions. The models predict yield under potential conditions, i.e. temperature- and radiation limited yield, assuming other production factors to be optimal.

After a review of concerns voiced in model-based applied research, the simulation models were described in a systematic manner to simplify access to the software code. A model analysis showed that the models contain switches, describing abrupt changes occurring in the crop (e.g. change of temperature driven leaf area growth to photosynthesis driven leaf growth; onset of leaf senescence). Some switches introduced discontinuities in the relation between state variables and parameters. Such properties make non-standard approaches for parameter estimation necessary.

Subsequently, the empirical basis of the simulation model was reviewed in terms of parameter values and their uncertainty, as derived from literature. The results were used to evaluate the predictive quality given the parameter uncertainty. Predictive quality given the parameter uncertainty was low; parameter estimation to adapt the model to local conditions was necessary.

Different procedures to calibrate the models were discussed and presented. For the combination of models and the data available in this case study, parameters had to be selected. Selection was based on the ranking of the parameters on the basis of their contribution to output uncertainty. Non-selected parameters were fixed at their default value. Calibration using a controlled random search algorithm for a point estimation procedure was executed for both models. In the estimation procedure a compromise was sought between different types of problems: estimation bias, parameter identifiability and local minima.

The parameter estimates were used to generate predictions. A comparison between predictions and measured data was used to evaluate the predictive quality of the models in terms that are relevant for the application. To do so, the concept of a link hypothesis was introduced. It defines the anticipated relation between prediction and measurement. Deviations from the anticipated relation were used to quantify predictive quality. Predictive quality was shown to depend strongly on the procedure used to generate predictions, i.e. procedures combining results based on multiple calibration sets yielded better predictions than predictions based on a single data set.

To translate predictive quality in terms of usefulness of the simulation model prediction errors were compared to those of benchmark predictors (simple statistical predictors). LINTUL and SUCROS87 differed in their performance in relation to the benchmark predictors.

Procedures developed in this thesis suggested that facilitating model evaluation requires actions that are not easily executed within the context of project-based, often time-limited, applied research. Investment in the methodological basis and in the empirical basis of the models prior to their application will be required.

Additional index words: elicitation, uncertainty analysis, sensitivity analysis, manual point (trial and error) calibration, Price algorithm, validation, prediction error, *Zea mays L.*, Netherlands, forage maize, potential yield, photosynthesis, maintenance respiration, growth respiration, partitioning, phenology, light use efficiency, light interception, senescence, specific leaf area.

Voorwoord

Een aantal mensen waren nauw bij dit onderzoek betrokken en hen wil ik graag bedanken. Allereerst de begeleidingscommisie, bestaande uit Jos (A.A.M.) Jansen, mijn copromotor Michiel (M.J.W.) Jansen, Pavel Kabat, Frits Penning de Vries, en mijn promotor Gerrit van Straten. De in dit proefschrift gebruikte versies van de modellen SUCROS87 en LINTUL werden beschikbaar gesteld door de vakgroep Theoretische Productie-Ecologie en het AB-DLO. Graag bedank ik de vakgroep en het instituut voor deze essentiële bijdrage aan dit onderzoek. Rond de modellen was door Daniel van Kraalingen, Cees Rappoldt en Willem Stol al een omgeving ontwikkeld waarbinnen de gepresenteerde procedures uitgevoerd konden worden. Deze was onmisbaar. Met name wil ik verder Klaas Scholte en Ad Schapendonk bedanken die experimentele gegevens ter beschikking gesteld hebben. Naast de formele begeleidingscommissie, de mensen achter de modellen, de mensen rond de modellen en de mensen achter de gegevens was er een schaduwbegeleidingscommissie. De harde kern van deze informele begeleidingscommissie bestond uit Jet Drenth, Hans Heesterbeek en Gon van Laar. De volledige informele begeleidingscommissie was veel groter, van wisselende samenstelling en gaf eigen bijdragen vanuit hun eigen invalshoek en op zeer diverse gebieden. Onder hen natuurlijk collega's en ex-collega's van AB, GLW en SC, de vakgroepen Agrotechniek en Agrofysica en Theoretische Productie-ecologie, mede-AIO's en OIO's, grootouders, ouders, familie en vrienden, mijn huisgenoten en hun kinderen, en de eetclubgenoten.

Zonder de formele en informele bijdragen, vanuit verschillende invalshoeken, op verschillende terreinen en op eigen manier gegeven en vormgegeven, had dit proefschrift er anders uitgezien, en was er misschien niet gekomen. Dank allen!

Een proefschrift is dan in naam het werk van een individu, maar wordt uitgevoerd binnen een structuur waarin veel mensen een steentje bijdragen, en het grootste gedeelte anoniem blijft. Dit proefschrift is daarom opgedragen aan een gedeeltelijk anoniem persoon: Teichler, Königlicher Obergärtner, een man die in het Duitsland van 1876 wel een functie, maar geen initialen had. Daarmee wil ik tenslotte iedereen bedanken die bewust of onbewust, formeel of informeel aan dit onderzoek heeft bijgedragen. De literatuurlijst geeft een deeloverzicht van de formelere bijdragen.

Table of contents

Chapter *1*

Introduction

1.1 Introduction

Research is an activity directed towards the extension of certified knowledge (Philip, 1991). Certified knowledge should be understood as logically consistent descriptions which allow predictions that can be and have been empirically confirmed. Research results are often presented as mathematical equations interwoven with a descriptive text. In present day research, these equations are then programmed to yield simulation models. The research process which results in simulation models is referred to as simulation modelling. Since the late 60s (e.g. de Wit, 1968) computer-based modelling has been an approach used in agricultural research.

Thoughts about development and use of simulation models in specific branches of agricultural research are persistent elements in inaugural addresses at the Wageningen Agricultural University (some examples: Rabbinge, 1985; Feddes, 1990; Stroosnijder, 1991; van Straten, 1991; Fresco, 1992; Goudriaan, 1993; Bot, 1994; van Keulen, 1995; Kropff, 1996; Leijnse, 1996; Oenema, 1996). One may conclude that models constitute a methodological focus in applied agricultural research. In a different formulation: Models are attractive tools to help thinking about the possible in an organized way (de Wit and Penning de Vries, 1985). However, the use of simulation models in agricultural research generates new research questions.

1.2 A systems view of model use in agricultural research

Agricultural research can be thought of as a system in which problems are input, a specific technology is used, and contributions to the problem solution based on this technology are output. Modelling and model use is regarded as a possible technology to yield these contributions. In this thesis, computer based modelling will be the only technology considered. A symbolic representation of the conceptual system which is at the heart of a large number of projects in model-based agricultural research is presented in Figure 1.1 (adapted from Nance and Balci, 1987). This representation of the research system is discussed by e.g. Baker and Curry (1976) and Rao et al. (1989) for agro-ecosystems, and by Shaeffer (1980) for environmental assessment models. Furthermore it is discussed in general terms by Banks et al. (1988) and in the context of operations research by Gass (1983). The research system contains subsystems: the problem definition, model development, evaluation and application phase, which can be further described in terms of their components and interactions (discussed in Sections 1.2.1 and 1.2.2).

There are several reasons to consider modelling to be a suitable tool in agricultural research. Modelling allows to describe and integrate knowledge regarding complex systems. Modelling requires that all assumptions, inputs and results are made explicit and quantified, and thus allows to make them subject of discussion. This discussion can allow to set research priorities, or may contribute to other policy decisions. Models can be characterized as applicable and quantitative research reviews, allowing quantifiable insights. These characteristics allow to use models as a means of knowledge transfer e.g. for training, education and decision-support systems (Penning de Vries et al., 1988; ten Berge, 1991; Uehara and Tsuji, 1991). Model-based decision-support systems are thought to become important in future developments in agriculture (Plucknett and Winkelmann, 1995).

1.2.1 Certification of model-based agricultural research

Agricultural research is applied research, where applied indicates that the model-based research system as sketched in Figure 1.1 is embedded in an agricultural production system.

All agricultural research has to be certified twice: the first time in a critical evaluation as scientific research; the second time under different conditions and in a different social context, by research clients. In Figure 1.1, research results are input to an evaluation phase, with construct evaluation as its main component. A construct is any artificial object which is transferred from one social context to another. Evaluation is defined as a process by which interested parties, who were not involved in the construct's origins, development and implementation, can assess the construct in terms of its structure, inputs and outputs so as to determine, with some level of confidence, whether and to what extent it can be used for a given objective or application (Henize, 1984).

Depending on differences between the research context and the context of application, the construct sometimes has to be reworked in a dialogue between researchers and clients. The role of models in applications is, therefore, supportive rather than prescriptive (van Latesteijn and Rabbinge, 1992; Leutscher, 1995; Smit, 1996). Constructs which pass evaluation are certified for application.

An example of actions in the evaluation phase is the extensive set of interviews executed to evaluate an agricultural decision-support system for advice on spraying to control various pests and diseases in wheat (Blokker, 1984). This evaluation lead to a modification of the decision-support system. Other examples which illustrate how models are evaluated outside the research context concern a model for prediction of aquifer pollution (Bair, 1994), a model to analyse air pollution used for policy implementation (de Nevers, 1973), and an evaluation of a model for water allowances in irrigated agriculture following an erroneous prediction (Glantz, 1982).

Models are also used to investigate policy options in so-called scenario studies. These scenario studies are subject to an evaluation phase. An investigation of development options in agriculture for the European community (Rabbinge and van Latesteijn, 1992; Wetenschappelijke Raad voor het Regeringsbeleid, 1992), led to an extensive discussion of the results (van Latesteijn and Rabbinge, 1992).

In these different discussions of model or scenario studies information is required e.g. regarding the choice of the model, the performance of the model, the arguments for the assumptions made, and the parameter values used. In all cases, the questions generally focus on the quality and credibility of the steps leading to model application.

For models and model-based research to be also credible in a different social context, a necessary condition is to make sure that all steps towards application can be evaluated and all information from these evaluations is available. It should be stressed that we already suppose that the interested and concerned parties have access to the evaluation and to the problem definition phase.

Providing the information requires a research protocol and evaluation methods for each of the research components. The necessity of systematically providing information for construct evaluation is the main motivation for the research in this thesis.

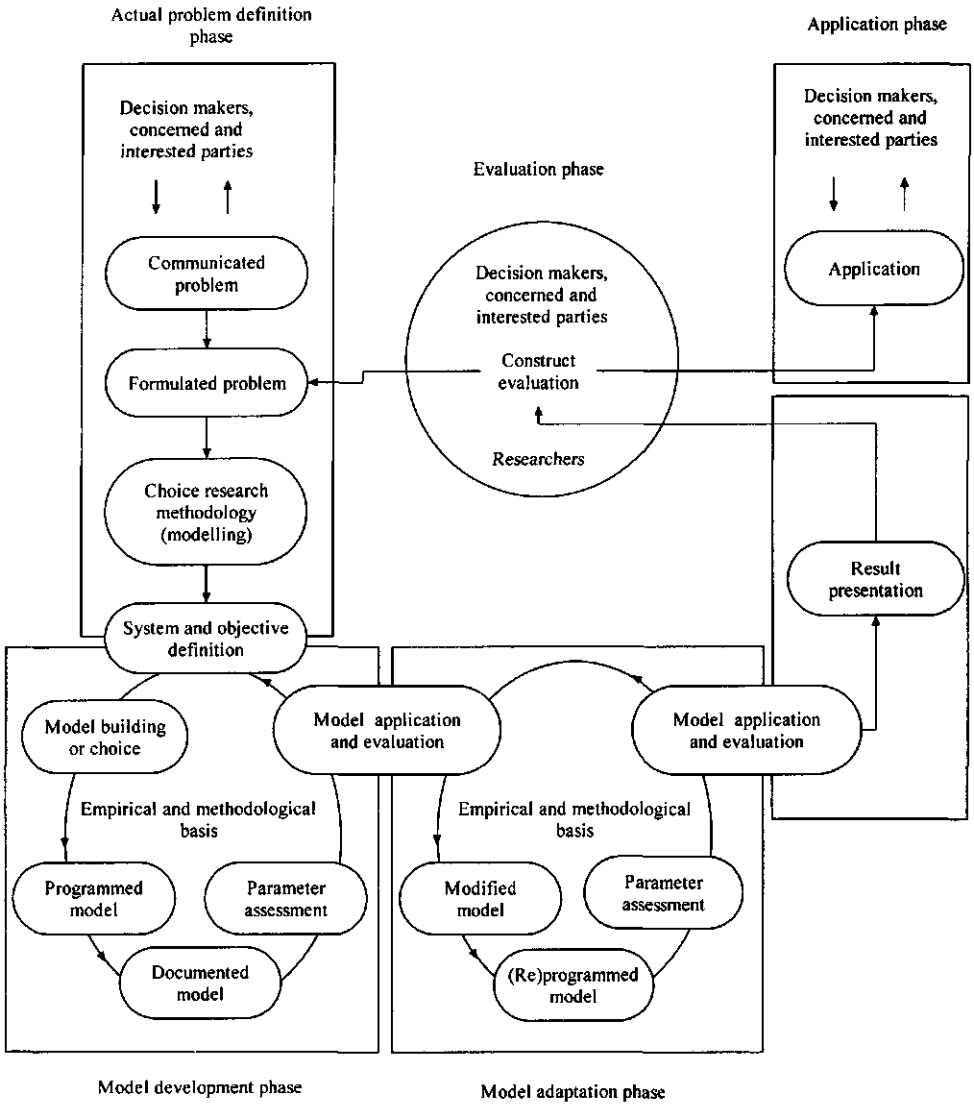

Figure 1.1 A view of components in model-based applied research.

1.2.2 The components of model-based research

Even if properly embedded in the agricultural production system, model-based agricultural research can only be successful if each of the subsystems, components and their interactions function properly. As the advantages of a model-based approach have been presented, an inventory of opinions regarding model-based research will serve to evaluate concerns in model-based research up to the present. This inventory will indicate possible topics of interest for research. A number of the opinions presented here are derived from the so-called 'stellingen' (propositions) associated with PhD-theses at Wageningen Agricultural University. When referring to a proposition the number is indicated in brackets ([]). The propositions are associated with theses in which model-based research was executed. The resulting overview is anecdotical and certainly not exhaustive.

The interaction between research components

Interaction between the different research subsystems and system components is brought about by the researchers. This requires that researchers are available who are well versed in combinations of different skills (Penning de Vries and Spitters, 1991), or that specialists are able to work together in a multi-disciplinary context.

Jansen et al. (1995) are concerned about the number of people who can effectively use quantitative methods (mathematics, statistics and operational research) and those who can envisage the potential applications of methods developed by 'pure' mathematicians. These skills are required in a large number of research system components such as model building, parameter assessment, model modification and model application and evaluation. Whereas Jansen et al. (1995) worry about the use of conceptual models with little support by quantitative methods, Klemeš (1986) worries about the use of quantitative methods in hydrology with little support by conceptual models. Veldkamp ([8], 1991) suggests that in earth sciences excessive priority is given to supporting sciences (statistics and informatics). To improve the situation, Jansen et al. (1995) propose to re-define education goals in quantitative methods, whereas Klemeš (1986) strongly advises to strengthen the conceptual basis of hydrology. This suggests that researchers are over-specialized or over-specializing.

Acquiring skills both in empirical research and in modelling is also regarded as necessary. Stuyt ([11], 1992) states that learning about the facts through empirical observation is forgotten. Philip (1991) characterizes a situation as follows: "Most raw PhD's seeking a job at our Centre proudly present themselves as computer jockeys, incurious about real-world phenomena and innocent of laboratory and field skills, yet blissfully unaware of their inadequacy for serious research." Bastiaans ([5], 1993) suggests that there is a lack of formal training in combining modelling and empirical research. The empirical basis of model concepts is sometimes weak, as voiced by several authors (Monteith, 1981; Hillel, 1987; Philip, 1991; Querner, [12] 1993). Smit ([1], 1996) asserts that knowledge of a crop is more often than not insufficient for a quantitative simulation model. Monteith (1981) proposed a moratorium on modelling to allow experimental research to catch up with the sophisticated modelling concepts used.

Notwithstanding that some of the above statements may be put somewhat more subtle,

concerns are expressed about the development and combination of quantitative methods, modelling and experimenting skills. This concern was also found in an inventory of problems in agricultural modelling in Dutch agricultural research (Verkenningscommissie, 1986).

The proposals to alleviate this concern focus on education. Strict adherence to de Wit's maxim "No simulation without experimentation" (Bastiaans, [5] 1993) and its reverse "No experimentation without simulation" (Leffelaar, [3] 1987) serves the purpose of more closely integrating experimentation and simulation based research. To alleviate the concerns presented in this section, skills in simulation must be combined with skills in quantitative methodology.

Individual research components: Definition of system and objective

Systems and objectives in model-based research are diverse. In the previous section we already introduced hydrology and earth sciences as fields in which models are used, and distinguished between different objectives: models used for training, education, and for decision support, models used for long-term policy analysis, and models for environmental assessment. The research objective has important consequences. This is very clear if we compare a model developed to be used in education, a model developed for yield prediction and a model developed to test a scientific hypothesis. Whereas the system considered may be identical, the level of detail required in model building, the model documentation required, and the criteria used for model evaluation change completely (Goudriaan, 1996b).

Individual research components: The empirical and methodological basis

Model-based research requires information regarding model inputs (forcing functions, such as weather variables, and parameter values).

The required input information is not always available (van Lanen, [8] 1991). This can be regarded as a shortcoming of the empirical basis; alternatively it can be regarded as a shortcoming in the model formulation, which requires model adaptation (Bouman, [1] 1991). In some cases the input information is not available for the scale on which the model has to be applied (Wopereis, [5] 1993). There is need for standardized datasets for use in simulation, either at a single accessible location or within a world-wide network (Smit, [10] 1996; Wopereis, [8] 1993).

The methodological basis contains the tools required for model application. Tools to check unit consistency, numerical integration procedures for stiff differential equations, and tools for mathematical manipulation of equations, as required by Goudriaan ([8], 1977), have been developed or are being developed. More recently (Querner, [6] 1993) noted the need for a method to schematize regional hydrological systems as an alternative for the expert judgement used.

In literature some remarks are made regarding the technical accessibility of the empirical and methodological research basis. Mous ([4], 1994) suggests that the accessibility of numerical and statistical software (the methodological basis) can be improved by incorporating them in the modelling environment. Rappoldt ([6], 1992) submits that modelling languages used may not allow interfacing with numerical and statistical software. These concerns show that technical

interactions between research components could be improved.

In both the empirical and methodological basis suggestions for improvement concern technical accessibility and standardization.

Individual research components: Model building and model choice

Ideas regarding the choice or building of a model can be described in terms of the properties of an ideal model. Although little theoretical methodology exists to assist in the formulation of model concepts (van Straten, 1991), some criteria are suggested.

According to Wösten ([2], 1990) and Smaling ([4], 1993) an ideal model requires relatively few input variables, which are easily and cheaply measured. The number of uncertain initial conditions used in a model should be minimal (Meinke, [4] 1996).

Conceptual content of the models is discussed in terms of complex vs. simple models. The optimal complexity for predictive use is subject of discussion (Botkin, 1977; Beck, 1981; Håkanson and Peters, 1995). It is generally assumed that an optimal model complexity for a specific system and objective exists. Discussing crop models, Whisler et al. (1986) conclude that data requirements, running time, and the number of possible sources of error increases with complexity. Prediction results suggest that conceptually complex models are not better predictors than simple models (Rogers, 1978; Versteeg and van Keulen, 1986; van Grinsven et al., 1995). Model conceptual content should be determined by the problem studied (Schouwenaars, [1] 1990; Whisler et al., 1986) or by the model's applicability (Querner, [4] 1993).

Simplicity and elegance (Meinke, [1] 1996) or beauty (Heinen, [12] 1997) are also suggested as criteria to use in model formulation. Simplicity is required to keep the model accessible and applicable (Smit, [9] 1996; Bouman, [9] 1991). A different conceptual requirement is time-invariance of the parameters. The search for time-invariant parameters has been a matter of explicit concern in model formulation (Meinke, 1996; Bastiaanssen, 1995).

Sheng ([4], 1994) states that models should be judged on utilitarian, rather than aesthetic, criteria. This discards elegance, beauty, and probably simplicity as criteria for model formulation. One such utilitarian judgement criterion is the prediction quality (Wösten, [2] 1990; Jansen et al., 1995).

The above suggests that developing the ideal model is a trial and error procedure in which input requirements, conceptual content, predictive quality, and aesthetic properties are simultaneously optimized. Ranking these goals as to their relative importance for specific objectives is essential; e.g. the maximization of conceptual content is a secondary goal for applied research (Schouwenaars, [2] 1990). Simplicity is an important criterion for the development of models for training and education (Goudriaan, 1996b), whereas elegance and beauty are often used to characterize highly formalized mathematical treatises. Given the diversity of criteria and the variation in their importance given the research objective, it is difficult to conceive how model development could be formalized.

Individual research components: The model program and model documentation

The presentation and documentation of models is cause for concern (Verkenningscommissie,

1986). Assumptions are not stated or are not well substantiated (Stroosnijder, [5] 1976; Miglietta, [10] 1992). Numerical procedures are not documented (van Dorp, [6] 1977). Processes are not easily recognized within the software code (Leffelaar, [4] 1987) or models are sometimes only presented as a software code (van Henten, [7] 1994; Boesten, [8] 1986). These problems give rise to projects for software certification and standardization to ensure software quality (Waldman and Rickman, 1990; van Stijn et al., 1994; Wijngaard et al., 1995).

Individual research components: Parameter assessment

Parameter assessment is the research component in which values are assigned to different parameters in the model. In the model development phase one has to assign values to model parameter to be able to use the model. This is not straightforward: parameter values are not precisely known, and vary between different situations.

From a model evaluation study of crop growth models, Goudriaan (1996a) concludes that it is necessary to assign new parameter values if the model is used in a different environment. Given this conclusion, a large number of research questions and problems is introduced: e.g. what constitutes a different environment, or which parameters - out of the large number used - should be assigned new values.

Parameter values can be determined from experiments using a part of the simulation model or the entire model. A view of the practice of estimating parameter values from measurements using the entire simulation model (calibration) is presented by Janssen and Heuberger (1995). They suggest that some of the methods used could suffer from lack of exactness, reproducibility and objectivity. Ideally the quality of the parameter estimation process should be explicitly quantified by determining the precision with which the parameters have been estimated from the data.

A problem in parameter estimation for complex models is the large number of parameters which are potential candidates for estimation (Philip, 1991; de Wit, 1970). In a comparison of calibration executed for a number of models for potato yield prediction, Kabat et al. (1995) note that for all models a selected subset of all parameters was estimated using a number of methods. Wösten ([8], 1990) submits that the assumption that selected parameters are the only parameters whose value is unknown is not generally justifiable. This introduces the problem which criteria for parameter selection to use.

Given the experimental data some parameters may be estimated, whereas others may not. Parameter estimation and experimental design interact, and the experimental design may well determine which parameters can be estimated (Mous, [1] 1994). This is an additional complicating factor in parameter selection.

The question is to which extent parameter assessment can be formalized, given these problems of parameter selection and determining the quality of parameter estimation.

Individual research components: Model evaluation

Model evaluation is the research component in which the quality of the model for a specific objective is established. Criteria for model evaluation are different for different objectives.

If the objective is predictive application, predictive quality is an important aspect of model

evaluation. Most discussions of model quality focus on this aspect. According to Jansen et al. (1995) predictive quality is an essential component in model-based research, as "Results without a degree of accuracy are worthless." Prediction quality can be established either by quantifying the prediction uncertainty, or by comparing model prediction and observed values.

Quantification of output uncertainty is regarded as an indispensable element in construct evaluation (Kropff, [9] 1989; Hordijk, 1993; Jansen et al., 1995). However, little attention is paid to error analysis or output uncertainty (ten Berge, [8] 1986; Bouman, [9] 1991).

Comparing model results and measurements or observations is a major part of construct evaluation. Concern is voiced about the fact that a model is only evaluated on the basis of a limited number of output variables (Kroon, [3] 1985). For a specific application this output variables might be less relevant. Care should be taken that the output variable conceptually corresponds to the variable measured (Kim, [3] 1995). Other problems are encountered in the evaluation of complex models, where specific subprocesses can not be tested at the level of the model output, because of the system's complexity (Klepper, [1] 1989). Verboom ([5], 1996) submits that complex models can not be evaluated in terms of their predictive quality, but only in terms of comparative usefulness.

1.2.3 Summary of the evaluation of model-based research

Model-based research generates a number of research questions and concerns. On one hand the evaluation shows a need for a set of qualitative methods regarding model presentation, documentation and the integration of the different research system components. These methods should result in standards for model presentation and documentation. Integration of the different research components requires communication between specialists, or execution of research by generalists. This is a problem of training and education.

On the other hand, a need for quantitative methods for parameter assessment and model evaluation exists. The methods for parameter assessment should provide criteria for parameter selection in parameter estimation. Parameter estimation methods should yield results regarding the parameter estimation error. Methods for model evaluation should yield a measure of predictive quality, and clarify the conditions which should be met for model evaluation to be possible.

Finally, the inventory showed that both the qualitative and quantitative methodological basis for the development of models which are best for a specific application is rudimentary.

1.3 Objectives and approach

The objective of the research presented in this thesis is to develop methods to assess and quantify the predictive quality of simulation models and to improve that quality.

The research in this thesis is restricted to the development and application of quantitative methods for the three research components 'parameter assessment', 'model application' and 'model evaluation'. All components occur both in the model development phase and the model adaptation phase (cf. Figure 1.1). The notion 'model auditing' is defined as a critical evaluation

of these three selected components executed with the intent to contribute to construct evaluation.

Given the diversity of subject fields, applications and objectives, the research is necessarily executed as a case study. In the case study we will be using two models, SUCROS87 (Spitters et al., 1989) and LINTUL (Spitters and Schapendonk, 1990), for the prediction of potential maize yield. SUCROS87 is a more detailed model in comparison to LINTUL (Spitters, 1990). This could shed some light on the question regarding the desired model complexity for predictive application.

In order to achieve the objective the following tasks have to be executed:

- Describe the models and analyse their structure;
- Assess parameter values used in the models;
- Assess the credibility of the model predictions given the information regarding the parameters;
- If necessary re-assess parameter values on the basis of additional information;
- Re-assess the credibility of model predictions;
- Summarize and review the results.

1.4 Outline of the thesis

The outline of the thesis is given in Figure 1.2. The starting point of the thesis is the software code of the models SUCROS87 and LINTUL. The programs are described in Chapter 2. The analysis in Chapter 3 focuses on structural properties of the two models. The guiding question is whether the structure of the model has consequences for re-assessing parameter values. In Chapter 4 the model parameters are subject of research to yield the data necessary for a credibility assessment of the model for prediction purposes. The credibility of the model is then assessed in Chapter 5, using uncertainty analysis. This analysis serves a dual purpose. It allows to conclude whether the prediction quality of the model using present knowledge of the parameter values is sufficient for the application envisaged. It also allows to propose a methodology to select parameters for parameter estimation in the adaptation phase. Estimating parameter values of the model from field experiments (calibration) is executed in Chapter 6. A credibility assessment through prediction (validation) using the adapted model is then made in Chapter 7. Chapter 8 reviews the chapters in relation to each other and summarizes the findings.

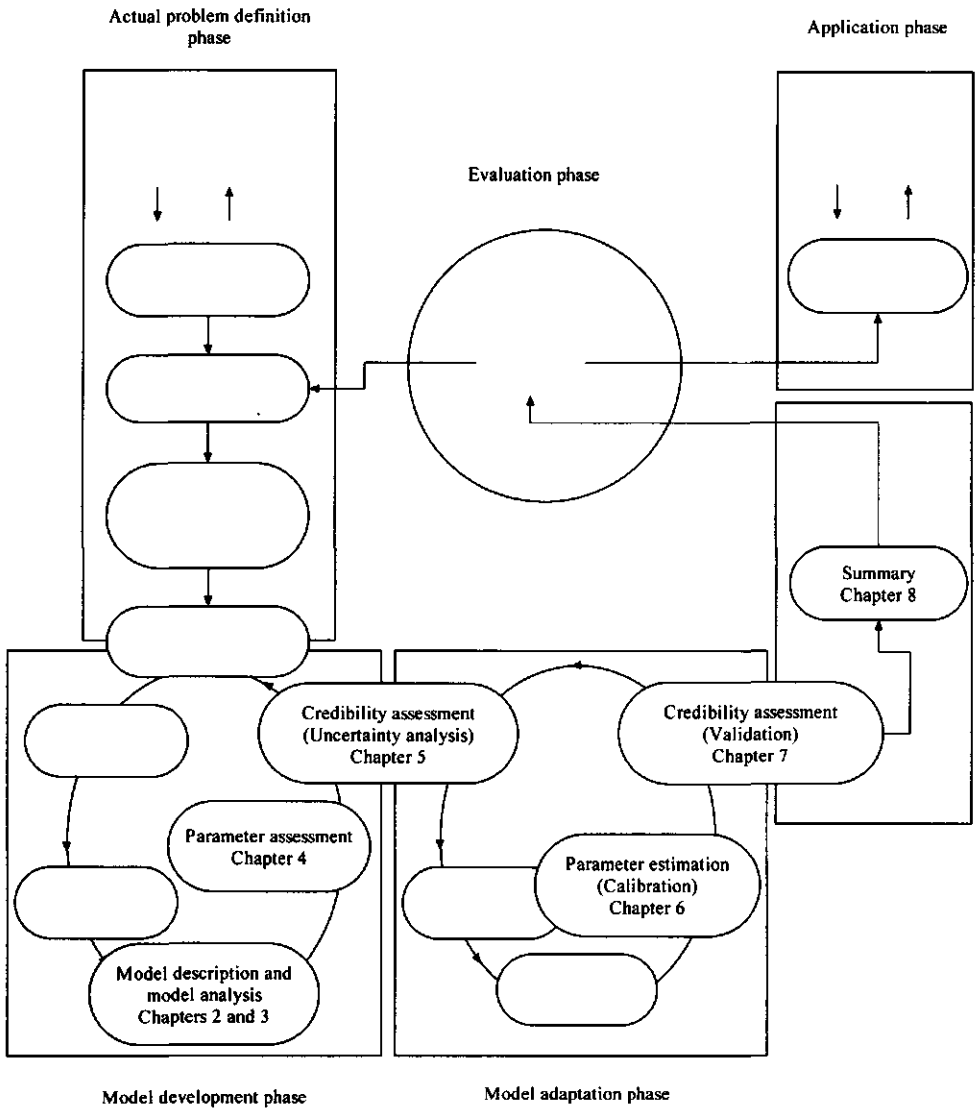

Figure 1.2 Outline of the thesis.

Chapter 2

Model description

2.1 Introduction

2.1.1 Options for description

In this case study the models SUCROS87 and LINTUL will be used. As stated in Chapter 1 the models were chosen because of their difference in complexity. The model SUCROS87 is described earlier by Spitters et al. (1989). The model LINTUL was derived from the model LINTUL developed for potatoes (Kooman and Spitters, 1995) and modified for maize for use in this thesis. LINTUL was described earlier by Spitters and Schapendonk (1990). The models calculate temperature- and radiation-limited production for an annual crop. This production is also called potential production: all other production factors (water, nutrients, management,) are assumed to be non-limiting. Damage by pests, diseases and weeds is assumed to be absent due to control measures. Both models are available as a computer program written in FORTRAN.

To present the models to an interested or concerned audience the models have to be described. A description is a translation from a specialist language to a more common language: in this case a model written in FORTRAN has to be translated. Model descriptions in scientific literature are often a combination of a formal language (mathematics) and natural language (English), as e.g. in the references quoted earlier. This combination is necessary as understanding theory is not possible without motivation and explanation in a natural language, whereas only formal languages (logic or mathematics) allow systematic analysis and exploration of the consequences of a theory (Bruggeman, 1996).

In this thesis the purpose is to describe the models in such a way that the description can be used as input to construct evaluation. This does not require a rigorous formalization. It does, however, require a structured description, which improves access to the prime source of both models, in this case the FORTRAN source codes. As in programming a topdown stepwise refinement procedure is often used, we will use the same procedure for model description, starting with the state variables used in the models. These considerations constitute the basis for the descriptions in this chapter.

2.1.2 Notation

The description given here is based on the computer algorithms. To allow easy reference to the original programs it was decided to retain the original acronyms used in the model program, but to incorporate prefixes to indicate the function of the acronym in the model. Different model elements are designated in the following way:

S_	: State variables	I_	: Input
R_	: Rate variables	P_	: Parameter
F_, G_	: Auxiliary calculations	L_	: Logical switch
T_	: Tabulated function	O_	: Output variable

In Appendix 1 (SUCROS87) and 2 (LINTUL) all acronyms with description, units and type are given. Inputs and states have the standard argument DAY; this is left out. As in computer programs the '=' sign signifies 'becomes'.

Characteristic for the implementation are look-up tables T_ and IF-THEN-ELSE structures. To allow a short description of IF-THEN-ELSE structures we introduce switches L_i. These switches are functions which can have the value 1 or 0. As an example, the structure:

```
IF (S_DVS.LT.P_DVSJUV.AND.S_LAI.LT.P_LAIJUV) THEN
   R_GLAI( ) = F_LAJUV( )
ELSE
   R_GLAI( ) = F_LAMAT( )
ENDIF
```

becomes

$$R\_GLAI( ) = L\_7( ) * F\_LAJUV( ) + [1 - L\_7( )] * F\_LAMAT( )$$

where

$$L\_7( ) = \begin{cases} 1, & \text{if (S\_DVS.LT.P\_DVSJUV.AND.S\_LAI.LT.P\_LAIJUV)} \\ 0, & \text{OTHERWISE} \end{cases}$$

Rewriting an IF-THEN-ELSE statement in a single equation and a logical switch is not perfectly identical: in the IF-THEN-ELSE statement statements are skipped, whereas in the single equation all elements are evaluated.

In this model description the brackets '( )' refer to the listing of arguments given in Appendix 3 (for SUCROS87) and Appendix 4 (for LINTUL). In these appendices the models are described in terms of argument trees, starting with the state variables and ending with the parameters, state variables and inputs which are used in their calculation.

Common characteristics of the models

Both models are written as finite difference equations. The timestep used for the rectangular integration is fixed at one day. The model calculations are repeated within a loop; this loop usually starts at sowing or at emergence. If activated - at crop emergence - the initialization section is used once. Under different conditions a different initialization section is used in which different initial values of the state variables are input. For normal termination of crop growth this loop ends, either if the crop is mature, or at harvest.

The calculation of any state variable in SUCROS87 and LINTUL can be written as:

$$S\_X(d) = S\_X(d-1) + R\_X[F\_X( ), S\_(d-1), I\_(d)]$$

where S_X(d) is the value of the state variable at the end of the regarded day d. The variable

R_X yields the increase in variable S_X over day d. S_(d–1) are the values of the state variables at the end of the previous day (d–1), which are used in the calculations; I_(d) are the inputs for day d; inputs are daily values. In the calculation of the rates auxiliary variables F_X( ) are used. This distinction between model elements introduces a hierarchy as follows: S_X depends on R_X; R_X depends on F_X; F_X depends on other auxiliary variables, which finally only depend on other state variables S_, parameters P_ and inputs I_. The model descriptions are based on a topdown hierarchical description of these different dependencies.

2.2 Description of SUCROS87

The SUCROS version (SUCROS87) and the parameters presented here are specific for a maize crop. SUCROS87 uses input data on a daily basis and simulates the dry matter accumulation of several organs. The dry matter accumulation of an organ is a resultant of different processes - assimilation, maintenance- and growth respiration, senescence, and partitioning - which are controlled by temperature.

The description proceeds in the following order: state variables, rate variables, auxiliary variables, logical switches, initialization section and operations on state variables. The description stops at the level of detail at which the subroutines for assimilation are called. In these subroutines hourly rates are integrated and daily rates are output to the main program. For their description the reader is referred to Spitters (1986), Spitters et al. (1986), Goudriaan (1988), and Goudriaan (1986). Acronyms, their description, units and type are given in Appendix 1.

2.2.1 The inputs

Inputs to SUCROS87 are daily global radiation (I_RDD), daily maximum temperature (I_TMMX) and daily minimum temperature (I_TMMN). The temperatures are used to calculate the daily average temperature (I_DAVTMP) and the daily daytime average temperature (I_DDTMP). The daily global radiation received at the top of the atmosphere (I_DS0) is calculated using the function ASTRO.

Daily input:

I_TMMN, I_TMMX, I_RDD

Calculations:

I_DAVTMP = (I_TMMX – I_TMMN)/2

I_DDTMP = I_TMMX – 0.25 * (I_TMMX – I_TMMN)

I_DS0 = F_ASTRO(P_LAT, S_DAY, P_SC)

2.2.2 The state variables

The different organ dry matter weights are regarded as state variables. The following state variables are distinguished:

WRT = Root dry matter weight
WLVG = Green leaves dry matter weight
WLVD = Dead leaves dry matter weight
WST = Stem dry matter weight
WCB = Cob dry matter weight
WSO = Grain (Storage organ) dry matter weight

Other state variables are:

DAY = Day of the year
DVS = Crop development stage
EMERG = Temperature sum for the crop to emerge
TSUMEM = Temperature sum controlling early leaf area development
LAI = Leaf area index

The temperature sum to emergence (EMERG) and the crop development stage (DVS) determine the timing of events during crop growth. DVS equals 0 at emergence; 1 at flowering and 2 at maturity. TSUMEM determines leaf area growth in an early phase. All three state variables have temperature as only input. The state variable DAY is used as a counter, associated with the control of in- and outputs. The basic equations used in SUCROS87 are:

S_WRT = S_WRT + L_2( ) * L_3( ) * R_GRT( )

S_WLVG = S_WLVG + L_2( ) * L_3( ) * [R_GLV( ) - R_DLV( )]

S_WLVD = S_WLVD + L_2( ) * L_3( ) * R_DLV( )

S_WST = S_WST + L_2( ) * L_3( ) * R_GST( )

S_WCB = S_WCB + L_2( ) * L_3( ) * R_GCB( )

S_WSO = S_WSO + L_2( ) * L_3( ) * R_GSO( )

S_DAY = S_DAY + 1

S_DVS = S_DVS + L_2( ) * L_3( ) * R_DVR( )

S_EMERG = S_EMERG + L_1( ) * [1 - L_2( )] * R_DEMERG( )

S_TSUMEM= S_TSUMEM + L_2( ) * L_3( ) * R_DTEFF( )

S_LAI = S_LAI + L_2( ) * L_3( ) * R_GLAI( )

where L_1( ), L_2( ), and L_3( ) are switches described in Section 2.2.5.

2.2.3 The rate calculations

The rate variables R_X( ) used in the calculation of the state variables are the result of complex calculations using parameters, auxiliary variables, inputs, and state variables.

R_GRT = [1 – T_FSH( )] * F_GTW( )

R_GLV = T_FLV( ) * T_FSH( ) * F_GTW( )

R_DLV = L_4( ) * F_RDR( ) * S_WLVG

R_GST = T_FST( ) * T_FSH( ) * F_GTW( )

R_GCB = T_FCB( ) * T_FSH( ) * F_GTW( )

R_GSO = [1 – T_FLV( ) – T_FST( ) – T_FCB( )] * T_FSH( ) * F_GTW( )

R_DEMERG= MAX [0, [I_DAVTMP( ) – P_TBSEM]]

R_DVR = [1 – L_8( )] * T_DVRV( ) + L_8( ) * T_DVRR( )

R_DTEFF = L_7( ) * MAX [0, [I_DAVTMP( ) – P_TBSJUV]]

R_GLAI = L_7( ) * F_LAJUV( ) + [1 – L_7( )] * F_LAMAT( )

where L_4( ), L_7( ) and L_8( ) are switches described in Section 2.2.5. At this point the description is complete for the rate variables R_DTEFF and R_DEMERG. Other rate variables require auxiliary variables for a complete description.

2.2.4 Auxiliary variables in the rate calculations

The tabulated functions

The different tabulated functions (T_FSH, T_FLV, T_FST, T_FCB) used in partitioning the daily growth over the different crop organs are presented in Figure 2.1. In SUCROS87 (Spitters et al., 1989, page 178) leaves are defined to be leaf blades; stems are defined to be leaf sheaths and stems; grains are regarded as a separate storage organ. The partitioning to the organ 'cob'

can be calculated from the requirement that the sum of all partioning functions must equal 1. These definitions of the other organs defines the organ 'cob' to be a heterogeneous organ, consisting of husks, shank, aborted cobs, and grains which can not be separated from the cob. The tabulated functions used in the calculation of the development rate (T_DVRV, T_DVRR, identical in the nominal parameterization) are presented in Figure 2.2.

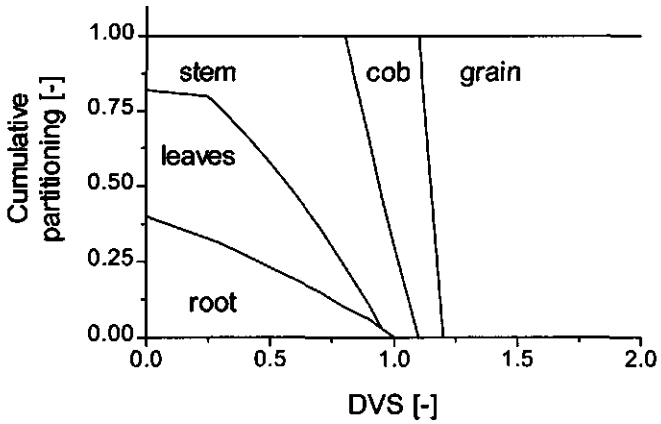

Figure 2.1 Cumulative dry matter partitioning coefficients as a function of development stage (DVS) for different organs (roots, leaves, stem, cob and grain) distinguished in SUCROS87 for maize.

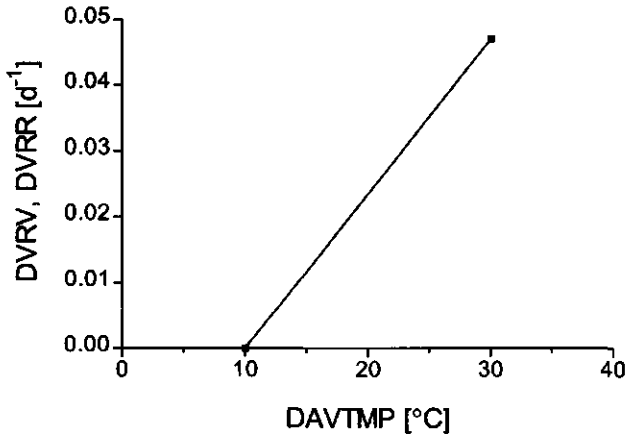

Figure 2.2 Development rate for vegetative and generative stage (DVRV and DVRR) for SUCROS87 as a function of daily average temperature DAVTMP.

The total dry matter growth F_GTW

The total dry matter growth F_GTW( ) is a complicated auxiliary variable, in which a number of auxiliary variables are introduced. In the version of SUCROS87 used the total dry matter growth is not allowed to become negative:

F_GTW( ) = MAX [0., G_GTW( )]

The total daily dry matter growth G_GTW is calculated as follows:

G_GTW( ) = [F_GPHOT( ) – F_MAINT( )] / F_ASRQ( )

where F_GPHOT( ) is the daily total gross assimilation. The maintenance respiration F_MAINT( ), the amount of assimilate required to maintain the living crop mass, is subtracted from the gross assimilation. Gross assimilation minus maintenance respiration is converted into the total dry matter growth, using the conversion function F_ASRQ( ), the partitioning weighed amount of carbohydrates (CH_2O) required to form an amount of structural dry matter.

The daily total gross assimilation F_GPHOT

The daily total gross assimilation in equivalents CH_2O (F_GPHOT) is calculated from the daily total gross assimilation in equivalents CO_2 (F_DTGA) using the conversion factor 30/44 kg CH_2O/kg CO_2:

F_GPHOT = 30/44 * F_DTGA( )

The auxiliary variable F_DTGA( ) is the basic element in the calculation of the crop organ growth rates. The daily total gross assimilation F_DTGA depends on the leaf area index (S_LAI), the global radiation (I_RDD), the average daytime temperature I_DDTMP, and the development stage (S_DVS). The total gross assimilation in SUCROS87 is determined by:

1) the amount photosynthetically active radiation (PAR) either direct or diffuse, which reaches the top of the canopy,
2) the amount of PAR (direct and diffuse) at a given depth in the canopy,
3) the gross assimilation rate for a leaf as a non-linear function of the level of PAR.

The maximum value of the gross leaf assimilation (P_AMX) is reduced as a function of development stage through the dimensionless tabulated function T_AMDVS(S_DVS). It is also reduced as a function of average daytime temperature modelled by the tabulated function T_AMTMP(I_DDTMP). The function T_AMDVS is presented in Figure 2.3, T_AMTMP is presented in Figure 2.4.

The results of these calculations are integrated numerically over leaf angle, depth in the canopy, and daytime. These calculations are executed in the subroutines TOTASS( ) and

ASSIM( ) and can be summarized as:

F_DTGA( ) = G_DTGA(TOTASS_(ASSIM_( )), T_AMTMP, T_AMDVS)

The parameters additionally used in this auxiliary function can be found in Appendix 3. Approximating functions are given by e.g. Spitters (1986).

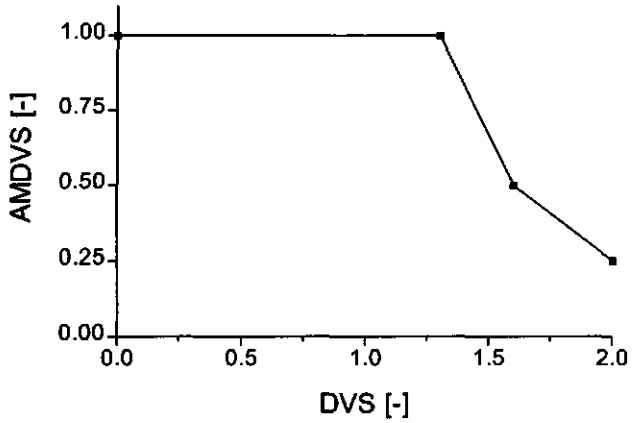

Figure 2.3 Value of leaf gross assimilation rate as a fraction AMDVS of its maximum value as a function of development stage DVS.

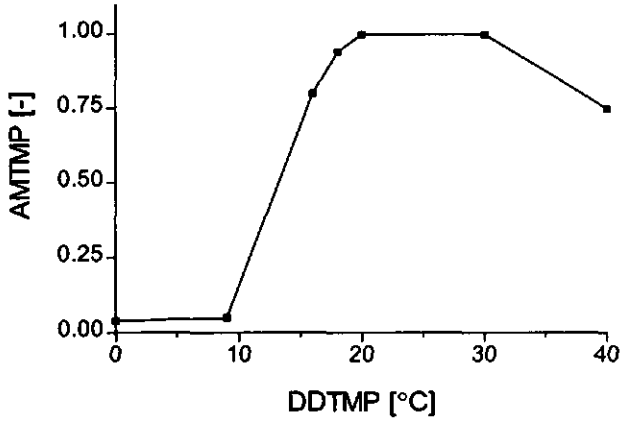

Figure 2.4 Value of leaf gross assimilation rate as a fraction AMTMP of its maximum value as a function of daytime average temperature DDTMP.

The maintenance respiration F_MAINT

The maintenance respiration, the amount of assimilates required for the upkeep of the living biomass, is a function of the living biomass F_MASS( ), temperature F_MTEFF( ) and development stage F_MNDVS( ):

```
F_MAINT  = F_MASS( ) * F_MTEFF( ) * F_MNDVS( )
```

Maintenance respiration depends on the composition of the living biomass:

```
F_MASS   = P_MAINLV * S_WLVG + P_MAINST * S_WST + P_MAINRT * S_WRT +
           P_MAINSO * S_WSO
```

its magnitude is exponentially related to temperature:

```
F_MTEFF  = P_Q10 ** [(I_DAVTMP - 25.) / 10.]
```

Once leaf senescence starts (S_WLVD > 0), maintenance respiration decreases proportional to the ratio of live leaf weight over total leaf weight (dead and living leaf weight):

```
F_MNDVS  = S_WLVG / (S_WLVG + S_WLVD)
```

The assimilate requirement F_ASRQ

The assimilate requirement F_ASRQ( ) depends on the chemical composition of the organs and their relative growth rate, and is calculated over the assimilates remaining after maintenance requirements have been met. The outcome of the function F_ASRQ is a value larger than 1, which represents the organ weighed assimilate requirement per unit dry matter growth. It is calculated as follows:

```
F_ASRQ   = T_FSH( ) * [ P_ASRLV * T_FLV( ) +
                        P_ASRST * T_FST( ) +
                        P_ASRCB * T_FCB( ) +
                        P_ASRSO * [1 - T_FLV( ) - T_FCB( ) - T_FST( )]] +
                        P_ASRRT * [1 - T_FSH( )]
```

The leaf relative death rate F_RDR

The relative death rate is the fraction of live leaf weight which dies per day. The process is calculated as the maximum of two auxiliary functions and a minimum value:

```
F_RDR    = MAX (F_RDRDV( ), F_RDRLT( ), 0.001)
```

The auxiliary function F_RDRDV( ) describes leaf weight senescence as a two step process controlled by the switch L_5( ):

F_RDRDV = [1 – L_5( )] * F_RDRL( ) + L_5( ) * F_RDRH( )

The switch L_5 is described in Section 2.2.5.

The auxiliary functions F_RDRL and F_RDRH describe leaf weight senescence as a function which increases with temperature if temperatures are above a specific threshold. The functions F_RDRL and F_RDRH model the effects of high temperatures on living leaf weight:

F_RDRL = P_SHRDRL * MAX [0, (I_DAVTMP – P_TBSRDR)]
F_RDRH = P_SHRDRH * MAX [0, (I_DAVTMP – P_TBSRDR)]

The second component of the function F_RDR, F_RDRLT, which describes leaf weight senescence due to low temperatures, is a fraction activated by the switch L_6( ) (described in Section 2.2.5):

F_RDRLT = L_6( ) * LIMIT [0., 1., G_RDRLT( )]

The auxiliary function G_RDRLT decreases linearly from 1 to 0 for temperatures between 0 °C and P_TBSKIL:

G_RDRLT = (P_TBSKIL – I_DAVTMP) / P_TBSKIL

Outside that temperature interval F_RDRLT is either 0 or 1.

The leaf area growth rate F_LAJUV and F_LAMAT

The calculation of the leaf area growth rate is switched between two auxiliary functions (F_LAJUV and F_LAMAT) by the switch L_7. The initially active auxiliary function describes leaf area index growth as an exponential function of a temperature sum:

F_LAJUV = G_LAJUV( ) * R_DTEFF( )

where

G_LAJUV = F_LAII( ) * P_RGRL * EXP (P_RGRL * S_TSUMEM)

and

F_LAII = P_NPL * P_LAINI * 10E – 4

where P_LAINI is the leaf area per plant (cm^2), and 10^{-4} a conversion to m^2. After the switch

has become active (cf. Section 2.2.5), leaf area index growth is proportional to the net green leaf weight growth using the tabulated function T_SLA:

F_LAMAT = T_SLA( ) * [R_GLV( ) – R_DLV( )]

This function is presented in Figure 2.5. In these calculations R_DTEFF, R_GLV and R_DLV are rates already introduced.

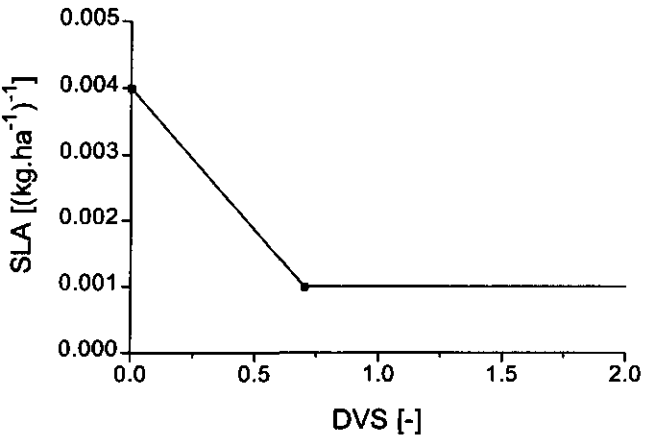

Figure 2.5 Specific leaf area (SLA) as a function of development stage DVS.

2.2.5 Description of switches

The switches are described giving their initial value and the condition(s) under which their value is changed. SUCROS87 contains 8 switches.

L_1 is the switch checking whether or not the crop has been sown:

L_1 = 0; IF (S_DAY.GE.P_DAYSOW) L_1 = 1

L_2, the switch checking emergence of the crop, is evaluated after integration:

L_2 = 0; IF (S_EMERG.GE.P_TSEMER) L_2 = 1

L_3 is the switch which flags whether the crop is mature or not. If the crop is mature the simulation is stopped:

L_3 = 1; IF (S_DVS.GE.2.) L_3 = 0

L_4 is the switch activating leaf weight senescence:

```
L_4 = 0;    IF (S_DVS.GE.P_DVSSEN) L_4 = 1
```

L_5 is the switch increasing senescence due to high temperatures:

```
L_5 = 0;    IF (S_DVS.GT.P_DVSRDR) L_5 = 1
```

L_6 is the switch activating senescence due to low temperatures:

```
L_6 = 0;    IF (S_DVS.GT.P_DVSKIL) L_6 = 1
```

L_7 is the switch activating leaf area growth proportional to leaf weight growth rate:

```
L_7 = 0;    IF (S_DVS.LT.P_DVSJUV.AND.S_LAI.LT.P_LAIJUV) L_7 = 1
```

L_8 is the switch activating the crop development rate in the reproductive phase:

```
L_8 = 0;    IF (S_DVS.GT.1) L_8 = 1
```

2.2.6 Initialization section

In the initialization section the initial values are assigned to the different state variables. In the case in which the simulation starts at the date of sowing, the initialization switch is evaluated before the rate calculations have started. The model is generally initialized using the initial LAI only; other initial values are zero, and are not presented here.

```
IF (L_2.EQ.1) THEN
   S_LAI = F_LAII
END IF
```

2.2.7 Operations on state variables

In SUCROS87 some additional calculations are executed on the basis of the calculated state variables. These variables are formally output variables.

```
O_WLV    = S_WLVG + S_WLVD

O_TADRW = S_WST + S_WLV + S_WSO + S_WCB

O_TDM    = S_WRT + S_TADRW
```

2.3 Description of LINTUL

The model LINTUL has been formulated to simulate the aboveground dry matter accumulation of vegetative organs and storage organs for maize on the basis of daily input values of temperature and radiation. The total dry matter increase is assumed to be proportional to the amount of light intercepted. The proportionality factor is controlled by temperature. The total dry matter increase is partitioned to the organs. Acronyms, their description, units and type are given in Appendix 2.

2.3.1 The inputs

As in SUCROS87 inputs are daily global radiation (I_RDD), daily maximum temperature (I_TMMX) and daily minimum temperature (I_TMMN). The temperatures are used to calculate the daily average temperature (I_DAVTMP).

Daily input:

I_TMMN, I_TMMX, I_RDD

Calculation:

I_DAVTMP = (TMMX – TMMN)/2

2.3.2 The state variables

The different organ dry matter weights are regarded as state variables. The following state variables are distinguished:

WSO = Grain dry matter weight
W = Aboveground dry matter weight other than grain

Other state variables are:

DAY = Day of the year
DVS = Crop development stage
EMERG = Temperature sum to emergence

The temperature sum to emergence (EMERG) and the crop development stage (DVS) determine the timing of events during crop growth. Temperature is the only input used in their calculation. The state variable DAY is used as a counter, associated with the control of in- and outputs. Given these introductory remarks the basic equations used in LINTUL are:

S_WSO = S_WSO + L_2( ) * L_3( ) * R_GSO( )

S_W = S_W + L_2( ) * L_3( ) * R_GW( )

```
S_DAY    = S_DAY + 1

S_DVS    = S_DVS + L_2( ) * L_3( ) * R_DVR( )

S_EMERG  = S_EMERG + L_1( ) * [1 - L_2( )] * R_DEMERG( )
```

where the switches L_1( ), L_2( ) and L_3( ) are described in Section 2.3.5.

2.3.3 The rate calculations

The rate variables R_X( ) used in the calculation of the state variables are themselves calculated using auxiliary functions, parameters, inputs and state variables.

```
R_DVR    = MAX [0, (I_DAVTMP - P_TBASE)]/P_TSDVR

R_DEMERG= MAX[0, (I_DAVTMP - P_TBSEM)]

R_GW     = [1 - T_FSO( )] * F_GTAW( )

R_GSO    = T_FSO( ) * F_GTAW( )
```

At this point the description is complete for the rate variables R_DVR and R_DEMERG. Other rate variables require auxiliary functions for a complete description.

2.3.4 Auxiliary functions in the rate calculations

The tabulated functions

The tabulated function used in the calculation of the daily organ growth T_FSO( ) is identical to the one used for SUCROS87. It was presented in Figure 2.1, and is calculated as the fraction of the growth rate remaining, after the biomass growth rate has been partitioned to the other crop organs.

The total dry matter growth F_GTAW

The total dry matter growth F_GTAW is the product of a parameter, the light use efficiency P_LUE, and two auxiliary functions, a reduction function F_FINT and a function to calculate the incoming photosynthetically active radiation F_PAR:

```
F_GTAW   = P_LUE * F_FINT( ) * F_PAR( )
```

The auxiliary function F_PAR

The photosynthetically active radiation F_PAR is linearly related to the incoming global radiation as:

F_PAR = 10E – 6 * P_FRPAR * I_RDD

where the factor 10^{-6} is a conversion factor, and P_FRPAR is the fraction photosynthetically active radiation in the global radiation.

The reduction function F_FINT

The reduction function F_FINT is calculated as the minimum value of two auxiliary functions:

F_FINT = MIN (F_FINTL( ), F_FINTS( ))

where F_FINTL is a logistic function of development stage calculated as:

F_FINTL = P_NPL * G_FINTL/(P_NPL * G_FINTL + 1 – P_NPL * P_FINTI)

with

G_FINTL = P_FINTI * EXP(P_RI * S_DVS)

F_FINTS describes a linear decrease in the interception of photosynthetically active radiation with development stage:

F_FINTS = [1 – L_4( )] * MIN(1, G_FINTS) + L_4( ) * MAX(0, G_FINTS)

where G_FINTS is calculated as:

G_FINTS = 0.5 – (S_DVS – P_DVHALF)/P_DVLGTH

and the switch L_4( ) is described in Section 2.3.5. The parameter P_DVHALF is the development stage at which the actual light use efficiency has decreased to half its maximal value, and P_DVLGTH is the period over which the light use efficiency decreases.

2.3.5 Description of switches

The switches are described giving their initial value and the condition(s) under which their value is changed. LINTUL contains 4 switches.

L_1 is the switch checking whether or not the crop has been sown:

```
L_1 = 0;    IF (DAY.GE.P_DAYSOW) L_1 = 1
```

L_2, the switch checking emergence of the crop, is evaluated at the end of the day:

```
L_2 = 0;    IF (S_EMERG.GE.P_TSEMER) L_2 = 1
```

L_3 is the switch which checks whether the crop is mature or not. If the crop is mature the simulation is stopped.

```
L_3 = 1;    IF (S_DVS.GE.2.) L_3 = 0
```

L_4 is a switch which controls the evaluation of the decrease in the light use efficiency:

```
L_4 = 0;    IF(S_DVS.GE.P_DVHALF) L_4 = 1
```

2.3.6 Initialization section

In LINTUL, the initial growth rate is calculated using the initial value of the development stage DVS (standard zero). The value of the reduction function for the light use efficiency for DVS = 0 equals P_FINTI.

2.4 Model modifications in the framework of this thesis

The model modifications were kept to an absolute minimum. Modifications focused on the scalar values and on the parameterization of the tabulated functions used in both models.

In both models the numerical values of scalars used as defaults in the models were replaced by acronyms (P_X). The parameter value actually used in a simulation run is calculated as P_X = P_X_rel×P_X_def, where P_X_def is the default value of the parameter. The value used in the actual run is calculated when P_X_rel is read from the input file. Model runs with all parameters in the input file equal 1 yield the nominal simulation result. In SUCROS87, the original scalar values were declared as parameter values. In LINTUL default parameterization was based on the literature review of parameter values in Chapter 4.

In order not to change the structure of the model, but to allow relatively simple modifications to the tables using a limited number of parameters, transformations of both argument and response of the tabulated function were chosen, which are appropriate given the scale on which these are measured. The single exception to this rule is the parameterization of the tables T_DVRV and T_DVRR. To reduce the number of parameters required, the tabulated linear function was replaced by a linear function within the source code. A detailed discussion and description of the transformations chosen and fitted (in Chapter 4) are presented in Appendix 7.

2.5 Conclusion

Apart from formal subroutines, the models SUCROS87, and to a lesser extent LINTUL, contain submodels, each of which describe a certain phase of growth. The changes between growth phases are modelled using logical switches, which are triggered if the state variable or state variables used in a switch meet the required condition. These phases are distinct: from one day to the next a phase in growth will end and a next phase start. This type of model is a switching model, or a composite model.

There are e.g. r growth phases in a model; the actual phase of growth is e.g. γ (γ = 1...r). In these phases different equations are used to calculate the growth rates of state variables. If we call the set of all growth rate functions F, the set of functions used for the state variable x_i is the set F_i (i = 1...n). In growth phase γ the growth rate function $f_{i\gamma}$ is used. The arguments of the function $f_{i\gamma}$ are state variables, parameters and inputs. The specific arguments depend on the growth phase. If X is the set of all state variables, U is the set of all inputs and Θ the set of all parameters used in the model, then $X_{i\gamma}$, $U_{i\gamma}$ and $\Theta_{i\gamma}$ are the arguments of the growth rate function $f_{i\gamma}$ for the state variable x_i in growth phase γ. Each single state variable in SUCROS87 and LINTUL is calculated using a difference equation, written as:

$$x_i(k+\Delta t) - x_i(k) = f_{i\gamma}\left(X_{i\gamma}(k+\Delta t), U_{i\gamma}(k+\Delta t), \Theta_{i\gamma}\right)\Delta t$$

where

$x_i(k)$	simulated state variable x_i at time k,
$f_{i\gamma}$	an element of the set of functions F_i: $f_{i1}...f_{i\gamma}...f_{ir}$, determining the rate of change of state variable x_i within growth phase γ,
$X_{i\gamma}(k)$	the subset of states influencing the growth of state x_i at day k,
$U_{i\gamma}(k)$	the subset of input variables influencing the growth of state x_i at day k,
$\Theta_{i\gamma}(k)$	the subset of parameter values influencing the growth of state x_i at time k,
Δt	the time step (Δt = 1 day in SUCROS87 and LINTUL),
γ	index determined by the growth phases of the crop (as modelled).

Clear-cut phases within both SUCROS87 and LINTUL are the submodel for emergence, and the model for growth. For SUCROS87, the growth phase itself is divided in the following phases: a juvenile leaf growth stage, a senescing stage with a number of subphases, and a mature growing period between the end of juvenile growth and the start of senescence. These phases are not present in LINTUL.

Given the number of steps required to completely describe SUCROS87, and the number of composite parts in comparison to those in LINTUL, SUCROS87 is certainly the most complex model of the two. Whether this has consequences for the following analyses remains to be seen.

Chapter 3

Analysis of model behaviour

3.1 Introduction

Assigning values to the parameters is an important action to prepare models for application (cf. Chapter 1). Once parameter values are assigned, qualitative analyses of model behaviour allow a comparison between expected and realized behaviour, which may lead to model reformulation.

In this chapter, the analysis will be concentrated on those aspects of model behaviour which are important when calibrating, i.e. finding parameters such that model output matches observed outputs as closely as possible according to some criterion value. Calibration itself will be treated in Chapter 6. The analysis of model behaviour also allows to test whether the model behaves as expected.

A simulation model can be characterized in terms of its structure. Model structure was described in the previous chapter. Model structure may have consequences for calibration, which will be discussed in Section 3.2.

Another aspect of the models with important consequences for parameter estimation is the nature of the relation between calibration variables and parameters, e.g. linear, or non-linear. The relation between calibration variables and individual parameters is analysed in Section 3.3.

3.2 Analysing the model structure for calibration

When considering model calibration it is useful to know whether the calibration can be divided into a number of smaller calibration problems, which may be solved sequentially. This is possible if the model contains independent submodels. Once the submodels are identified one calibrates parameters in the submodels separately and sequentially using the submodel output and the corresponding measurements. Each smaller calibration problem still poses its own problems, the most important that of parameter choice. The procedure is the following: In a first step parameters used in the first submodel are calibrated using observations corresponding to the output of the first submodel. In the second step, parameters in a second submodel are calibrated. In this step the observations already used are ignored, and the parameters calibrated in the first submodel are kept fixed. For the second calibration step not to influence the fit achieved in the first calibration step, the model structure has to meet specific conditions, and the parameters selected should not be common to both submodels.

Obviously, the question to be answered is: does the model structure allow this type of calibration? Analyses of model structure as presented by Beumer et al. (1978) of Forrester's model of urban dynamics and Klaassen et al. (1980) of Forrester's model of world dynamics allow to answer this question. To analyse whether one model variable influences another one, the model is represented using indicator variables, which indicate use (value = 1; influence present) or non-use (value = 0; influence absent) of a model variable in the calculation of other model variables. We can illustrate the principle using Appendix 3 for SUCROS87. Taking the state variables DVS and WRT as an example, we observe (as presented in Appendix 3, Table A3.1) that the following state variables (values from integration step k–1) are used in the calculation of the state variable WRT at integration step k: EMERG, DVS, LAI, WLVG, WLVD, WRT, WST, WSO, and WCB. The state variable DVS(k) only depends on DVS(k–1) and EMERG(k–1). The arguments

used in the calculation of each state variable can be represented in the matrix presented below, where the column variables refer to the arguments at integration step k–1 and the row variables is the response at integration step k:

	EMERG	DVS	TSUMEM	LAI	WLVG	WLVD	WRT	WST	WSO	WCB
EMERG										
DVS	1	1	0	0	0	0	0	0	0	0
LAI										
TSUMEM										
WLVG										
WLVD										
WRT	1	1	0	1	1	1	1	1	1	1
WST										
WSO										
WCB										

For an explanation of the abbreviations cf. Appendix 1.

The remaining rows can be filled in the same manner. The resulting matrix now summarizes the coupling between different model variables. The matrix can be used to represent and analyse consequences of coupling between variables in the model, using a linear description, in vector representation: $x(k) = P\,x(k–1)$, where P is the above matrix filled with indicator values 0 and 1.

Although the above example refers to state variables, this method of describing couplings can be extended to accommodate other (non-state) variables, e.g. those considered for calibration. In any analysis of this type, however, all state variables should be incorporated, as they determine the carry-over of information from one integration step to the next. In the following we will use the term 'model variables' to indicate the extension of the analysis to include non-state variables.

If the model equations do not change over the integration period, a single matrix P suffices to describe the coupling between model variables. The matrix describing the coupling between the relevant variables between integration steps $k = 0$ and $k = n$ would be calculated as the matrix P to the power n: P^n, where elements with values larger than 1 are interpreted as 1. The elements $(P^n)_{ij}$ (i = row; j = column) in the matrix indicate which variables are used in the calculation of a specific variable over n integration steps. If a matrix element $(P^n)_{ij}$ is 0, the conclusion is that over n integration steps considered the model variable does not influence variable x_i. In that case x_i is independent of x_j over n integration steps. It may be shown that calculation of matrix P to the power b–1, where b is the number of variables used in the evaluation, shows all dependencies and independencies which will ever be realized in the model calculations.

This structural analysis can result in different types of P matrices. Names of matrix types are based on Press et al. (1992, page 64). We will discuss some examples of matrices to illustrate the possible conclusions this analysis allows.

If the analysis results in a diagonal matrix, the conclusion would be that all model variables are calculated independent from each other. In that case, provided parameters are available that allow to modify the individual model variables, parameter values can be found by sequentially

matching individual model variables to their corresponding measurements. In this case, a diagonal structural matrix, calibration of the selected parameters can be executed in any order, and the absence of relations between the model variables ensures that the match between measurements and model variables is not influenced if the next model variable is calibrated.

If the analysis results in a matrix P^{b-1} where all elements $(P^{b-1})_{ij}$ are non-zero (all indicator variables 1) no model variable can be calculated independent from the other. Obviously, for a model with such a structure it is not possible to split the calibration problem into a set of smaller problems.

Results of a structural analysis may also be intermediate between these extremes. For a hypothetical model after 4 (= b –1) evaluations, the analysis of couplings between state variables x_1,...x_5 may e.g. result in:

$$P^4 = \begin{array}{c|ccccc} & x_1 & x_2 & x_3 & x_4 & x_5 \\ x_1 & 1 & 1 & 0 & 0 & 0 \\ x_2 & 0 & 1 & 0 & 0 & 0 \\ x_3 & 1 & 1 & 1 & 1 & 1 \\ x_4 & 1 & 1 & 0 & 1 & 1 \\ x_5 & 1 & 1 & 0 & 0 & 1 \end{array} \qquad (3.1)$$

where non-zero values $(P_{ij})^4$ have been given an indicator value 1 (influence present). In this case, sorting rows, using an increasing number of non-zero values in the rows as the sorting criterion, and subsequently permutating the columns to ensure that the diagonal elements refer to the same (in this example state) variable, yields a triangular matrix:

$$\begin{array}{c|ccccc} & x_2 & x_1 & x_5 & x_4 & x_3 \\ x_2 & 1 & 0 & 0 & 0 & 0 \\ x_1 & 1 & 1 & 0 & 0 & 0 \\ x_5 & 1 & 1 & 1 & 0 & 0 \\ x_4 & 1 & 1 & 1 & 1 & 0 \\ x_3 & 1 & 1 & 1 & 1 & 1 \end{array} \qquad (3.2)$$

This matrix suggests that parameters contained in individual state variables may be calibrated in a sequence determined by their occurrence in the following sequence of state variables x_2, x_1, x_5, x_4, x_3. If a parameter contained in x_2 has been calibrated, the fit of x_2 is not modified in the next step, the calibration of a parameter contained in x_1. This calibration procedure can be continued sequentially until all state variables are calibrated.

A slight modification of the model structure, if e.g. variable $x_4(k–1)$ is used in the calculation

of $x_5(k)$, introduces a new type of relation in the model. This modification yields a block-triangular matrix:

	x_2	x_1	x_5	x_4	x_3
x_2	1	0	0	0	0
x_1	1	1	0	0	0
x_5	1	1	1	1	0
x_4	1	1	1	1	0
x_3	1	1	1	1	1

(3.3)

This matrix suggests that in this example parameters contained in the individual state variables could be calibrated in a sequence determined by their occurrence in the sequence x_2, x_1, (x_5, x_4), x_3, where parameters contained in the group (x_5, x_4) can only be calibrated simultaneously, as one state variable cannot be changed without changing the fit of the other. These examples refer to couplings between state variables, but could be extended to incorporate output variables, or rate variables. In both cases the value on the corresponding diagonal element would be zero.

The question whether a given coupling matrix can be made triangular does not have a straightforward answer. Beumer et al. (1978) refer to algorithms for row and column permutation to analyse this problem. It is obvious that a $b \times b$ matrix can not be made triangular, if more than $\frac{1}{2}b(b+1)$ elements are non-zero. A coupling matrix can furthermore not be made triangular, if the model contains loops, i.e. if the calculation of state variable 1 depends on state variable 2, which in turn depends on state variable 1 (Beumer et al., 1978).

A structural analysis of SUCROS87 and LINTUL in terms of a coupling matrix P is not straightforward. As shown in Chapter 2, SUCROS87, and to a lesser extent LINTUL, contain submodels, each of which describe a phase of growth. The model equations change between phases. Different matrices P would have to be used to describe the couplings in the model for each phase. The resulting analysis is complicated, even more so as the changes in the model equations depend on state variables, in which parameters which might be considered for calibration are used. This complication can be illustrated using the above example. According to the source code of SUCROS87 grain weight (WSO) is an argument in the calculation of the root weight (WRT). However, using the default parameter values, WSO has no effect on root weight, because the partitioning functions define root growth to stop before grain growth starts.

3.2.1 Structure analysis using a simulation study

Given that the relations between model variables in SUCROS87 and LINTUL depend on the parameter values actually used, e.g. on those for partitioning (cf. previous section), the value of a paper and pencil structural analysis of both models is limited. This leads to the execution of the analysis as a simulation study aiming to clarify couplings between calibration variables, i.e.

model variables for which measured data are available. The results of the analysis may be used to tentatively structure the calibration procedure.

The analysis requires that we derive a coupling matrix similar to P^n from a simulation study. To do so we execute the following steps:

- Introduce new parameters θ_j (j = 1...b), where b is the number of calibration variables. For all calibration variables c_j, add the line $c_j = \theta_j c_j$ at one single relevant location in the source code. For state variables which are to be calibrated, the parameters are introduced in the rate equation. In those cases where measurements of auxiliary variables are available for calibration, the new parameters are introduced as soon as the auxiliary variable has been calculated.
- The analysis itself is executed changing the parameters θ_j one by one. If θ_j is increased or reduced 100%, the change in the calibration variable corresponds to an on-off analysis. The changes result in $2b$+1 runs, the factor 2 for two values of θ_j; b for the number of calibration variables, and one run as reference (all θ_j equal 1).
- The changes in a calibration variable c_i given the on-off perturbation in variable c_j are then analysed at time t e.g. as the average absolute change in the variable c_i relative to value of c_i (which should not become 0) in the reference run, given the changes in the variable c_j:

$$a_{ij}(t) = \frac{|c_i(t,-1) - c_i(t,0)| + |(c_i(t,+1) - c_i(t,0)|}{2\,c_i(t,0)} \tag{3.4}$$

where the second argument of c_i refers to the following perturbations:

-1 : $\theta_j = 0$,
0 : $\theta_j = 1$,
$+1$: $\theta_j = 2$.

The results can be represented as a matrix A, a square matrix containing elements a_{ij}, which describe the effect of a change in parameter θ_j on calibration variable c_i at time t, at which the output is evaluated.

- The matrix A (containing elements a_{ij}) is transformed to a matrix Q containing indicator variables θ_{ij}. This transformation is executed using a function defined as (if $a_{ij} > \varepsilon$, then $\theta_{ij} = 1$ else $\theta_{ij} = 0$). The choice of the threshold parameter ε determines the interpretation of the matrix Q. A value θ_{ij} of 0 where $\varepsilon = 0$ ($j \neq i$) signifies that no effect of c_j through θ_j on c_i has been observed. Within the restrictions of this analysis, the variables are perfectly independent until the time at which the output is evaluated. When sequentially calibrating the variables, perfect independence suggests that the calibration result for one variable may not be modified by the calibration of another variable. Note that in the previous sentence the notion 'calibration of a variable', is used as shorthand for 'finding values for parameters used in the calculation of the model variable and not used in the calculation of other model variables in such a way that the model variable matches the corresponding measurements as closely as possible'. This shorthand notation will be used in this and the following sections.

Instead of analysing the results in terms of perfect independence ($\varepsilon = 0$), one may accept a weak dependence of calibration variable c_i on calibration variable c_j ($j \neq i$, $a_{ij} \approx 0$). A procedure based on

a value of $\varepsilon \neq 0$ can be interpreted as accepting imperfect independence between c_i and c_j. When sequentially calibrating the variables, this could imply that calibrating the next variable has a small effect on the calibration quality of the previous variable. This effect may be acceptable in those instances where the advantage of dividing the problem into a sequence of smaller problems outweighs the loss of calibration quality and possibly a loss of predictive quality.

A sequential approach to model adaptation (either parameter estimation or the introduction of new model elements) of a simulation model for water limited crop growth is suggested by Goudriaan (1996a) for the components water balance and crop growth. The above procedure may offer an option to formalize this argumentation.

3.2.2 Simulation experiments

Two examples of structural analysis are presented. In the first example the coupling between all state variables was evaluated at final harvest. Results for both LINTUL and SUCROS87 are presented.

As the structure of SUCROS87 changes in time, the second example shows the relation between state variables at different harvesting days for SUCROS87. For the second example a historic measurement schedule is used.

The simulation experiments were executed using the default parameter values for maize (cf. Appendices 5 and 6); the experimental characteristics (management characteristics and experimental design) of a maize experiment at Sinderhoeve in 1985 (Ouwerkerk and Drenth, 1986) were used; weather data (1985) were from the meteorological station at Wageningen (Dept. Physics and Meteorology, 1954-1992). The analysis was executed as described in Section 3.2.1.

Requiring perfect independence: results

The relations between state variables at final harvest were investigated for SUCROS87 and LINTUL. In both analyses a threshold ε of exactly zero is used. The results for the two examples are presented in Tables 3.1 to 3.3. Columns and rows are permuted to show the most independent state variables in prominent position, while ensuring that row i and column i refer to the same state.

Table 3.1 shows the results for SUCROS87 at final harvest. The block triangular matrix indicates that a sequential calibration is possible for a few state variables. The sequence in which variables may be calibrated is the following: emergence (EMERG); development state (DVS), and the temperature sum (TSUMEM), which is associated with the calculation of the juvenile leaf area index (LAI). All other state variables are dependent on a larger number of state variables, culminating in the cob weight and grain weight (WSO and WCB), which are influenced by all state variables. It is interesting to note that part of this sequence of calibration variables corresponds to that used in calibration of the model WOFOST which is similar to SUCROS87: Boons-Prins et al. (1993) and Rötter (1993) sequentially calibrate development stage and dry matter partitioning.

Table 3.2, which presents the results for LINTUL at final harvest shows that the different

Table 3.1 The effect of a one at a time on-off perturbation of the rate variables on the state variables in SUCROS87 at final harvest. An indicator value 1 denotes that row-variable i is influenced by column-variable j during the simulation period. For an explanation of the abbreviations cf. Appendix 1.

	EMERG	DVS	TSUMEM	WRT	WLVG	LAI	WST	WLVD	WCB	WSO
EMERG	1	0	0	0	0	0	0	0	0	0
DVS	1	1	0	0	0	0	0	0	0	0
TSUMEM	1	1	1	0	0	0	0	0	0	0
WRT	1	1	1	1	1	1	1	0	1	0
WLVG	1	1	1	1	1	1	1	1	1	0
LAI	1	1	1	1	1	1	1	1	0	0
WST	1	1	1	1	1	1	1	1	1	0
WLVD	1	1	1	1	1	1	1	1	1	0
WCB	1	1	1	1	1	1	1	1	1	1
WSO	1	1	1	1	1	1	1	1	1	1

Table 3.2 The effect of a one at time on-off perturbation of the rate variables on state variables at final harvest for LINTUL. An indicator value 1 denotes that row-variable i is influenced by column-variable j during the simulation period. For an explanation of the abbreviations cf. Appendix 2.

	EMERG	DVS	W	WSO
EMERG	1	0	0	0
DVS	1	1	0	0
W	1	1	1	0
WSO	1	1	0	1

state variables can be arranged in a triangular matrix. The state variables in LINTUL might then be calibrated sequentially.

The relations of the state variables in SUCROS87 at different harvesting times were also determined following the above procedure. The simulation was executed using the design of an existing experiment in which the state variables EMERG, DVS, LAI and TADRW (total aboveground dry weight) were measured. Table 3.3 shows that the relevant structural matrix is triangular at time = 156. At day 170, a feedback relation between leaf area index (LAI) and total aboveground dry matter (TADRW) is active; the resulting matrix is block-triangular. After day 170 the relations between state variables do not change any more. This reflects two different phases in SUCROS87: juvenile leaf area development, during which leaf area development is independent of leaf weight, and mature leaf area development, during which leaf area index is calculated from the growth rate of leaf dry matter. In the juvenile phase leaf area can be calibrated independent of total aboveground dry matter; in the mature phase this is no longer possible. This supposes of course that the parameters determining the shift from juvenile to mature leaf area growth are not themselves calibrated.

Table 3.3 The effect of a one at a time on-off perturbation of the rate variables on state variables in SUCROS87 in the course of a growing season. An indicator value 1 denotes that row-variable i is influenced by column-variable j during the simulation period. For an explanation of the abbreviations cf. Appendix 1.

	EMERG	DVS	LAI	TADRW	
EMERG	1	0	0	0	at day 156
DVS	1	1	0	0	
LAI	1	0	1	0	
TADRW	1	1	1	1	
EMERG	1	0	0	0	at day 170 and after
DVS	1	1	0	0	
LAI	1	1	1	1	
TADRW	1	1	1	1	

Requiring imperfect independence: results

To analyse the structure of the model under the assumption of imperfect independence ($\varepsilon \neq 0$), we have reanalysed the first example for SUCROS87 (cf. Table 3.1). The results of this re-analysis are presented in Table 3.4. If we assume that for our purpose coefficients a_{ij} less than 10% are negligible ($\epsilon = 0.1$), transformation and re-arrangement of the matrix Q yields the matrix presented in Table 3.5. The block-triangular matrix in Table 3.5 suggests that within the homogeneous block of crop organ variables as found in Table 3.1 the state variables leaf area (LAI), green leaf weight (WLVG) and dead leaf weight (WLVD) are the ones most strongly interacting. If one considers sequential calibration (notwithstanding losses of calibration quality due to interaction) calibration of emergence (EMERG), development stage (DVS), juvenile leaf area (LAJUV), followed by a simultaneous calibration of the other leaf associated state variables and in a next step cob weight (WCB), root weight (WRT), and grain weight (WSO) is an option.

3.3 Relations between parameters and calibration variables: single parameter - single output

3.3.1 Introduction

The type of relation between parameters and model response has important methodological consequences. The relation between parameters and a model response is e.g. either characterized as linear or as non-linear, resulting in the application of either linear or non-linear regression methodology (Seber and Wild, 1989; Montgomery and Peck, 1992; Payne and Lane, 1993).

In this section a specific aspect of relations between individual parameters and model responses for SUCROS87 and LINTUL will be analysed. The aspect selected is continuity. The selection is based on the importance of this criterion for parameter estimation algorithms (e.g. Gill et al., 1981).

Table 3.4 The influence of a one at a time on-off perturbation of the rate variable on the state variables in SUCROS87 at final harvest: the percentage change induced by a 100% change in the calculation of the rate (rounded). A dot indicates a value rounded to 0. A value larger than 0 denotes that row-variable i is influenced by column-variable j during the simulation period. For an explanation of the abbreviations cf. Appendix 1.

	EMERG	DVS	TSUMEM	WRT	WLVG	LAI	WST	WLVD	WCB	WSO
EMERG	48	0	0	0	0	0	0	0	0	0
DVS	51	57	0	0	0	0	0	0	0	0
TSUMEM	51	21	100	0	0	0	0	0	0	0
WRT	51	65	53	97	8	65	4	0	.	0
WLVG	48	268	54	3	93	65	4	100	.	0
LAI	50	355	82	2	44	141	3	19	0	0
WST	51	15	25	4	10	54	93	.	.	0
WLVD	53	5	54	3	93	65	4	71	.	0
WCB	50	10	6	4	13	47	13	1	96	.
WSO	53	11	13	5	6	50	15	22	11	94

Table 3.5 The effect of a one at a time on-off perturbation of the rate variable on the state variables in SUCROS87 at final harvest: the transformation and permutation of the results from Table 3.4 assuming a threshold value ε of 10%. A value larger than 0 denotes that row-variable i is influenced by column-variable j during the simulation period. Rows and columns are permuted in comparison to Table 3.4. For an explanation of the abbreviations cf. Appendix 1.

	EMERG	DVS	TSUMEM	LAI	WLVG	WLVD	WST	WCB	WRT	WSO
EMERG	1	0	0	0	0	0	0	0	0	0
DVS	1	1	0	0	0	0	0	0	0	0
TSUMEM	1	1	1	0	0	0	0	0	0	0
LAI	1	1	1	1	1	1	0	0	0	0
WLVG	1	1	1	1	1	1	0	0	0	0
WLVD	1	0	1	1	1	1	0	0	0	0
WST	1	1	1	1	1	0	1	0	0	0
WCB	1	1	0	1	1	0	1	1	0	0
WRT	1	1	1	1	0	0	0	0	1	0
WSO	1	1	1	1	0	1	1	1	0	1

The relations between responses at system level and individual parameters are rarely presented in literature. Nonhebel (1994) and Kocabas et al. (1993) present studies of the effect of systematic variations in temperature on model responses; both conclude (for different models) that the model predictions are occasionally very sensitive to small changes or inaccuracies in temperature. This suggests that small changes in parameters influencing the response of the model to

temperature may also cause relatively large changes in model predictions.

It is difficult, due to the numerous relations between calibration variables, to discuss and analyse continuity of the relation between parameter and calibration variables for SUCROS87 and LINTUL case by case and analytically on the basis of the description in Chapter 2. Therefore, an analysis of continuity is best executed on the basis of a simulation study, using reference parameter values.

3.3.2 The continuity of functions

The question to be answered in this section is whether the models contain structures which may induce discontinuities in the relation between a state variable and a parameter at a given time. Continuity of a relation excludes the possibility of relatively large state variable changes induced by infinitely small changes in a parameter. The logical functions (switches) defined in the description of the models (Chapter 2) are natural candidates for model elements which may introduce relatively large changes in state variables for infinitely small changes of a parameter.

A switch is a device that converts a smooth input into a discontinuous output (Thornley and Johnson, 1990). The following model elements are typical switches:

1) initialisation procedures for state variables, in which the switch is incorporated in the state equation, which is literally an on-off switch, and therefore less interesting than
2) conditional constructions incorporated in the rate equation, e.g. if-then-else constructions referred to as logical switches in the model description (cf. Chapter 2).

An example in which a switch is used is the calculation of the growth rate. Growth is conditional on emergence in LINTUL and SUCROS87. The state variable emergence (after n integration steps) is calculated as:

$$\mathrm{EMERG}(n) = \sum_{i=1}^{n} \max\left(0;\ \mathrm{TAVG}(i) - \mathrm{TBASE}\right) \tag{3.5}$$

and is initialised at sowing. It is input to a logical switch (L), which determines start of crop growth:

$$L = \begin{cases} 1 & \textit{if} \quad \mathrm{EMERG} \geq \mathrm{TSEMER} \\ 0 & \textit{if} \quad \mathrm{EMERG} \leq \mathrm{TSEMER} \end{cases} \tag{3.6}$$

To illustrate the effect of this switch on crop growth, assume that the crop growth rate is constant, the final harvest date is fixed, and that the average temperature TAVG is constant at a value above the base temperature for emergence TBASE. As one day is the smallest time step in both models, the crop emerges after a number of days equal to TSEMER/(TAVG–TBASE), rounded upwards to the nearest integer.

Changing either TSEMER or TBASE gradually, while everything else remains constant, the value of TSEMER/(TAVG–TBASE) will gradually change too. As soon as TSEMER/(TAVG–TBASE) passes an integer value, the number of days to emergence changes by one. At the jump,

yield changes by the daily growth rate (assumed constant in this example). Plotting final yield against TBASE or TSEMER will yield a step function: constant over a small range of the parameter, suddenly changing, and then remaining constant again. For a varying TAVG, this function will be less regular, and may even show changes larger than a daily growth rate. This occurs e.g. if a specific value of TBASE allows a crop to emerge just before a cold spell during which TAVG is smaller than TBASE. Higher values of TBASE will cause the crop to emerge after the cold spell, whereas lower values will allow the crop to emerge before the cold spell. As the models, notably SUCROS87, contain more switches, as these different switches may also interact, and as the daily growth rate is not constant, it is most practical to analyse possible discontinuities in relations between model outputs and parameters using a numerical approach.

3.3.3 A numerical procedure to detect discontinuity of functions

A numerical procedure to test a function for discontinuous behaviour may be executed in a straightforward way: plot a state variable, simulated for different values of a parameter, against the parameter and inspect these plots visually. If these plots show a break on successively more precise scales, it may be strongly suspected that the function is not continuous. Pinpointing a single discontinuous event is sufficient to prove discontinuity. On the other hand it may be very hard to find a single discontinuity. An example in which a single function value of an otherwise continuous function evaluated in an interval is shifted down- or upwards may readily illustrate the difficulty of finding this discontinuity.

If a computer algorithm is used to test a function $f(x)$ for continuity at any value x_0 where a discontinuity is suspected, a formal evaluation criterion has to replace visual inspection. A criterion can be based on the evaluation of the function at four equidistant values near x_0:

$$\begin{aligned} x_1 &= x_0 - 1\tfrac{1}{2}h, \quad f_1 = f(x_1); \quad x_2 = x_0 - \tfrac{1}{2}h, \quad f_2 = f(x_2) \\ x_3 &= x_0 + \tfrac{1}{2}h, \quad f_3 = f(x_3); \quad x_4 = x_0 + 1\tfrac{1}{2}h, \quad f_4 = f(x_4) \end{aligned} \tag{3.7}$$

The differences between the four function values are designated as d_l, d_c, d_u, where d_l equals f_1-f_2, d_c equals f_2-f_3 and d_u equals f_3-f_4. All three differences are evaluated over the distance h. The criterion is based on these function evaluations, as follows:

$$C = \frac{|d_l + d_u - 2d_c|}{|d_u - d_l| + \varepsilon} \tag{3.8}$$

where ε is a small positive value introduced to prevent division by zero.

The rationale behind criterion C is the following. If the function f is smooth on the interval from x_1 to x_4, C will tend to 0 if step size h tends to 0. If, on the other hand, f has a jump discontinuity between x_2 and x_3, C will become large, if h tends to 0. Thus, a large value of C is an indication of a possible discontinuity, whereas a small value is reassuring. In the present study a value of $C \geq 1$ will lead to a closer scrutiny. The value $C = 1$ is required to distinguish between large and small values, and allows to reduce the number of simulations.

One may sketch the behaviour of C for different limit cases. Assume that f is smooth on the interval studied. Denote the first, second and third derivatives at x_0 by $f^{(1)}$, $f^{(2)}$ and $f^{(3)}$. It may be shown by Taylor expansion of f around x_0, that $d_u - d_l$ is approximately equal to $2h^2 f^{(2)}$. Similarly, $d_u - 2d_c + d_l$ is approximately equal to $h^3 f^{(3)}$. Approximation errors are in the order of $h^3(f^{(2)})$ resp. $h^4(f^{(3)})$. for a smooth function $C \approx h^3 |f^{(3)}|/(2h^2 |f^{(2)}| + \varepsilon)$, which tends to 0 if h tends to 0, irrespective of the small positive value of ε. Now consider the case that f makes a jump of size δ between x_2 and x_3, staying smooth elsewhere. In that case, C will tend to the limit $2|\delta|/\varepsilon$, which is large provided that ε is small compared with the jump δ.

Using this criterion as a diagnostic for possible discontinuities, a procedure to analyse the relations between calibration variables and parameters in terms of discontinuity could be formulated as follows:

Divide the range of the parameter in a number of equal intervals (say between 10 and 100). Execute model runs for each value of the parameter. Evaluate the changes in function value by comparison with the neighbouring points using the above criterion. If the threshold value of the criterion is exceeded, select the parameter interval (x_1, x_4) yielding the largest criterion value and repeat the procedure using this interval as the starting point for a new evaluation. Repeat the evaluation(s) a number of times (thus evaluating the function for values of h tending to 0), but stop if d_c becomes too small to be relevant. Plot the results.

3.3.4 The effect of the parameters in SUCROS87 on final yield: discontinuities

An analysis of continuity along the lines described above was executed for SUCROS87, using weather for Wageningen (1985, Dept. Physics and Meteorology, 1954-1992) and default values for the parameters (Appendix 5). The output variable evaluated was the total aboveground dry matter (TADRW) at final harvest.

Initially 50 runs were made for all parameters in SUCROS87, which were varied plus or minus 50% in steps of two percent. The parameters to which TADRW showed no response at all (LAIJUV, TBSKIL and DVSKIL) were varied 98% in a second attempt and the results were analysed again. The criterion values C were calculated for each of the parameters over the whole range.

The maximum criterion values of this first run are given in the first column of Table 3.6. Those parameters with a criterion value larger than or equal to 1 were chosen for closer scrutiny. The interval enclosing the largest criterion value was selected. Within this smaller range an additional 25 simulation runs were made. This procedure was repeated three times. The criterion value for the second, third and fourth set of runs are given in Table 3.6. Parameters yielding criterion values smaller than 1 are dropped from further analysis.

The final results show that the parameters associated with large criterion values are either used in switches (TSEMER, LAIJUV) or are used in state variables that are input to switches (RGRL and TBSJUV in LAI, TSDVRR and TBSDVRR in DVS). An exception is the parameter ATRL, which is used in the subroutine ASTRO. The variation of fifty percent exceeds a limit for this parameter, causing the subsequent analyses to focus on a discontinuity outside the possible range.

Table 3.6 An analysis of the occurrence of discontinuous relations between parameters and state variables: the example of the relation between TADRW at final harvest and parameters used in SUCROS87. The maximum criterion values (C, Eqn. 3.8) are presented for 4 iterations in which the interval size was decreased from one iteration to the next. Model parameters that are not included in this table or that are dropped from one iteration to the next had a maximum criterion value less than 1. For an explanation of the abbreviations cf. Appendix 1.

Parameter	Iteration 1	2	3	4
TSEMER	92156	92156	92156	92155
TBSEM	92156	92156	92156	92155
DVSJUV	7280	7280	7280	7280
DVSSEN	2533	1329	1329	1329
TSDVRV	236	286	1144	1306
TSDVRR	141	1514	2120	2650
TBSVRV	68.0	179	1690	321
RGRL	60.5	1778	2488	3110
TBSVRR	35.1	1323	2650	2650
ADVRV	27.6	408	1269	2405
ADVRR	26.1	527	1403	2405
DVSRDR	25.8	2951	2951	2951
TBSJUV	24.4	1538	2461	3076
BDDTMP	20.3	11.5	1.00	0.500
LAIJUV	19.7	9531	9531	9531
FRDFBA	18.5	6.00	0.600	
TBSKIL	14.8	1.34	0.900	
TBSRDR	4.89	1.00	0.500	
FRDFBB	3.60	1.33	0.600	
ATRL	3.56	42.5	48.5	48.5
AADVS	3.53	2.60	0.500	
BADVS	2.32	3.40	0.500	
ABDVS	1.82	2.00	0.500	
ASLA	1.72	2.14	1.00	0.200
SHRDRH	1.57	0.97		
ADDTMP	1.26	1.64	0.600	
FRDFCB	1.00	0.778		

For some of the parameters with large criterion values in the last evaluation the results of the first and the second set of runs are presented in Figures 3.1 to 3.4. The figures are selected to illustrate some properties of the analysis. First of all, Figure 3.1 illustrates that sometimes 'zooming in' on

the assumed discontinuity is not necessary, because the first set of data already allows a clear conclusion. Figure 3.2a shows a curious effect, because the interval chosen in the subsequent analysis (Figure 3.2b) is not the interval in Figure 3.2a that draws the most attention.

Figures 3.3 and 3.4 show the complex relations arising if a parameter is used in a state variable which in its turn is input in one or more switches. The parameter TSDVRR is used in the state variable DVS which is (among others) input to a number of senescence switches. The parameter TBSJUV is input to the state variable LAI, which is input to the switch from juvenile to mature leaf area growth.

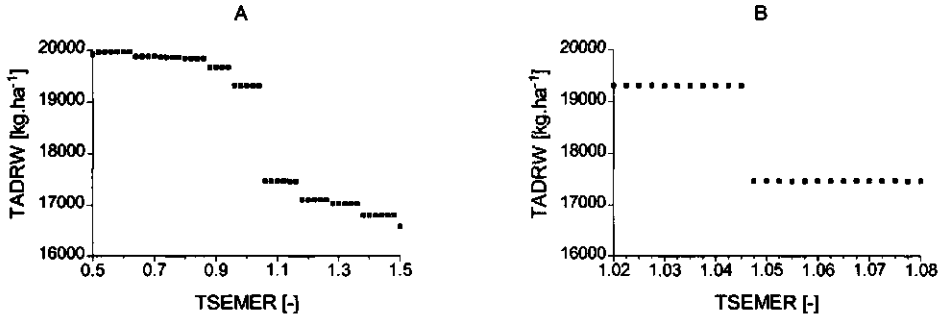

Figure 3.1 Result of the analysis of discontinuity for the relation of total aboveground dry matter weight (TADRW) and the parameter TSEMER (temperature sum from sowing to emergence; divided by its nominal value) over the full range (a) and focusing on the possible discontinuity (b).

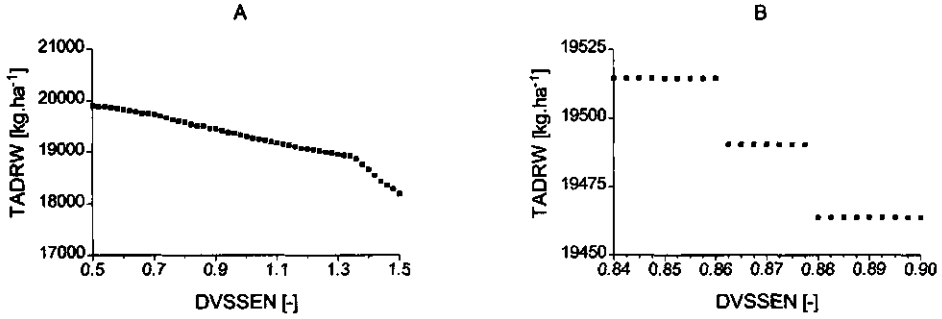

Figure 3.2 Result of the analysis of discontinuity for the relation of total aboveground dry matter weight (TADRW) and the parameter DVSSEN (onset of leaf senescence; divided by its nominal value) over the full range (a) and focusing on the possible discontinuity (b).

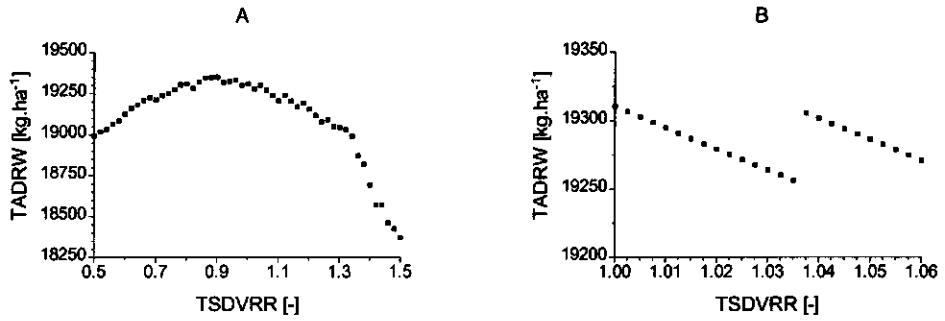

Figure 3.3 Result of the analysis of discontinuity for the relation of total aboveground dry matter weight (TADRW) and the parameter TSDVRR (Temperature sum from silking to maturity; divided by its nominal value) over the full range (a) and focusing on the possible discontinuity (b).

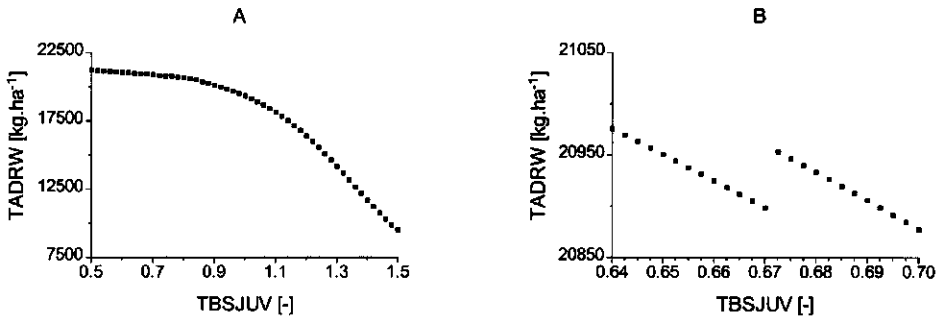

Figure 3.4 Result of the analysis of discontinuity for the relation of total aboveground dry matter weight (TADRW) and the parameter TBSJUV (Base temperature for juvenile leaf area development; divided by its nominal value) over the full range (a) and focusing on the possible discontinuity (b).

3.3.5 The effect of the parameters in SUCROS87 on final yield: relative extreme values

A qualitative analysis of the occurrence of extreme values of TADRW for a given parameter can be executed simultaneously with the analysis of continuity. For this qualitative analysis to be interesting, the extreme value should not be situated on the parameter interval bounds, but within the range of the parameter investigated (an internal extreme). Such an internal extreme is

interesting as it suggests that under the given meteorological conditions there are parameter values for which crop forage yield is maximal (or minimal). The occurrence of internal extremes will also be of some importance in Chapter 6.

The function evaluations used in the analysis of continuity were therefore also used to evaluate the occurrence of internal extrema in the relations between the state variable TADRW and the parameters used in SUCROS87. In the analysis no internal minima were observed. The parameters presented in Table 3.7 have a value for which the TADRW is maximal within the analysed range.

Table 3.7 Parameters for which internal maxima in TADRW (total above-ground dry weight) at final harvest were observed. For an explanation of the abbreviations cf. Appendix 1.

Parameter			
ABDVS	ADVRV	KDFTHE	TSDVRV
ADDTMP	ASLA	TSDVRR	TBSVRV
ADVRR	KDIF	TBSVRR	

3.4 Discussion and conclusions

The structural analysis of SUCROS87 suggests that it is possible to calibrate a small number of state variables (EMERG, DVS, TSUMEM and the closely associated LAI in the juvenile phase) sequentially, thus reducing the size of the calibration problem. Accepting a small dependency between different state variables does not increase the number of state variables that can be sequentially calibrated. For LINTUL the structural analysis suggests that it is possible to calibrate the state variables sequentially (Section 3.2). Suggested independencies should be checked against model source code. If structural properties are used in a calibration, the quality of fit of sequentially calibrated variables should be monitored. This is necessary to guard calibration quality against dependencies induced by parameter changes.

An analysis of relations between state variables and parameters shows that switches introduce discontinuities. The analysis of relations between TADRW (final yield) and model parameters shows that these relations are discontinuous for 15 parameters in SUCROS87. This means that the derivative of TADRW with respect to these parameters does not exist everywhere. This makes a sensitivity analysis on the basis of the derivatives of TADRW with respect to a parameter impossible (Sections 3.3.2-3.3.4). The results of this analysis are summarized in Table 3.8. Because of the occurrence of switches in LINTUL, discontinuities in state variable-parameter relations are to be expected in LINTUL as well.

Most often, the analysis of the relations between multiple parameters and multiple calibration variables is based on linearization which requires that derivatives are defined (Bard, 1974). As this is not the case for SUCROS87, and are to be expected in LINTUL, such an analysis is not possible (Section 3.3.4).

Given the problems caused by the discontinuities, it is conceivable to reformulate or modify

the model to remove the discontinuities from the relations between calibration variables and parameters. Discontinuities can be removed or their magnitude decreased, if the logical functions introduced in the model description are replaced by alternative switch functions with the following property: they gradually change their value from zero to 1 over a number of integration steps. The derivatives of this function should be defined for all values of the parameters used in the switch. Possible candidates for such a switch function are functions such as the logistic or alternatives proposed by Thornley and Johnson (1990), Richter and Söndgerath (1990). However, introduction of these functions introduces additional parameters, thus further increasing the number of parameters. From this point of view the approach is not attractive. Furthermore, changing the models solely for mathematical convenience, without a discussion of the conceptual justification of the switches or the use of alternatives is not justifiable. Even more important was the consideration that consequent analyses on the basis of a reformulated model would change the methodological choices made. It was therefore decided not to modify the models for this case study, and regard the source code as given.

A numerical analysis of the relations between TADRW and the parameters in SUCROS87 shows that these relations are not all strictly monotonous. The analysis shows that a number of parameters show an optimal yield within the investigated range (cf. Section 3.3.5). The results of both analyses are summarized in Table 3.8.

The number of simulations required to analyse the continuity of the relations between state variables and parameters is high (more than 3000 for SUCROS87). For the models used in this case study, this is not a problem, but it might be one for other simulation models. Therefore it is efficient to develop a single procedure which combines an analysis of the structure of the model, analysis of the continuity of the relation between state variables and parameters and the occurrence of relative extreme values in the relation between state variables and parameters.

Table 3.8 Summary of the analysis of SUCROS87. Characterization of the relation between the parameters and total aboveground dry matter weight (TADRW) at final harvest: discontinuous (D) and/or internal extreme values (E). Parameters not included in this table did not give rise to a detailed analysis of discontinuity, nor did they show a clearly defined internal extreme. For an explanation of the abbreviations cf. Appendix 1.

	D	E		D	E
TBSEM	*		LAIJUV	*	
TSEMER	*		RGRL	*	
ADVRV	*	*	TBSJUV	*	
TBSVRV	*	*	ABDVS		*
TSDVRV	*	*	ADDTMP		*
ADVRR	*	*	BDDTMP		
TBSVRR	*	*	DVSSEN	*	
TSDVRR	*	*	KDFTHE		*
ASLA		*	KDIF		*
DVSJUV	*		DVSRDR	*	

Chapter 4

Parameter uncertainty

4.1 Introduction

4.1.1 Parameter uncertainty and prediction

An important argument to choose a simulation model for predictive purposes is a high predictive quality. Predictive quality is determined by the quality of each of the steps leading to model application. Overviews of steps leading to model application were discussed in Chapter 1. When using a model for predictive purposes, one attempts to control the quality of these different steps, but complete certainty regarding the success of that attempt can not be offered. Residual uncertainty remains. Extensive classifications of this residual uncertainty are presented in literature (O'Neill and Gardner, 1979; Loehle, 1987). For a fully developed model where quality of all previous steps has been ensured, this uncertainty is reduced to uncertainty regarding the model structure and the model input (boundary conditions, parameter values and input data). In this thesis, the effect of uncertainty of parameter values on prediction quality will be analysed. Model structure and inputs (daily temperature and radiation) are assumed to be perfectly known in this chapter. The effects of uncertainty in input data will be discussed in Chapter 5.

In the previous chapter, a parameter is defined as a scalar which is constant during a single simulation run. But in reality parameter values will vary in time and from one instance to another. To assess prediction quality on the basis of relevant variation, it is necessary to observe the variation in the parameter values, to ascribe it to different sources and to quantify the variation. We will use the notion 'parameter variation' to summarize the lengthy 'variation in parameter values due to different sources of variation'. We will use 'parameter uncertainty' for that part of the parameter variation which is relevant for prediction quality, i.e. for the possible variation of parameters between system realizations (and the simulation of these systems using the model).

In a hierarchical approach to modelling the variation of a parameter value can be regarded as the output of a non-modelled subsystem. If formulation of submodels to replace model parameters is not possible or feasible, other descriptions of parameter variation have to be used. Modelling parameter variation takes many forms. The progression from empirical functions to extensive subsystem models, explaining the variation in the original parameter, is gradual. Look-up tables of parameter values are presented (e.g. Bignon, 1990; Hough, 1990). Empirical functions are used, such as pedotransfer functions, which translate soil characteristics to parameter values used in hydrological models (Wösten, 1988).

The following questions are subject of research in this chapter: how to describe parameter variation; how to determine parameter uncertainty, and what is the magnitude of parameter uncertainty for the parameters of the models SUCROS87 and LINTUL?

4.1.2 Describing parameter variation

The analysis of parameter variation and parameter uncertainty in this thesis is executed for parameters in two simulation models (SUCROS87 and LINTUL). Additional research to extend the two simulation models by dynamic simulation of the processes that are the cause of variation

in the value of the parameter is not an option in the context of this thesis. Therefore, parameter variation has to be described in terms of black box models.

One type of black box model which allows to describe parameter variation is a probability distribution. In that case, the variation of a single parameter is described in terms of a probability distribution, or if simultaneous observations of multiple parameters are available, a multivariate distribution. If a number of observations on the value of an individual parameter is available, a particular probability distribution should be chosen and estimated.

Sometimes too few observations are available to estimate a parameter distribution. For our application we do not expect this to be the case. If it occurs, the so-called set-theoretic approach provides an alternative. In this approach, parameter variation is described by a membership function. A membership function defines which parameter value belongs to the set of admissible parameter values. There are 'crisp' and 'fuzzy' sets. In crisp sets, the membership function can only assume two values: 0, inadmissible, or 1, admissible. In fuzzy sets, a membership function is defined which allows for intermediary degrees of admissibility. The value of the membership function is now specified as a value on the interval [0,1], where this value specifies the degree of membership.

Methods to formulate membership functions are introduced by Dubois and Prade (1989). Some examples of an application of fuzzy sets in an agronomic context are Kuzelka (1990), Woldt et al. (1995), and Paul and Witte (1995). An example of a set-theoretic approach to parameter uncertainty in crop growth modelling using so-called biologically plausible ranges (crisp sets) is presented e.g. by Klepper and Rouse (1991), and Klepper (1989).

The discussion which description of uncertainty to use for which type of problem is not yet solved (see e.g. Goicoechea, 1988; Dubois and Prade, 1989; Laviolette et al., 1995). Formal criteria to choose between a set-theoretic approach or a probabilistic approach are not available. Given that a lack of information regarding parameter values for maize is not expected, and that combining the two methods would yield undefined results, it was decided to describe parameter uncertainty in terms of probability distributions only.

4.1.3 Defining parameter uncertainty

A population of production instances

In this chapter, the uncertainty in potential yield of maize due to parameter uncertainty will be quantified. The potential production level was defined in Chapter 2. The characteristic temporal scale for the application of crop growth simulation models is the growing season; the spatial scale is a farm field. Realizations of potential maize production at these scales will be referred to as production instances. One production instance is a single unit of specific temporal and spatial scale, or, in other words, one growing season at one maize field. All possible production instances over the world jointly constitute the population which yields the parameter values relevant for this research. This population comprises different maize cultivars, cultivation practices and climates.

Models contain a number of scalars. Two types of scalars can be distinguished on the basis of their variation between production instances.

'Constants' are scalars which are known or defined not to vary between production instances. Examples are the Euler constant e, the ratio of the circumference of a circle to its diameter (π), the latitude of the tropic of Capricorn and of Cancer, the solar constant, but also the reference temperature at which the maintenance coefficients are defined.

The other type of scalars vary between production instances. These are called 'parameters'. An example of such a scalar is the maximum rate of leaf photosynthesis, which depends on many factors. An example of an inaccurately named scalar is the psychrometric 'constant' which, far from being a constant, is a non-linear function of atmospheric pressure. Parameter values may be calibrated using data for a single instance (cf. Chapter 6), constants may (and should) not be calibrated.

An example by Doucet and Sloep (1992) illustrates the classification of scalars and its dependence on model application: a model describing a rocket standing on earth uses a constant value of the gravity acceleration. In this case (a single instance) the gravity acceleration may be called a constant. If the same model is used to calculate and compare model results of rockets standing on different planets (multiple instances), gravity acceleration is a parameter. Finally, if the same model is used to describe a travelling rocket, gravity acceleration is no longer a scalar, but the output of a submodel.

4.1.4 Estimating parameter uncertainty

A description of parameter variation should account for variation between production instances and variation within production instances. We have defined a parameter to be a constant in (and during) a production instance. Uncertainty is then defined as the variation of a parameter between production instances. Constants are not uncertain, as they do not vary between instances. Variation within an instance is measurement variation (measurement error, or variation-within). This variation is 'noise', as we defined parameters to be constant within a production instance.

As parameter variation is described in terms of probability distributions, parameter uncertainty is measured in terms of the variance of the probability distribution. The reported value of a parameter (where reported refers to all 'somehow' determined values of a parameter; we will often use measured in this sense) can be described under a number of assumptions.

Assume that a production instance i is characterized by a parameter value π_i. Define the reported value for that instance (P_i) as π_i with a 'measurement' error η_i:

$$P_i = \pi_i + \eta_i \tag{4.1}$$

Define π_i as the sum of θ (θ is the mean over the population of instances), and a variable ε_i for each instance (ε_i describes the variation between instances):

$$\pi_i = \theta + \varepsilon_i \tag{4.2}$$

The combination of both definitions describes a parameter measured in a production instance i which varies over instances and is determined with a measurement error:

$$P_i = \theta + \varepsilon_i + \eta_i \tag{4.3}$$

Parameter uncertainty and the measurement error are assumed to be described by probability distributions with expected values of 0, and variances:

$$\begin{aligned} Var(\varepsilon_i) &= \tau^2 \\ ar(\eta_i) &= \sigma_i^2 \end{aligned} \tag{4.4}$$

The different distributions are assumed to be independent.

Estimators under the assumption of homogeneous measurement variance

If all measurement errors are assumed to be realizations from a single error distribution, the estimators derived then may serve to illustrate some of the consequences of the above description of parameter variation.

Homogeneous measurement variance means that $\sigma_i^2 = \sigma^2$ for all instances i. A measurement P_i is then described as:

$$P_i = \theta + \varepsilon_i + \eta_i \tag{4.5}$$

and has an expected value θ and a variance $Var(\varepsilon) + Var(\eta) = \tau^2 + \sigma^2$. We will assume that in each of the n instances there are m measurements.

In the following equations the hat (^) indicates the estimate, the bar ($\bar{}$) an average. The constant characterizing the population, $\hat{\theta}$, is estimated as the mean of all P_i:

$$\hat{\theta} = \frac{1}{n}\sum_{i=1}^{n} P_i = \overline{P} \tag{4.6}$$

This estimate is itself uncertain. Its variance, the variance of a mean, is:

$$ar(\hat{\theta}) = \frac{Var(P_i)}{n} = \frac{(\tau^2 + \sigma^2)}{n} \tag{4.7}$$

The variance of P_i is estimated as:

$$\hat{ar}(P_i) = \sum_{i=1}^{n} \frac{(P_i - \overline{P})^2}{n-1} \tag{4.8}$$

It contains both measurement variance and parameter uncertainty. The measurement variance can be estimated as the mean of all s_i^2, the reported measurement variance for the individual

instances:

$$\hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n} s_i^2 \tag{4.9}$$

To estimate parameter uncertainty ($\hat{var}(\pi_i)$) one should subtract the estimate of the measurement variance from the variance of P_i. To exclude negative (impossible) values of parameter uncertainty, a possible estimator is:

$$\hat{\tau}^2 = \max((var(P_i) - \hat{\sigma}^2);\ 0) \tag{4.10}$$

This is the parameter uncertainty in a descriptive situation. To describe the parameter uncertainty in a new instance, the uncertainty of the constant characterizing the population should be added:

$$\hat{MSEP}(\pi_i) = \frac{var(P_i)}{n} + \max((var(P_i) - \hat{\sigma}^2);\ 0) \tag{4.11}$$

The set of equations to analyse this type of data (parameter value in a production instance and its estimation error) are summarized in Table 4.1. The square root of Equation 4.11 is the root mean square error of prediction, the RMSEP, which is often used in this chapter.

Estimators under the assumption of heterogeneous measurement variance

If distributions of the measurement error vary between production instances, e.g. because different measurement methods are used in different instances, measurement errors in different instances will be realizations from different distributions. In those cases, the estimate of the mean and that of the uncertainty should be weighed according to measurement error. Estimates can be obtained using different statistical methods, such as e.g. residual maximum likelihood (REML) (Payne and Lane, 1993). REML will be used in the analysis of the data.

Table 4.1 Estimators to analyse uncertainty of scalar model elements (under assumption of homogeneous measurement variance).

Case	Estimation	Variance in a new situation (*MSEP*)
Constant: Only measurement error	$\hat{\pi}_i = \frac{1}{n}\sum P_i$	$\hat{MSEP}(\pi_i) = \frac{var\ P_i}{n}$
Parameter: Measurement error and parameter uncertainty	$\hat{\pi}_i = \frac{1}{n}\sum P_i$	$\hat{MSEP}(\pi_i) = \frac{var\ P_i}{n} + \max((var\ P_i - \frac{1}{n}\sum s_i^2);\ 0)$

4.1.5 Other aspects of parameter uncertainty

Evaluating the constancy of a scalar

Invariance over instances can not be proven. However, the data may provide an indication that invariance occurs. If sample variation is of the same magnitude as the measurement error, or in other words the quotient $(\tau^2+\sigma^2)/\sigma^2$ is close to 1, natural variation could be absent (e.g. Hedges, 1987), and the scalar is regarded as a constant, until further information becomes available. Of course the quotient may also be close to 1 because a parameter has been measured inaccurately. If the estimated sample variation $(\tau^2 + \sigma^2)$ is very large compared to the estimated measurement variance σ^2, natural variation is certainly present, and the scalar is to be treated as a parameter.

Correlation

Relations between reported parameter values may exist, either because of unmodelled relations or because of joint estimation from measurements. The simplest way to describe such a relation is in terms of correlation. Depending on the cause of their variation or the type of measurement errors made within an instance, parameters may be correlated within instances. The common and simplest assumption regarding correlation is that it is absent as long as it has not been established. Constants can not be correlated.

4.2 Methods to review parameter uncertainty

4.2.1 An overview

After defining the statistical model for the parameter uncertainty (Section 4.1), the methods to collect the data to determine parameter uncertainty have to be presented. Apart from the experimental determination of these data (primary research) the following three methods are possible:

1) Elicitation. The knowledge of experts is used to quantify the distributions of parameter values.
2) Reviewing parameter values presented in literature. A set of values from experiments presented in literature is used to quantify the distribution of a parameter.
3) Extended-model analysis. A set of parameter values is derived using a complex model. These values are used in a simple model for prediction.

In this section (Section 4.2) these three methods are presented and their advantages and disadvantages reviewed.

4.2.2 Elicitation of uncertainty using expert knowledge

When data on the parameters are missing or hard to come by, we are interested in methods to

quantify the knowledge of experts. This can be done using methods which allow to question persons (assessors, experts) in such a way that the parameter uncertainty can be quantified using a selected description of parameter uncertainty, in this thesis probability distributions. These methods are referred to as elicitation methods. If uncertainty is described in terms of probability distributions, the elicitation method should be formulated in such a way that the belief of a certain event occurring can be reformulated as a probability. This probability is referred to as subjective probability, because it is not based on observed frequency distributions. The Dutch Ministry of Health regards expert judgement as a possible option for risk assessment (VROM, 1989, p. 17). An extensive monograph on elicitation in a probabilistic context has been written by Cooke (1991).

Elicitation of parameter uncertainty, a short review

Two phases are distinguished in an elicitation session: preparation and execution.

In the preparation phase elicitation methods are chosen, and the questions to be asked are defined. The procedure is well-documented. The methods used are simple and preferably indirect, that is to say the expert is not asked to answer: What is the probability that...? Instead uncertainty is expressed in measures which can be translated in probabilities. If possible, several elicitation methods are used (Hogarth, 1975; Spetzler and Staël von Holstein, 1975; Lourens, 1984; Terlouw, 1989). The questions asked are carefully and concisely formulated and have been tested in a preliminary session. The duration of the session should be limited (Lourens, 1984). In the preparation for the session the interviewer selects a number of assessors. The task of the ideal assessor regularly includes assessments of the variables of interest within the target population. The ideal assessor evaluates his or her assessments on a regular basis, using empirical data. In an agricultural context, an example is the visual estimate of the yield of natural grassland presented by Tadmor et al. (1975).

During the execution phase a number of steps are distinguished (Spetzler and Staël von Holstein, 1975). In the first step, the interviewer should motivate the assessor and explain and discuss the approach. Before and during the session, the assessor has to be trained and reminded to present his or her knowledge and the associated uncertainty as 'proper' as possible. Proper is the technical term which indicates that the opinion has properties which allow to translate it in terms of a subjective probability. To ensure proper results, the methods proposed in literature and in some cases presented as software (e.g. van Lenthe, 1993) contain training questions in the form of a game, which assessors lose if they are overconfident or excessively uncertain.

After encoding his or her knowledge, the assessor is allowed to reconsider. The interviewer is present during the session, for clarification and to ensure that all questions regarding the uncertain quantities are unambiguously interpreted.

Examples

An example of parameter uncertainty elicitation for the type of models examined here is presented by van der Voet and Mohren (1994) for a simulation model of forest stand growth. Thirty-four parameter distributions and their correlations were calculated from expert know-

ledge. The variability of 6 other parameters was derived from measurements. To reduce the task, the nominal (default) values of the parameters were assumed to be equal to the mean of the distribution. The distribution was assumed to be normal and was transformed after assessment to satisfy theoretical bounds. The expert assessed either the upper or the lower 95% confidence bound of the parameter and the correlations between the parameters. The elicited parameter distributions were used in an uncertainty analysis of the model. A structured interview session in which the uncertain parameters of an atmospheric dispersion model were elicited is presented by Cooke (1992) and van Steen (1992). In a case study of fish response to lake acidification, Reckhow (1988) elicits parameter distributions. Other examples are presented by e.g. Kadane et al. (1980) for a linear model and Goossens et al. (1992) for a non-linear risk model.

4.2.3 Literature reviews

References in scientific articles, reviews and monographs provide a basis for parameter values e.g. in the form of an integrative research review (Cooper, 1984) defined as combining, integrating and summarizing information regarding hypotheses across multiple studies. A related notion is that of meta-analysis. Meta-analysis designates statistical methods and procedures to combine quantitative information across studies (Hedges and Olkin, 1985; Cook and Leviton, 1980). This notion of meta-analysis is used in medical science and in social sciences. In natural sciences no specific terminology is used for meta-analysis; the term parameter review will be used. Examples of parameter reviews are quoted by Hedges (1987).

The relatively recent introduction of meta-analysis in social and medical studies gave rise to methodological discussions in literature. In contrast, methodological evaluations of review sessions in natural sciences could not be found in literature. As parameter reviews in both branches of science serve the same purpose (Hedges, 1987) the description of the review methodology will be based on evaluations of review sessions in social sciences (Jackson, 1980; Cooper, 1984; Light and Pillemer, 1984).

Aspects of a review session

In this summary review, methodological guidelines are reformulated in terms of assessing parameter uncertainty.

The following tasks are distinguished within a review session: 1) problem formulation, 2) retrieving research results, 3) abstracting research results, 4) analysing review results using quantitative methods, 5) reporting.

1) Problem formulation. How can the parameter value be derived from the raw data; which methods are used. Do not restrict the definition of a parameter to values associated with a specific method of determining a parameter. The method(s) should be defined in the model documentation. Note that differences in method, in definition or formulation reflect methodological or conceptual uncertainty.
2) Retrieving research results. It is as important to indicate what was included as it is to indicate what was excluded. This will allow a critical assessment of the review by the reader and will

allow extending the review without duplication by future reviewers. Be overly inclusive, at least initially. Be exhaustive in the literature search and be explicit about the methods used to gather the literature; present information services used, keywords used. Be exhaustive in the type of publications searched.

3) Abstracting research results. Regard the review as if it were an experiment; try to define experimental factors which might influence the parameter value and regard the parameter to be analysed as the response. Be specific about experimental design; extract experimental factors from the sources as well. Note that the description of parameter uncertainty in Section 4.1.2 requires that standard errors of the measured parameter values (if available) are abstracted.

   A rather trivial, but necessary, advice is to check the abstracted results for copying errors and differences in units. Do not eliminate sources on the basis of a quality judgement. If systematic errors are assumed to occur, introduce the factor associated with the alleged error as an experimental factor in the analysis. Definition of these factors may already be possible in the phase of problem definition.

   Research results are sometimes presented in a way which requires additional (secondary) data analysis, to retrieve the parameter values. If secondary data analysis was necessary, describe the method(s) used.

4) Analyse review results using quantitative methods. Execute the analysis as if it were an experiment with treatments (experimental factors) and units (in our case production instances), but keep in mind that parameter reviews are observational research (Draper, 1987). This means that correlations suggested by the data should not be regarded as sufficient reason to believe that these associations always exist.

5) Reporting. Structure the presentation. Present parameter reviews as primary research papers with standard structure: introduction, methods, results, and discussion.

Examples

An example of a meta-analysis in both physical sciences and social science is presented by Hedges (1987). Some examples of reviews of parameter uncertainty are Bresler and Dagan (1988a,b) and Dagan and Bresler (1988, irrigation), Aggarwal (1995, crop growth), Rossing et al. (1994a, crop disease), Blower et al. (1991, AIDS epidemiology), Klepper (1989, ecosystem modelling). The parameter reviews quoted are part of papers dealing with uncertainty analysis. In reviews of parameter values not specifically aiming at parameter uncertainty, parameter values are presented as ranges derived from the source (Gosse et al., 1986) or as point values (Penning de Vries et al., 1989; Kiniry et al., 1989).

4.2.4 Extended-model analysis

The simultaneous existence of models which differ in complexity offers a possibility to investigate parameter uncertainty, as the output of a subsystem in a complex model may be modelled as a parameter in a simpler model. Generally, relations between complex and simple models will

not be that straightforward. However, the distinction between simple models and complex models could be used to evaluate the deterministic aspects of parameter uncertainty in the simpler model which is then used for prediction.

The methodology is based on the assumption that (the most important part of) parameter variation in the simple model is explained in the deterministic complex model. The complex model is then used to simulate an experiment which can be used to estimate the parameter distribution for the parameter used in the simple model. The following tasks are distinguished within extended model analysis: 1) model selection, 2) formulating the experimental design, 3) simulating pseudo-experimental data, 4) estimating parameter values from these data and analysing the results, and 5) Reporting. In the following, methodological suggestions are summarized.

Methodology

In the selection step existing models of different complexity are considered. In our case LINTUL is the simple model to be used for prediction, and SUCROS87 is the complex model to be used to simulate 'measurements'. The experimental design should be based on the operational definition of the parameter. This requires that input and output variables in the model are selected, a measurement schedule is developed, and the relevant variables are selected to become output. In the simulation step the complex model is used to generate a dataset according to the experimental design formulated. These data are used to estimate parameters and their uncertainty in the simple model. One must take care that this methodology does not become a purely academic exercise: the complex model has to be calibrated and validated for the application of interest.

An example

Examples for an analysis of parameter uncertainty on the basis of a comprehensive model and a model for application could not be found in literature. However, relatively complex models are sometimes used to draw conclusions regarding concepts used in simpler models. Spitters (1990) discusses the constancy of the light use efficiency (LUE) as used in LINTUL, on the basis of results derived from SUCROS87. On the basis of simulated data de Ridder and van Keulen (1997) derive a regression model.

The approach is illustrated for the light use efficiency LUE in LINTUL, using SUCROS87 as the complex model. The light use efficiency is determined from a linear regression between experimentally determined aboveground crop weight (TADRW) and cumulative intercepted PAR (Spitters, 1990). SUCROS87 was used to simulate such an experiment. The synthetic data used consist of total aboveground dry weight (TADRW) and cumulative intercepted PAR generated in 10-day intervals. The simulations are based on management practice near Wageningen for a number of years, the associated meteorological data, and the reference parameter values. Intercepted PAR is estimated assuming exponential extinction in the canopy and cumulated on a daily basis. Linear regression forced through the origin yields an estimate of LUE and its estimation error. The results are presented in Table 4.2. The estimation error is

relatively small. The residuals (results not shown) were sometimes systematic, due to the non-linearity of the relationship in SUCROS87 (Spitters, 1990).

In principle we have generated parameter estimates similar to those in Section 4.1.4, for a part of the target population. However, there is a conceptual difference: the estimation error now reflects structural differences between SUCROS87 and LINTUL, instead of variation within an instance. It is not relevant to regard this variation as measurement error or as natural variation. In the analysis of the results we choose to disregard this variation (assuming σ^2 to be 0). In that case, using the equations presented in Section 4.1.4, the estimator of the variance in a new instance becomes:

$$\hat{ar}(\pi_i) = (1+\frac{1}{n})\tau_e^2 = (1+\frac{1}{n})Var(P_i) \tag{4.12}$$

Analysing the data in Table 4.2, neglecting variance within a production instance, the LUE used in LINTUL can be estimated as 2.91 ± 0.07 gMJ^{-1} intercepted PAR (mean ± standard error) from data generated by SUCROS87 for Dutch meteorological conditions.

This number of data is not sufficient to fit a probability distribution. Given the facts that the light use efficiency should be positive on a seasonal basis, and that the theoretical upper bound is not yet known, the parameter values are assumed to be described by a gamma distribution, with the above mean and standard error.

Table 4.2 Values of the light use efficiency (LUE, g TADRW MJ^{-1}) estimated from simulated total aboveground dry matter weight (TADRW) and intercepted PAR with standard error of the parameter estimate for a number of years.

Year	LUE ($g\ MJ^{-1}$)	standard error ($g\ MJ^{-1}$)	Year	LUE ($g\ MJ^{-1}$)	standard error ($g\ MJ^{-1}$)
1973	2.82	0.02	1982	2.96	0.01
1974	2.83	0.02	1983	2.93	0.02
1975	2.86	0.01	1984	2.92	0.01
1977	2.97	0.02	1985	2.95	0.02
1978	2.79	0.02	1986	2.93	0.02
1980	2.93	0.01	1988	2.99	0.01
1981	3.01	0.01			

4.2.5 Concluding remarks

The methods to describe and quantify parameter uncertainty are summarized in Table 4.3.

Expert judgement provides and summarizes knowledge regarding parameters. As such it is a basis for estimates of parameter uncertainty, but a formal procedure to elicit this knowledge is rarely used. The review in Section 4.2.1 shows that formal procedures based on a theoretical

frame exist and have been executed in a number of cases. Software is available which allows efficient execution. In risk analysis (notably in nuclear technology) expert estimation of failure probabilities is considered to be inevitable. It requires expertise in developing and executing interview sessions.

Parameter reviews provide and summarize primary research results. Parameter reviews are the closest alternative to directly measuring parameter values, and offer insights in concepts and methods used. It often requires a number of pragmatic assumptions and secondary analyses, such as reading data off figures. If restricted to articles that were subjected to peer reviews, there has been a check on data quality.

The extended-model methodology uses and summarizes secondary research results. It requires a number of simulation studies, which are easily executed by scientists familiar with simulation models. The procedure is fast compared to the alternatives - provided that the methodology can be based on existing models. A further advantage is that the results are based on qualitatively consistent data, and can be used to specify parameter uncertainty for different experimental factors. However, the quality of the parameter uncertainty thus derived is determined by the conceptual completeness and the predictive quality of the complex model. More specifically, for the models analysed in this thesis, the consequence of applying this method is that we are only able to derive the uncertainty for LINTUL's parameters. Furthermore, as LINTUL differs from SUCROS87 only in its calculation of the total dry matter growth, uncertainty can only be established for some of the parameters.

The only methodology which allows a complete overview of parameter uncertainty for both models and requires no additional expertise is a parameter review. The combination of a parameter review with a description of parameter uncertainty in terms of probability distributions is the approach chosen in this thesis.

Table 4.3 Options to describe and determine parameter uncertainty. The order (left to right, top to bottom) reflects our preference.

	Probability distribution	Set-theoretic
Parameter review	approach chosen in this thesis	
Parameter elicitation		
Extended model analysis		

4.2.6 Parameter uncertainty in SUCROS87 and LINTUL: review methodology

Based on the above discussion of options, parameter uncertainty for both models (SUCROS87 and LINTUL) is based on a probabilistic description of parameter values retrieved from literature.

The search strategy to retrieve literature initially aimed at recovering specific literature for maize quoted in the basic description of SUCROS87 (Rabbinge et al., 1989). In the second step, for the evaluation of parameter uncertainty, the references in these articles were retrieved and reviewed.

Access to literature was through a bibliography, an abstract journal, two monographs, two reviews, a thesis (cf. Table 4.4) and different literature retrieval systems. Additionally all available volumes of the Agronomy Journal (from 1971 onwards) were searched. Less complete searches were made in Agricultural and Forest Meteorology, Crop Science, Plant and Soil, Netherlands Journal of Agricultural Science, and Agronomie. Relevant articles and references were retrieved and reviewed.

The procedure followed was intended to be exhaustive with respect to published literature in international or national journals, as available in Wageningen libraries. No attempt was made to be exhaustive in the retrieval of reports ('grey literature'). The actual search strategies for each parameter, characterized by keywords and years searched, are not available, as the search strategy was not formalized. The extent to which this study is exhaustive can therefore not be evaluated.

During the project the accessibility of literature changed and improved enormously. Apart from the Wageningen Agricultural University library catalogue search program and abstract journal originally used, the availability of general bibliographies on computer allowed very specific searches. The importance of this development for parameter reviews cannot be rated high enough, and allows a systematic and well defined procedure for literature retrieval and reviews of parameter uncertainty.

Table 4.4 Sources used as starting points for the literature review.

Anonymus, 1941. Maize bibliography for the years 1917 to 1936, inclusive. Contributions from Iowa Corn Research Institute, volume 2, nr 1.

Anonymus, 1948. Maize bibliography for the years 1888 to 1916, inclusive. Contributions from Iowa Corn Research Institute, volume 3, nr 2.

Anonymus, 1951. Maize bibliography for the years 1937 to 1945, inclusive. Contributions from Iowa Corn Research Institute, volume 3, nr 3.

Heemst, H.D.J. van, 1988. Plant data values required for simple crop growth simulation models: review and bibliography. (Simulation report CABO-TT 17) Wageningen: CABO, 100 p.

Ledent, J.F., 1978. Proceedings of the European maize meeting held at Louvain-la-Neuve. 18-19 october 1977. Louvain-la-Neuve: Université Catholique, 150p.

Picard, D., 1991. Physiologie et production du maïs: communications au colloque la vie du maïs, physiologie du maïs, application à la production, organisé par l'INRA, l'AGPM et l'Université de Paris-Sud. Pau, 13-15 november 1990. Paris: INRA, 501 p.

Schröder, J.J., 1991. (in Dutch). De benutting van stikstof door maïs met speciale aandacht voor de wortels. (Verslag 152) Wageningen: CABO, 53 p.

Struik, P.C., 1983. Physiology of forage maize in relation to its production and quality. PhD-thesis, Wageningen Agricultural University, 252 p.

Selection of production instances

Initially as much information as possible was gathered on the parameters. Given that potential yield is to be simulated, those parameter values were selected which were associated with the treatment resulting in the highest yields in case of the growth factors other than radiation and temperature. In cases where radiation and temperature were varied, all treatments were included in the analysis. Parameter values derived from experiments in glasshouses and climate chambers under non-limiting conditions were included in the sample. The reviews were not limited to a certain period. Experiments from 1980 were considered to be as relevant as experiments from 1880. In the analysis, data were discarded if definitions were not met (e.g. FRPAR based on a radiation range other than 400-700 nm). Parameter values derived using different methods were retained.

Methods of secondary analysis

Secondary analyses were required to derive parameter values from research which was executed for different purposes. The methods of secondary analysis ranged from linear to non-linear regression.

Fitting a distribution to the data

Distributions were chosen on the basis of the assumptions regarding theoretical bounds. If lower bounds are known, the gamma distribution was chosen; if upper and lower bounds are known, a beta distribution was chosen. For parameters related to crop development the fit of a distribution to the data was investigated.

Data analysis

The data consisted of parameter estimates, in some cases with their estimated precision. For the analysis the typical data structure was as follows (*: missing value):

observation of parameter value	measurement	std. error measurement
instance 1	P_1	s_1
instance i	P_i	s_i
instance n	P_n	*

Note that the notion 'measurement' is used to refer to sometimes very complex parameter estimation procedures. The estimated parameter value and its variance were determined and chosen according to the following branching decision process:

Whenever the number of instances in which the standard error of measurement was given at least equalled 2, the analysis was executed over the number of data with given standard error of the measurement. Other data were disregarded in this analysis. The retained data were tested

regarding the significance of uncertainty. If the uncertainty was non-significant, the analysis was continued as if uncertainty equalled zero. In that case the scalar is assumed to be a constant. If uncertainty was significant, the analysis was executed using REML.

A second analysis was executed for comparison. In that analysis the data for which the standard error of the measurement was missing were included in the analysis. The measurement precision was taken equal to the average of the given measurement precision and the analysis executed using REML.

A third analysis was executed when no measurement precision was given. In these cases the missing measurement precision was set to a negligibly small number (10^{-9}, depending on the computer) and executed using REML.

Large differences between estimates in the first, second and third analysis only serve to stress the importance of measurement error in this analysis. The program used was implemented in GENSTAT.

Correlation

Correlations between parameters were analysed using Spearman's rank correlation test (2-sided, 5%). Note that the correlation determined here refers to correlations between instances.

No values found

Those scalars for which no values could be retrieved from literature were defined to be constants and kept at the nominal values used in the original model.

4.3 Summary of the parameter uncertainty for SUCROS87 and LINTUL

The results of the parameter review for the defined population of production instances are summarized in Table 4.5. A short characterization of the population is 'parameter values for potential maize production instances all over the world'. The parameter specific procedures, the definitions and the results for individual parameters will be discussed in Section 4.4.

A number of scalars used in SUCROS87 could not be retrieved from literature. In the calculation of leaf senescence the parameters SHRDRL, SHRDRH, DVSRDR, TBSKIL, and DVSKIL, and in the calculation of maintenance and growth respiration the parameters MAINSO, MAINCB, MAINST, ASRLV, ASRCB, and ASRST could not be retrieved from literature. In partitioning (FSO and FCB) the number of data within each production instance was too small to allow fitting the partitioning functions to the data. These scalars are regarded as constants.

The remaining parameters can be grouped in terms of the processes they are associated with. These groups can be classified on the basis of the number of instances actually found, reflecting availability, and on the basis of the coefficient of variation, reflecting relative uncertainty. Based on group averages of data availability and relative uncertainty, the different processes are classified in Table 4.6 (SUCROS87) and Table 4.7 (LINTUL). Two classes are distinguished: below and above average.

Table 4.5 Results of the review of parameter values in SUCROS87 and LINTUL for the defined population. The mean, the coefficient of variation calculated from the root mean squared prediction error (cv, %), and the total number of data retained for analysis are given for SUCROS87 and LINTUL. The parameters are sorted according to their coefficient of variation. For explanation of the abbreviations and for units, cf. Appendix 1 (SUCROS87) and 2 (LINTUL).

parameter	mean	cv	n(all)	lower bound	upper bound	distribution
SUCROS87						
FRDFCB	0.06	143	4	0		gamma
FRDFCA	0.2	88	4	0		gamma
MAINRT	0.1	78	4	0		gamma
LAINI	0.0005	73	11	0		gamma
MAINLV	0.02	64	11	0		gamma
DLYATR	0.4	59	8	0		gamma
BSH	1.2	59	100	0		gamma
Q10	2.3	49	4	1		gamma
AMX	76	46	8	0		gamma
BSLA	1.2	43	19	0		gamma
AAMTMP	0.6	41	9	0		gamma
EFF	0.4	31	8	0	0.91	beta
ATRL	0.2	30	4	0		gamma
DVSSEN	1.1	29	10	0	2	beta
BLV	1.5	29	51	0		gamma
AAMDVS	0.7	23	2	0		gamma
TSDVRV	24	23	16	0		N(sqrt)
TSEMER	8	22	114	0		N(sqrt)
FRDFBB	1.8	21	4	0		gamma
TSDVRR	20.2	19	81	0		N(sqrt)
KDIF	0.67	13	14	0		gamma
FRDFBA	1.59	10	4	0		gamma
SCV	0.22	9	4	0	1	beta
FRPAR	0.47	5	12	0	1	beta
RGRL	0.028	4	17	0		gamma
ASRSO	1.39	4	2	1		gamma
ASRRT	1.59	3	14	1		gamma
LINTUL						
DVLGTH	2.3	85	10	0		gamma
FINTI	0.00008	52	20	0	1	beta
RI	10.9	31	20	0		gamma
DVHALF	1.9	22	10	0		gamma
LUE	3.0	12	134	0		gamma

Ideally data availability is high and relative uncertainty low for all processes. This is not the case. For both models the results show that knowledge regarding most of the processes is less available, and that relatively uncertain parameters are often relatively unknown. An interesting example is that of the group of parameters which describe radiation in the atmosphere in SUCROS87: these are not abundantly available, and relatively uncertain, but are nevertheless regarded as constants in the original source code.

Comparison of uncertainty and availability between both models is not possible, because the reviews are only partly independent. The average coefficient of variation over parameters is similar for both models (about 30%), but differs strongly between individual parameters.

Only a limited number of correlations between parameters could be estimated in the review. Correlations between parameters are presented in Table 4.8.

Table 4.6 A classification of parameter groups for SUCROS87, based on parameter availability (number of instances, average = 20) and relative uncertainty (c.v., average = 38 %). (Classification is based on values for the individual parameters averaged over the group).

Relative uncertainty	Availability of data	
	above average	below average
below average	- phenology (TSEMER, TSDVRV, TSDVRR)	- senescence* (DVSSEN) - radiation in canopy (KDIF, SCV) - assimilate requirement* (ASRRT, ASRSO)
above average	- partitioning* (BSH, BLV) - leaf area index (BSLA, RGRL, LAINI)	- maintenance respiration (MAINRT, MAINLV, Q10)* - radiation in atmosphere (ATRL, DLYATR, FRDFBA, FRDFBB, FRDFCA, FRDFCB, FRPAR) - photosynthesis (AMX, EFF, AAMDVS, AAMTMP)

* These groups contain parameters for which no information was available.

Table 4.7 A classification of parameter groups for LINTUL, based on parameter availability (number of instances, average = 42) and relative uncertainty (c.v., average = 32%). (Classification is based on values for the individual parameters averaged over the group).

Relative uncertainty	Availability of data	
	above average	below average
below average	- phenology (TSEMER, TSDVR**) - LUE	- FRPAR
above average		- senescence (DVLGTH, DVHALF) - light use efficiency reduction function (FINTI, RI)

** assumed to be equal to TSDVRV

Table 4.8 Spearman rank correlation between parameters used in SUCROS87 and in LINTUL.

Parameters	Correlation	Model
TSDVRV-TSDVRR	0.67	SUCROS87
LAINI-RGRL	−0.68	SUCROS87
RI-FINTI	−0.72	LINTUL
DVLGTH-DVHALF	0.93	LINTUL

4.4 A review of the uncertainty of parameters used in SUCROS87

In this section, the parameter values and their uncertainty for the model SUCROS87 are reviewed on the basis of data in literature and summarized in terms of probability distributions. The definitions for the analyses are given in Section 4.1. The motivation for this approach was described in Section 4.2, and Section 4.3 presented the summary of the results. References used are presented separately for each parameter at the end of this chapter.

Two sets of results are presented for each parameter. The first is the result of the analysis in which parameter estimates without measurement error were discarded. These were presented in Table 4.6. The second set of results are those from the analysis based on all instances retrieved from literature using average measurement precision in all instances. They are presented for comparison only, and are shown in brackets in Tables 4.12 - 4.30.

4.4.1 Parameters associated with development stage and emergence

Phenological development stage in the period between sowing and emergence is represented by the state variable EMERG. Calculation of the rate of development DEMERG requires one parameter, the temperature TBSEM, a so-called base temperature. The parameter TSEMER is required to flag 50% emergence of the crop.

Development of the crop after emergence is determined by the development rate which requires the following parameters for its calculation: TBSVRV, the base temperature before silking (female flowering), TBSVRR, the base temperature after silking. TSDVRV, the increase in development stage per degree Celsius above the base temperature between emergence and flowering, and TSDVRR, the increase in development stage per degree Celsius above the base temperature between silking and physiological maturity. The nominal values of these four parameters are equal. ADVRV and ADVRR are scalars which allow to incorporate non-linear responses to temperature (cf. Appendix 7). They were set at a constant value of 1, and not further considered.

A data file was generated containing different types of development observations. Data on emergence, tassel initiation, cob initiation, tassel emergence, tasselling, silking, and maturity were included. Temperature sums presented as averages over years, locations and varieties were excluded.

Parameter review

A wealth of methods to calculate temperature sums has been proposed. Comparisons between methods are presented by several authors (Gilmore and Rogers, 1958; Cross and Zuber, 1972; Mederski et al., 1973; Derieux and Bonhomme, 1982a,b, 1990). The approach used in SUCROS87 is common and referred to as the GDD-method. It has a base temperature of 10°C (TBSVRV, TBSVRR). A modification of this method, in which temperatures above 30°C also result in halting phenological development, is the so-called MGDD-method.

Due to the variety of methods encountered in literature in which the base temperatures for development are regarded as fixed, the need was felt to choose a single method on the basis of which all data could be presented. Whenever possible the available data were transformed to a temperature sum with a base temperature of 10°C. The literature sources for which this was not possible were discarded.

The consequence of this transformation is that errors may have been introduced during the transformations from temperature sums at one base temperature to temperature sums at another base temperature. In the primary data and the following analyses MGDD and GDD temperature sums are not distinguished; this may have increased uncertainty.

The definition of maturity used in the selected values is the date at which dry matter accumulation ceases (Hanway, 1963). This is also known as physiological maturity. Physiological maturity is equated with the harvest date of silage maize (Groot et al., 1986; Daynard, 1972). Under Dutch conditions, the optimal harvest date for silage maize coincides with physiological maturity (van der Schans et al., 1993). Coincidence of optimal harvest date with

physiological maturity is also suggested by Bunting (1976) and Groot et al. (1986).

The number of data remaining after the different operations are presented in Table 4.9. The following parameters (all temperature sums) are used: sowing to 50% emergence (TSEMER); 50% emergence to 50% silking (TSDVRV); and 50% silking to maturity (TSDVRR).

Table 4.9 Overview over the number of development data.

Development period	Number of data
sowing to emergence	114
sowing to tasselling	3
sowing to silking	224
emergence to tasselling	147
emergence to silking	16
emergence to physiological maturity	3
tasselling to silking	17
tasselling to physiological maturity	3
silking to maturity	81

Transformation and fitting of distributions

To be able to summarize parameter uncertainty in the parameters for crop development we assumed that the temperature sums were determined with negligible measurement error. No theoretical considerations are known which would lead us to expect a certain type of distribution for the uncertainty in the development parameters. Initially we assumed a normal distribution. However, the data showed the distribution to be skew and to have a degree of kurtosis (skewness characterizes the asymmetry of a distribution around its mean; kurtosis characterizes the peakedness or flatness of a distribution; both properties are dimensionless and defined relative to a normal distribution (e.g. Press et al., 1992)). A normal distribution could, therefore, not be used to summarize the data. A square root transformation of the data lead to a normal distribution with no significant skew, and no significant kurtosis. The parameter values have to be retransformed for use in the model.

Results

The final results of the analysis for the non-transformed data are presented in Table 4.10. As argued in the previous section a square root transformation allowed a better description of the data by a normal distribution. The transformed results are presented in Table 4.11 (note that these are the results of a separate analysis, and are not directly derived from Table 4.10). Analysis of correlations between different parameters are presented in Table 4.12.

Table 4.10 Parameters characterizing the distribution of phenology parameters (no transformation), and parameter values as used in SUCROS87.

Parameter	TSEMER	TSDVRV	TSDVRR
mean	61	599	423
number of data	114	16	81
SUCROS87	80*	425	425

* based on Sibma (1987).

Table 4.11 The uncertainty of the phenological parameters (square root transformation) in SUCROS87, as used to generate the parameter values for the uncertainty analysis.

	$\hat{\pi}$	RMSEP	number of data	unit
TSEMER	7.62	1.66	114	$(°Cd)^{½}$
TSDVRV	23.94	5.46	16	$(°Cd)^{½}$
TSDVRR	20.20	3.89	81	$(°Cd)^{½}$

Table 4.12 Correlation between parameters used in development.

	TSEMER	TSDVRV	TSDVRR
TSEMER	-		
TSDVRV	0	-	
TSDVRR	0	0.67	-

4.4.2 Parameters associated with light in the canopy and atmosphere

The theoretical behaviour of light in a crop and its relation to photosynthesis as used in SUCROS87 is presented in a series of publications (Goudriaan, 1986; Spitters et al., 1986; Spitters, 1986; Goudriaan, 1988, and e.g. Goudriaan and van Laar, 1994).

Leaf scattering coefficient (SCV)

The leaf scattering coefficient is calculated as the sum of the leaf hemispherical transmittance and reflectance. An overview of theoretical aspects is given in e.g. Goudriaan (1977), and Bunnik (1978). The scattering coefficient is the average of the scattering coefficient over different wavelength intervals in visible light. Results are presented by Maas and Dunlap (1989), Sellers (1989), Yocum et al. (1964), and Woolley (1971). Changes in leaf chlorophyll content and leaf moisture content affect the scattering coefficient. The variability of this parameter is presented in Table 4.13.

Extinction coefficient of canopy (KDIF)

Light extinction in the canopy of both direct and diffuse radiation obey an exponential law as a function of leaf area index. Its magnitude depends on the type of radiation considered. For diffuse light extinction, the parameter in this exponential function is the effective extinction coefficient (KDIF). KDIF as used in SUCROS87 should be derived from measurements of extinction of photosynthetically active radiation (PAR) under diffuse light conditions (Rabbinge et al., 1989). These measurements are rare. A single reference (Allen et al., 1964) allows to estimate KDIF as being roughly 0.66 under these conditions. Extinction coefficients presented in literature are often derived from different combinations of diffuse and direct radiation. The extinction coefficient is calculated as:

$$KDIF = -\frac{1}{LAI} \ln\left(\frac{PAR}{PAR_0}\right) \tag{4.13}$$

The difference between the extinction coefficient based on absorbed PAR (APAR) and on intercepted PAR (IPAR) is assumed to be small as the difference between absorbed and intercepted PAR is small (Gallo and Daughtry, 1986). This allows to equate extinction coefficients derived from APAR measurements and those derived from IPAR measurements.

The extinction coefficient for total global radiation and the extinction coefficient for PAR are related: the extinction coefficient for total global radiation is about 3/4 of the extinction coefficient for PAR. The difference is caused by differences in scattering coefficient for different wavelengths (Monteith and Unsworth, 1990). Planting density (row width) influences KDIF through a change in the leaf angle distribution (Flénet et al., 1996). KDIF also changes in the course of development (Birch, 1996).

Parameter review

Varlet-Grancher et al. (1989) offer a review of values of the extinction coefficient for different crops. Other sources allowed to estimate KDIF using the equation given above. Data from literature were selected if the effective extinction coefficient was based on intercepted PAR measured around solar noon on a single day. A selection on the basis of fraction diffuse and direct radiation was not possible. The results are presented in Table 4.13.

Fraction photosynthetically active radiation (FRPAR)

The fraction photosynthetically active radiation (FRPAR) is that fraction of global radiation which activates photosynthesis. FRPAR is defined as the fraction of the global radiation within the range of wavelengths between 400 and 700 nm. FRPAR-values based on other wavelength ranges were not used in this review. The nominal value for FRPAR in SUCROS87 is 0.5. The value of FRPAR has been reviewed by Varlet-Grancher et al. (1989) and Lauciani and Ponticiello (1993). Stigter and Musabilha (1982) show an increase in FRPAR with an increase in the ratio of diffuse radiation to global radiation. Britton and Dodd (1976) show an increase of

FRPAR with a decrease in incident radiation. Part of the variation in FRPAR may be systematical and related to the net effect of the following three factors (e.g. Monteith and Unsworth, 1990): 1) preferential absorption of infra-red radiation by cloud droplets, 2) preferential scattering of visible radiation, also by cloud droplets, 3) the decrease of scattering with solar elevation. The results of the analysis are presented in Table 4.13.

Table 4.13 The uncertainty of the parameters SCV, KDIF and FRPAR.

	$\hat{\pi}$	RMSEP	number of data	unit
SCV	0.22	0.02	2	-
	(0.21)	(0.02)	(4)	
KDIF	0.67	0.09	14	ha^2 (ground) ha^{-2} (leaf)
FRPAR	0.47	0.02	9	-
	(0.46)	(0.02)	(12)	

Parameters determining the ratio diffuse over total global radiation

Distinguishing between diffuse and direct radiation is necessary. Neglecting diffuse radiation underestimates crop photosynthesis. Diffuse radiation is calculated from the global radiation using the fraction diffuse radiation. The fraction diffuse radiation is calculated using an empirically established piecewise defined function (De Jong, 1980; Spitters et al., 1986, Eqns. 20a-c, also e.g. Kropff et al. 1994) with the ratio of global radiation (measured) over the radiation outside the atmosphere (calculated) as its argument. The function is evaluated at selected solar hours to calculate diffuse and direct radiation as input to the photosynthesis subroutine. The equations and the nominal value of these parameters are based on de Jong (1980):

$$F_1 = \frac{S_{df}}{S_g} = \begin{cases} 1 & \text{for} \quad \frac{S_g}{S_0} \le 0.22 \\ 1 - 6.4\left(\frac{S_g}{S_0} - 0.22\right)^2 & \text{for} \quad 0.22 < \frac{S_g}{S_0} \le 0.35 \\ 1.47 - 1.66\frac{S_g}{S_0} & \text{for} \quad \frac{S_g}{S_0} > 0.35 \end{cases} \tag{4.14}$$

This function (F_1) was reformulated in SUCROS87 as follows:

$$F_S = \frac{S_{df}}{S_g} = \begin{cases} 1 & \text{for} \quad \frac{S_g}{S_0} \le \text{ATRL} \\ 1 - \text{FRDFAA}\left(\frac{S_g}{S_0} - \text{ATRL}\right)^2 & \text{for} \quad \text{ATRL} < \frac{S_g}{S_0} \le \text{ATRM} \\ \text{FRDFBA} - \text{FRDFBB}\frac{S_g}{S_0} & \text{for} \quad \frac{S_g}{S_0} > \text{ATRM} \end{cases} \tag{4.15}$$

where

S_g	Global radiation (J m^{-2} s^{-1}) at selected solar hour, estimated from the daily measured values
S_0	Radiation at the top of the atmosphere (J m^{-2} s^{-1}) at selected solar hour, based on theoretical calculations
S_{df}	Diffuse radiation (J m^{-2} s^{-1}) at selected solar hour
F_S	Fraction diffuse radiation (-)

and where

ATRL	coefficient determined using regression (0.22)
ATRM	coefficient determined using regression (0.35)
FRDFAA	coefficient determined using regression (6.4)
FRDFBA	coefficient determined using regression (1.47)
FRDFBB	coefficient determined using regression (1.66)

In SUCROS87 the function is further modified in comparison to Spitters et al. (1986) to yield the fraction of diffuse radiation:

$$\text{FRDIF} = \max\left[F_S;\ \text{FRDFCA} + (1 - \text{FRDFCA})\left(1 - e^{-(\text{FRDFCB}/\sin\beta)}\right)\right] \tag{4.16}$$

where

$\sin\beta$	the solar elevation at the selected solar hour (-)
FRDFCA	coefficient determined using regression (0.15)
FRDFCB	coefficient determined using regression (0.1)

The function F_S contains a large number of parameters, the number of which can be reduced under conditions of continuity. This function and its first derivative should be continuous at ATRM. This yields the following conditions:

$$\text{FRDFBA} - \text{FRDFBB} * \text{ATRM} = 1 - \text{FRDFAA} * (\text{ATRM} - \text{ATRL})^2 \tag{4.17}$$

and

$$\text{FRDFBB} = 2 * \text{FRDFAA} * (\text{ATRM} - \text{ATRL}) \tag{4.18}$$

These conditions allow to express FRDFAA and ATRM as functions of the other parameters:

$$\text{ATRM} = \text{ATRL} + \frac{\text{FRDFBB}}{2\text{FRDFAA}} \tag{4.19}$$

and

$$\text{FRDFAA} = \frac{\text{FRDFBB}^2}{4(\text{FRDFBA} - 1 - \text{FRDFBB} * \text{ATRL})} \tag{4.20}$$

An additional condition (FRDFAA > 0) imposes the following condition on ATRL:

$$ATRL < \frac{FRDFBA - 1}{FRDFBB} \tag{4.21}$$

The conditions allow to eliminate two parameters. The conditions would furthermore ensure that when executing an uncertainty analysis for the parameters in this function one does not introduce discontinuities.

Parameter review

In SUCROS87, the fraction diffuse calculated (FRDIF) is used to estimate hourly values of diffuse and direct radiation. Literature data for hourly measures were selected and reviewed. The most important difference between the different functions presented in literature is the inclusion of the solar elevation angle as a regression variable. Those results were retained in which the fraction of diffuse radiation was also a function of solar elevation.

The parameters of the function used in SUCROS87 were fitted on functions presented in literature. The point estimates of the parameters were averaged over solar elevation and the resulting variance regarded as variation within a situation. The results for these functions are presented in Table 4.14. Missing variation within a situation was due to discontinuities which occurred if the solar elevation was varied. Correlations between the parameters were found to be not significant; the dataset was too small.

Table 4.14 The uncertainty of the parameters describing the relation between the fraction diffuse radiation and the fraction global radiation as used in SUCROS87.

	$\hat{\pi}$	$\hat{RMSEP}$	number of data	unit
ATRL	0.24	0.07	3	-
	(0.23)	(0.06)	(4)	
FRDFBA	1.6	0.16	3	-
	(1.6)	(0.14)	(4)	
FRDFBB	1.8	0.39	3	-
	(1.8)	(0.32)	(4)	
FRDFCA	0.17	0.15	3	-
	(0.17)	(0.12)	(4)	
FRDFCB	0.07	0.09	3	-
	(0.07)	(0.08)	(4)	

The dependence of atmospheric transmission on solar elevation

In the estimation of global radiation at selected solar hours a sinusoidal course of radiation over the day is assumed to calculate instantaneous values. Transmission through the atmosphere

influences the instantaneous value. Transmission depends on solar elevation. (Spitters et al., 1986). The correction of the instantaneous radiation for the course of transmission over the day is based on the following relation:

$$\frac{S_g}{S_0} = a + b \sin\beta \tag{4.22}$$

where

S_g	Instantaneous global radiation (J m^{-2} s^{-1})
S_0	Instantaneous radiation at the top of the atmosphere (J m^{-2} s^{-1})
a,b	Coefficients determined using regression
$\sin\beta$	The instantaneous solar elevation (–)

The ratio b/a is the parameter DLYATR used in SUCROS87, with a nominal value of 0.4 (Spitters et al., 1986) for solar elevation angles larger than 20 degrees. Data which allow to determine this parameter are presented by Lumb (1964), also quoted in Monteith and Unsworth (1990) for different cloud types. Re-interpretation of data presented by de Jong (1980) and other literature sources (Carroll, 1985; Olseth and Skartveit, 1993) yield values of this parameter. The parameter was assumed to be estimated without measurement error. Results are presented in Table 4.15.

Table 4.15 The uncertainty in the parameter describing the dependence of atmospheric transmission on solar elevation (DLYATR) used in SUCROS87.

	$\hat{\pi}$	RMSEP	number of data	unit
DLYATR	0.42	0.25	8	-

Leaf angle distribution: leaf inclination and leaf azimuth distribution

Light absorption and interception characteristics of the canopy are determined by leaf properties and by the orientation of leaves. The orientation has to be characterized in an horizontal plane (leaf azimuth, 0-360 degrees, 0 = North), and in a vertical plane (leaf inclination, 0 (leaf horizontal) to 90 degrees (leaf vertical)). Both are characterized in terms of the relative frequencies of different leaf angles.

In SUCROS87 the leaf surfaces are assumed to be randomly arranged on a sphere. The resulting azimuth angle distribution is uniform; the resulting leaf inclination distribution is spherical. On the basis of this assumption, two parameters are required: OGEM and KDFTHE.

The parameter OGEM is the ratio of the surface area of the sunlit leaves to its projection, the shaded surface (cf. Goudriaan, 1988). For a spherical distribution this ratio is 0.5, and independent of the solar elevation. For other leaf inclination distributions the parameter OGEM depends on the solar elevation. If the leaf azimuth distribution is not uniform, the solar azimuth

angle (sun path) would also have to be taken into account. Defining other leaf inclination distributions would require that we incorporate the dependence of OGEM on solar elevation. A detailed description is given by Goudriaan (1988).

KDFTHE is the effective extinction coefficient in the leaf canopy for diffuse radiation from a uniform overcast sky assuming a spherical distribution of the leaf angles in the crop canopy. The theoretical relation between leaf area index and relative radiation can be written as a sum of exponential functions which describes extinction in the canopy (Goudriaan, 1988). This compound function can be approximated by a single exponential function defined over a specified interval. The extinction coefficient used in this approximation is the parameter KDFTHE. Over the interval LAI 0-3, the nominal value of KDFTHE is estimated as 0.80, using non-linear regression (Goudriaan and van Laar, 1994). The value of KDFTHE has to be revised for different leaf inclination distributions, according to Goudriaan (1988).

For both parameters a leaf inclination distribution other than spherical requires including a dependency on solar elevation in the model. As we prefer not to modify the model, the parameters are assumed to be constant.

4.4.3 The parameters used in leaf photosynthesis

Leaf photosynthesis in SUCROS87 is separately calculated for areas which are sunlit, and for areas which receive diffuse radiation using the following equation:

$$PL = \mathrm{AMX} \cdot \mathrm{AMTMP} \cdot \mathrm{AMDVS} \cdot \left(1 - e^{-\left(\frac{\mathrm{EFF}}{\mathrm{AMX} \cdot \mathrm{AMTMP} \cdot \mathrm{AMDVS}} \cdot \mathrm{PAR_x} \right)} \right) \qquad (4.23)$$

where

PL	Leaf photosynthesis (kg CO_2 ha^{-1} (leaf) h^{-1})
AMX	Potential CO_2 assimilation rate at light saturation for individual leaves (kg CO_2 ha^{-1} (leaf) h^{-1})
AMDVS	Factor accounting for effect of development stage on AMX (-)
AMTMP	Factor accounting for effect of daytime temperature on AMX (-)
EFF	Initial light use efficiency for individual leaves ([kg CO_2 ha^{-1} (leaf) h^{-1}][J m^{-2} (leaf) s^{-1}]$^{-1}$)
PAR_x	Flux density of incoming photosynthetically active radiation, either direct or diffuse (J m^{-2} (leaf) s^{-1}), using the relevant extinction coefficient.

Theoretical limits to leaf photosynthesis

Initial light use efficiency for individual leaves (EFF)

Theoretical limits to the initial light use efficiency EFF exist. The upper limit is set by the number of light quanta that are necessary to reduce one molecule CO_2. The theoretical value can

be calculated as (de Wit et al., 1978):

$$EFF = \frac{abtm\lambda}{nhcN_a} \tag{4.24}$$

where

EFF	Initial light use efficiency ([kg CO_2 ha^{-1} h^{-1}][J m^{-2} s^{-1}]$^{-1}$)
a	Conversion factor (10^{-3} kg.g^{-1})
b	Conversion factor (10^4 m^2.ha^{-1})
t	Conversion factor (3600 s.h^{-1})
m	Conversion factor (44 g.mol^{-1})
λ	Light wavelength ($550*10^{-9}$ m)
n	Number of light quanta required to reduce a molecule CO_2 (-)
h	Planck's constant ($6.626*10^{-34}$ J.s^{-1})
c	Light velocity ($3*10^8$ m.s^{-1})
N_a	Avogadro's number ($6.0225*10^{23}$ mol^{-1})

Substitution of the different factors shows the value of EFF to be equal to $7.277/n$ in the units used above. Lawlor (1987, page 50) states that "it is now accepted that eight photons are needed per CO_2 as a theoretical minimum and more may be required, depending on conditions". Hence, values of EFF above 0.91 (units as above and evaluated at λ=550 nm) are incompatible with present theories. As the lower limit of EFF is 0, EFF is bounded by 0 and 0.91.

Maximum gross photosynthetic rate for leaves (AMX)

The maximum photosynthesis rate of leaves results from biochemical and physiological processes in cells, and from the geometry of the leaves. Some upper bounds of AMX for a C_4 crop like maize are given in literature. Collatz et al. (1992) use a maximum fixation capacity of 143 kg CO_2 ha^{-1} h^{-1} in their model for photosynthesis of a C_4 crop. A maximum value of 118 kg CO_2 ha^{-1} h^{-1} for a C_4-grass is used by Chen et al. (1994). Whereas AMX can not be negative, the theoretical limit of the maximum photosynthesis rate appears to be variable.

Parameter review

To determine the uncertainty in the parameters AMX, AMDVS, and AMTMP the following data are extracted: the average temperature at which the experiment was executed; the development stage (estimated from phenological stage, based on Groot et al., 1986); the value of apparent leaf photosynthesis in kg CO_2 ha^{-1} leaf h^{-1}; its standard sample error; the number of days after sowing; the pre-treatment temperature; the radiation (W m^{-2} PAR); and the dark respiration rate (kg CO_2 ha^{-1} (leaf) h^{-1}). The initial photosynthetic efficiency, EFF and its standard sample error was also extracted from literature. Leaf photosynthesis PL is calculated from apparent leaf photosynthesis (PL_A), the measured variable and dark respiration R_d as:

$$PL = PL_A + R_d \tag{4.25}$$

In the calculations we assume that dark respiration R_d is only maintenance respiration. R_d then reacts to temperature in the same way as maintenance respiration:

$$R_d = R_{d25} \times Q_{10}^{(T_a-25)/10} \tag{4.26}$$

where

R_{d25}	Dark respiration at 25 °C (kg CO_2 ha^{-1} (leaf) h^{-1})
Q_{10}	Temperature coefficient (-)
T_a	Experiment temperature (°C)

No data-sets were found which allow simultaneous estimation of all parameters as a function of measurement temperature, development stage and radiation. Literature was selected in which AMX and EFF were determined simultaneously. Leaf photosynthesis was estimated using the average value over available dark respiration values. The conditions imposed were optimal temperature, as defined by the nominal values of AMTMP used in SUCROS87; optimal DVS, as defined by the nominal values of AMDVS used in SUCROS87. Only radiation levels higher than 300 W m^{-2} PAR were accepted.

The dependence of AMX on temperature (AMTMP) and development stage (AMDVS) was determined separately. For both AMDVS and AMTMP a single scaling factor (AAMDVS resp. AAMTMP) was estimated which modifies the response value (Appendix 7).

The data set was divided in one that would allow to determine the relation between measurement temperature and leaf photosynthesis and one that would allow to determine the relation between development stage and leaf photosynthesis. Leaf photosynthesis was estimated using interpolated dark respiration values in the case of the dependence on temperature and the average leaf dark respiration value in the case of the dependence on development stage.

Since temperature and development stage are not independent in field experiments, only two experiments could be retained, which allowed to determine AMDVS independent of AMTMP. Other - laboratory - experiments were available in which both development stage and temperature were varied. However, in those cases temperature and development stage were varied over ranges which according to the nominal tables used in SUCROS87 would show no effect on leaf photosynthesis. Simultaneously fitting AMDVS and AMTMP was not considered, but could increase the number of instances retained. The two transfomation parameters, AAMDVS and AAMTMP, were estimated from the retained data. The correlation between AMX and EFF was not significant; other correlations could not be investigated. Results are presented in Table 4.16.

4.4.4 Respiration parameters, maintenance coefficients and Q10

The calculation of the crop assimilate requirement requires the values of 5 parameters: ASRLV, ASRST, ASRCB, ASRRT and ASRSO. The parameters ASR'x' are the assimilate requirements for the different crop organs. The 'x' are leaves (LV), stem (ST), cob (CB), storage organs (SO) and roots (RT). Maintenance respiration requires the same number of parameters MAINLV,

Table 4.16 The uncertainty of parameters in leaf photosynthesis used in SUCROS87. EFF is the initial light use efficiency and AMXis the potential CO_2 assimilation rate at light saturation, both for individual leaves. AAMDVS and AAMTMP are scaling parameters introduced to transform the tabulated functions for the dependence of AMX on development stage and temperature respectively.

	$\hat{\pi}$	RMSEP	number of data	unit
AMX	76.	35.	7	$kg(CO_2)\ ha^{-1}(leaf)\ h^{-1}$
	(71.)	(34.)	(8)	
EFF	0.37	0.11	7	$[kg(CO_2)\ ha^{-1}(leaf)\ h^{-1}]\cdot$
	(0.37)	(0.11)	(8)	$[Jm^{-2}(leaf)s^{-1}]^{-1}$
AAMDVS	0.72	0.17	2	-
AAMTMP	0.56	0.23	9	-

MAINST, MAINCB, MAINRT and MAINSO, and an additional parameter, Q10, which characterizes temperature dependency of maintenance respiration.

Parameter review

Literature was reviewed for data regarding the assimilate requirement for the different organs, ASRQ'x', the maintenance coefficients for different organs, MAIN'x', and the parameter Q10.

In literature maintenance respiration coefficients could be found for grain (storage organs, MAINSO) and roots (MAINRT). Other maintenance coefficients (MAINCB, MAINST, and MAINLV) could not be retrieved from literature.

The literature review of the assimilate requirement coefficients or its inverse (1/ASRQ), the conversion efficiency Y_g, yielded values for roots (ASRRT) and grains (ASRSO). Direct measurement of the other coefficients (ASRLV, ASRST, and ASRCB) could not be retrieved from literature. A theoretical approach to calculate the parameter values on the basis of composition of dry matter is possible. An empirical regression between assimilate requirement coefficients and ash content is also available. These require modification of the model.

The value of Q10 was determined in the following manner: 1) an analysis of measured maintenance respiration vs. temperature; 2) an analysis of the maintenance coefficients vs. temperature. The only source found in which the temperature dependency of maintenance respiration is presented is the analysis by Kaše and Čatský (1984), for maize leaves. The result for Q10 (presented in Table 4.17) was based on a pooled analysis of these two types of determinations, and to the largest extent reflects the uncertainty inherent in the second method. Correlations between parameters could not be investigated. The results of the literature review are given in Table 4.17.

4.4.5 Partitioning

In SUCROS87 new dry matter is partitioned over the different plant organs. The actual partitioning functions is a function of phenological development. The definition of the partitioning functions F_i in SUCROS87 is the following:

$$\frac{\Delta W_i}{\Delta t} = F_i \frac{\Delta W}{\Delta t} \tag{4.27}$$

where ΔW_i is the organ weight increase and ΔW is the total weight increase to be partitioned.

Parameter review

The data gathered to derive the partitioning coefficients were generally time series of dry matter weight. The partitioning coefficient was calculated as the difference quotient of organ growth over shoot growth in the case of all aboveground organs and as the difference quotient of root growth over total dry matter growth for the roots. The calculation was executed under the constraint that both enumerator and denominator should be positive. Values larger than 1 (due to measurement errors in case of independent measurements of organ weight and total weight) were set to 1. The calculated difference quotient was associated with the average of the DVS-values.

Table 4.17 The uncertainty of the assimilate requirement ASR'x' and the maintenance coefficients MAIN'x' (at 25 °C, standardized with a Q_{10} of 2) used in SUCROS87.

	$\hat{\pi}$	RMSEP	number of data	unit
ASRRT	1.59	0.05	6	$kg(CH_2O)$
	(1.47)	(0.04)	(14)	$(kg\ (d.m.))^{-1}$
ASRSO	1.39	0.05	2	$kg(CH_2O)$
				$(kg\ (d.m.))^{-1}$
MAINLV	0.02	0.01	5	$kg(CH_2O)$
	(0.03)	(0.02)	(11)	$(kg\ (d.m.)d)^{-1}$
MAINRT	0.11	0.08	4	$kg(CH_2O)$
				$(kg\ (d.m.)d)^{-1}$
Q10	2.3	1.1	4	-

The most important problem encountered in the literature review is the definition of the organ. Part of the variation in leaves and stem partioning functions is explained by differences in definition. Some researchers define leaves = leaf laminae + leaf sheaths; other researchers use the definition leaves = leaf laminae. This explains some of the variation in early development in both leaves and stem. Correction for these effects is possible if the definition of the 'organ' is given and results regarding the constituting parts are known, which is rarely the case. In those

cases where no definition was given, the data were excluded from analysis.

A second problem encountered in the analysis is the determination of root weight. In several sources the authors warn that errors are present in this determination. Root weight and therefore root partitioning is generally underestimated. Additional variation arises from the estimation of DVS values from phenological observations and time of observation.

The statistical description of parameter uncertainty requires that the partitioning functions are estimated from a single instance. No instance was retrieved from literature, which allowed to fit all shoot partitioning functions to the data simultaneously. Restricting the argument values to a specific DVS-range allowed to select instances which could be used to fit leaf and stem partitioning simultaneously. A single parameter (BLV) was estimated, transforming the response value of the partitioning function to leaves (Appendix 7). In the shoot-root partitioning function the same transformation was used and the associated parameter (BSH) estimated. Results are presented in Table 4.18.

Table 4.18 The uncertainty of parameters modifying the leaf partitioning function (BLV) and the shoot-root partitioning function (BSH).

	$\hat{\pi}$	RMSEP	number of data	unit
BLV	1.54	0.20	26	-
	(1.48)	(0.13)	(51)	
BSH	1.15	0.46	50	-
	(1.14)	(0.48)	(100)	

4.4.6 Leaf area index

Leaf area growth is initially independent of photosynthesis and a function of temperature only. The equation used to describe the initial leaf growth is the following:

$$LAI = NPL \cdot LAINI \cdot e^{RGRL \cdot TSUMEM} \tag{4.28}$$

where

LAINI Initial leaf area (m^2 $plant^{-1}$)

RGRL Relative leaf area growth rate ($(°Cd)^{-1}$)

NPL Planting density ((plants) m^{-2})

TSUMEM Temperature sum after emergence; for default parameter values proportional to DVS (°C d)

The duration of this exponential growth phase is determined by the combined condition that DVS < DVSJUV and LAI < LAIJUV, where both LAIJUV and DVSJUV determine the end of the juvenile stage. After exceeding this condition the leaf area index growth is proportional to leaf weight. The proportionality factor the specific leaf area (SLA) is a tabulated function of

development stage. Its values are determined from measurements of the growth rates of leaf area (ΔLAI) and leaf weight (ΔWLVG) as follows:

$$SLA = \frac{\Delta LAI}{\Delta WLVG} \tag{4.29}$$

Parameter review

The relationship between LAI and DVS based on a time series of harvests allows to derive RGRL and LAINI characterizing leaf area growth in the first phase. Verheul (1992) presented data regarding the value of RGRL, which were transformed to a base temperature of 10°C. Additional RGRL values were derived from the fit of an exponential function to the LAI-DVS relationship used in deriving the tabulated function SLA. The initial leaf area index was converted to an initial leaf area per plant, using available planting densities.

The available data did not allow to fit DVSJUV and LAIJUV, given that only a limited number of measurements were available in the initial leaf growth phase. The scalars DVSJUV and LAIJUV are therefore assumed to be constant and the initial leaf area LAINI and RGRL should be determined using these default values. Due to the fact that leaf area index was only available as a function of DVS, the base temperature for juvenile leaf area development (TBSJUV) was fixed at 10°C.

The tabulated function SLA was parameterized through a transformation of the SLA-values. SLA has a positive value; the table was transformed using a multiplicative transformation ($SLA^* = BSLA*SLA$, Appendix 7). The calculation was identical to the calculation of the partitioning coefficients. The results are presented in Table 4.19 and Table 4.20.

Table 4.19 The uncertainty of the parameter BSLA modifying the specific leaf area function and those determining juvenile leaf area growth and their uncertainty.

	$\hat{\pi}$	$\hat{RMSEP}$	number of data	unit
BSLA	1.24	0.54	18	-
	(1.24)	(0.54)	(19)	
RGRL	2.8E–2	0.1E–2	7	$(°C\ d)^{-1}$
	(2.8E–2)	(0.9E–2)	(17)	
LAINI	5.3E–4	3.9E–4	11	$m^2\ plant^{-1}$

Table 4.20 The correlation between parameters deterring the juvenile leaf area growth.

	RGRL	LAINI
RGRL		
LAINI	–0.68	

4.4.7 Senescence

The literature on ageing and senescence is not easily expressed or translated into the parameters used in SUCROS87. The parameters used to describe leaf senescence in SUCROS87 are TBSSEN, SHRDRL, SHRDRH, DVSSEN, DVSRDR, TBSKIL and DVSKIL. A base rate of senescence was assumed to be constant at a value 0.001. The concepts used to describe senescence in SUCROS87 for maize are based on the approach used in CERES-maize (Jones et al., 1986); the nominal parameter values are partly based on calibration (Rabbinge et al., 1989). The description of the process used in CERES-maize does not contain references to other literature. Literature is therefore not accessible.

Of the parameters used the chilling temperature TBSKIL is the only one accessible to experimental research. Chilling research is often executed on very young seedlings, and does not seem relevant for a mature crop. Other parameter values could be estimated from measured leaf senescence in terms of leaf weight loss, or leaf area decrease. Daily temperatures during the experiment which would allow estimation of the parameters, are generally not published. Therefore the analysis of the uncertainty in the senescence parameters was not possible. The estimation of parameters in the light interception function used in LINTUL (Section 4.3.1) allowed the estimation of the start of leaf senescence, the parameter DVSSEN. To simplify the analysis only point estimates were derived (cf. Table 4.21).

Table 4.21 The uncertainty of the parameter determining the start of leaf senescence (DVSSEN).

	$\hat{\pi}$	RMSEP	number of data	unit
DVSSEN	1.134	0.110	10	-

4.5 A review of the uncertainty of parameters used in the model LINTUL

In this section, the parameter values and their natural variation for the model LINTUL are reviewed on the basis of data in literature and summarized in terms of probability distributions. The definitions required to execute the analysis are given in Section 4.1. The motivation for this approach was described in Section 4.2.

The uncertainty of some of the parameters used both in LINTUL and SUCROS87 was discussed in the previous section, notably partitioning and parameters determining development stage. Apart from these parameters LINTUL contains five parameters determining the light interception function FINT and a sixth parameter, the light use efficiency LUE.

4.5.1 The light interception

Under optimal conditions the daily growth rate of a crop can be calculated as the amount of

intercepted radiation multiplied by a conversion coefficient. The latter is fairly constant over the growing season. This concept was critically discussed by Demetriades-Shah et al. (1992). The function FINT describes the fraction photosynthetically active radiation intercepted by active leaves as a function of development stage. Multiplication of radiation by FRPAR yields the amount of PAR; multiplication of PAR by FINT yields the amount of intercepted PAR.

Parameter review

The function describing the fraction intercepted PAR as a function of development stage (FINT) is generally defined as the fraction intercepted by active leaves. To define 'active' leaves, some researchers (e.g. Muchow, 1994) remove senescent material; other researchers (Gallo and Daughtry, 1986) define the active leaf surface in terms of remote sensing indices. Spitters (1990) proposes to use the fraction soil covered by the projection of the green leaves as an estimate of the fraction of light intercepted. Another option is to compute FINT from LAI using data extracted for determination of the specific leaf area. This fixes the base temperature for DVS to be used in the model at 10°C.

Assuming extinction of light intercepted within the canopy with an exponential extinction coefficient of 0.65, the leaf area index was converted into a fraction intercepted PAR. The difference between intercepted and absorbed radiation was assumed to be negligible. This will introduce an error if leaf area index included the area of senescing and dead leaves. This was often not indicated in the data. The equation fitted to the data is the following:

$$F = \min\left[\frac{\beta\alpha}{\alpha+(1-\alpha)e^{-\mathrm{RI}\cdot\mathrm{DVS}}},\ 0.5-\frac{\mathrm{DVS}-\mathrm{DVHALF}}{\mathrm{DVLGTH}}\right] \tag{4.30}$$

where α, β, DVHALF, DVLGTH, and RI are coefficients, and DVS is the development stage.

The equation fitted is a slight modification of the original function suggested by Spitters (1986) to allow for datasets with maximal light interception values smaller than 1, introducing an additional parameter β. The equation was fitted to the data using a GENSTAT program. The parameter α was rewritten as a constant FINTI times the planting density NPL in each instance after estimation. FINTI is used in LINTUL.

In a few cases (4 out of 14) the leaf area index did not decrease (DVLGTH and DVHALF could not be estimated), although the final DVS was larger than or equal to 1.5. In these cases senescence was probably not observed and leaf weight refers to the sum of living and dead leaves. The results of the analysis (parameter values and their correlations) are presented in Tables 4.22 and 4.23.

4.5.2 The light use efficiency

The light use efficiency is the proportionality constant which relates crop growth rates to the amount of PAR intercepted (g d.m.. MJ^{-1} intercepted PAR).

Table 4.22 The uncertainty of the parameters determining the course of light interception over development stage.

	$\hat{\pi}$	RMSEP	number of data	unit
FINTI*	80.E−6	42.E−6	20	m^2 $plant^{-1}$
β	0.9	0.2	20	-
RI	11.	3.	20	-
DVHALF	1.9	0.4	10	-
DVLGTH	2.3	1.9	10	-

* FINTI = (α/NPL)

Table 4.23 Correlations between parameters determining the course of light interception over development stage.

	FINTI	β	RI	DVHALF	DVLGTH
FINTI					
β					
RI	−0.72				
DVHALF		0.84			
DVLGTH		0.70		0.93	

Parameter review

The light use efficiency is estimated from the plot of total aboveground dry matter weight (TADRW) against cumulated values of intercepted PAR from emergence to maturity. Different estimation procedures are used in literature. LUE is sometimes estimated as TADRW at finalharvest divided by the cumulative PAR intercepted over the growing season (e.g. Otegui et al., 1995). Alternatively LUE is estimated as the linear regression slope of the plot TADRW vs. cumulative intercepted PAR.

Most of the data presented in literature present values of LUE based on above-ground dry matter weight. Exceptions are results in the articles by Varlet-Grancher and Bonhomme (1982) and Williams et al. (1965). If LUE-values were determined over two or three development periods, the average value was used and the standard error of the mean was taken as measurement error. The results are presented in Table 4.24, together with the results derived from simulation using SUCROS87. The uncertainty in LUE as derived from simulations (cf. Section 4.2.4) has a mean value close to that found in literature, but is much more certain than LUE derived from literature. Given that the variation is derived for Dutch conditions, the two results are not comparable.

Table 4.24 The uncertainty of the light use efficiency (LUE). The estimate LUE^* is based on data simulated by SUCROS87.

	$\hat{\pi}$	RMSEP	number of data	unit
LUE	3.0	0.4	30	g (d.m.) MJ^{-1}
	(2.7)	(0.4)	(134)	
LUE^*	2.91	0.08	13	g (d.m.) MJ^{-1}

Chapter 5

Uncertainty analysis

5.1 Introduction

5.1.1 Determining research priorities

In the previous chapter, parameter uncertainty in LINTUL and SUCROS87 was described. Some fundamental choices were made: only parameter uncertainty was considered as source of prediction error, and was described in terms of probability distributions. In this chapter, the effect of parameter uncertainty on some output variables will be quantified. Once the output uncertainty is given, conclusions can be drawn.

The magnitude of the uncertainty in the model output caused by all parameters collectively allows to answer the question whether it is sufficiently small for the application considered. If it is too large, values of the parameters for the local production instance should be estimated. The output uncertainty may also be analysed for groups of parameters used in a submodel or for individual parameters. Such analyses should yield the contributions of uncertainty in the parameters to the output uncertainty. A result of this analysis would be a ranking of parameters or groups of parameters in terms of their contribution to the total uncertainty. This ranking suggests priorities in parameter estimation, but also in research aiming to reduce prediction uncertainty. Other uses of this ranking are possible, for instance a low ranking of a model component with a relatively large number of parameters suggests simplification. Such analyses are necessary as the ranking of the parameters can not be based directly on a ranking by coefficient of variation of the individual parameters because the relations between state variables and parameters in models are possibly non-linear or discontinuous.

After the uncertainty analysis, the first part of the analysis sketched in Chapter 1 (Figure 1.2) is completed. The model structure has been described and analysed; we have analysed properties of relations between state variables and parameters (Chapters 2 and 3). After the inventory of parameter uncertainty (Chapter 4) and after applying the methods described in this chapter it should be possible to indicate those parameters that contribute most to the prediction uncertainty in a specific situation.

There are three questions to be answered in this chapter:

- What is the uncertainty in the output given the uncertainty in the parameters and what consequences does this have?
- How can we estimate the contribution of different parameters and groups of parameters to output uncertainty?
- What is the contribution of parameters and groups of parameters to output uncertainty and what consequences does this have?

These questions are discussed in Sections 2, 3 and 4 of this chapter.

5.1.2 Sensitivity analysis and uncertainty analysis

The terminology regarding uncertainty and sensitivity analysis is not standardized and warrants some extra attention. In this thesis both the notion sensitivity analysis and uncertainty analysis are used for analyses of changes in model input (input signals (as e.g. meteorological data),

initial conditions and parameters) on model output. In sensitivity analysis these changes need not be realistic nor do they have to be derived from empirical data. The analysis of discontinuity as presented in Chapter 3 may be regarded as a sensitivity analysis. In uncertainty analysis, these changes are derived from empirical data using methods presented in Chapter 4.

5.1.3 Uncertainty analysis: screening and sampling

In uncertainty analysis, one would like to assess the effects of different parameters on the probability distribution of the output variable. In case of a simulation model, evaluating uncertainty in the output is most practically executed using simulation. The simulation study can be based on a number of model runs using different parameter values sampled from the multivariate distribution of the parameters. This is known as a Monte Carlo approach, which results in an approximation of the output distribution. A comparison of different methods is presented e.g. by Downing et al. (1985) and Iman and Helton (1988).

The accuracy with which the effect of different parameters on output can be estimated increases with the number of runs in a Monte Carlo approach. The number of runs necessary to analyse the effect of individual parameters with equal precision furthermore increases with the number of parameters. The number of parameters, the associated sample size and consequently the total time required may in some cases prohibit a Monte Carlo approach. This aspect of uncertainty analysis is discussed by Downing et al. (1985). They focus on the efficient approximation of model output, using two criteria: time required for a model run and number of parameters. For parameter-rich models they propose a step prior to uncertainty analysis: screening. The result of screening is a reduction of the number of parameters that is taken into account in the uncertainty analysis. According to Downing et al. (1985), screening procedures can be based on expert knowledge, or on a sensitivity analysis, e.g. with two levels for each parameter. Options for screening are also discussed by Kleijnen (1987), Morris (1991), and Welch et al. (1992). An example of screening in water management is given by Bresler and Dagan (1988). The possible approaches for different types of models are presented in Table 5.1, where the qualifications low and high are used in a qualitative sense. Any attempt at quantification depends on the allowable time, the computer used, the model used and the accuracy with which the parameter contributions have to be estimated.

Table 5.1 Approaches to uncertainty analysis for different types of models characterized by the number of parameters and the time needed for a simulation run.

	number parameters low	number parameters high
cost simulation run low	no special methodology required	select number of parameters through screening
cost simulation run high	use efficient sampling schemes	use efficient sampling schemes for limited number of parameters selected through screening.

Although the models SUCROS87 and LINTUL have numbers of parameters for which screening is relevant, runtimes are sufficiently small to allow large total numbers of simulations within manageable periods (in the order of day in total). Screening is therefore not necessary. Research in this thesis is restricted to the uncertainty analysis proper, concentrating on the questions: how does one sample the probability distribution and how does one evaluate the results?

Sampling from the uncertainty distribution

To determine the uncertainty in the output as a function of the uncertainty in the parameters, the parameter distribution should be sampled. Two methods are often considered: simple random sampling and latin hypercube sampling. The most straightforward sampling method is simple random sampling, which consists of consecutive independent draws from the multivariate parameter distribution.

In latin hypercube sampling, all individual parameter distributions are divided into M equiprobable intervals (e.g. McKay et al., 1979; Stein, 1987; Owen, 1992; but also Press et al., 1992). The trick in latin hypercube sampling is to generate only one parameter value in each interval. If the parameters are independent, the M parameter values for the first parameter are randomly paired with the M values for the second parameter and so on. The sampling procedure ensures that each parameter has been sampled evenly over its entire range.

Latin hypercube sampling has to be modified to yield a sample in which the correlations between parameter estimates as presented in the previous chapter are reproduced. For arbitrary multivariate distributions an approximate procedure to introduce rank correlations is available (Iman and Conover, 1982). This procedure can also be used to introduce rank correlations in simple random samples independently drawn for individual parameters.

The advantages and disadvantages of different sampling methods are compared by e.g. McKay et al. (1979) in terms of the accuracy with which the mean of an output distribution is estimated. Accuracy can be estimated from an analysis of the results of repeated samples. Whereas latin hypercube sampling was shown to have good efficiency (a higher accuracy of the estimate of the mean for a smaller sample size), random sampling has some theoretical advantages (Owen, 1992). As moreover simple random sampling is easier to execute and given that the simulation model runs take little computation time, we will use simple random sampling in combination with the method of Iman and Conover (1982) to generate correlated samples.

5.2 Results of uncertainty analysis

5.2.1 Method

The uncertainty analysis was executed using a simple random sample of 1000 parameter vectors for each model. The sample was used to establish the uncertainty in final yield for 13 years of experiments around Wageningen using the relevant meteorological data, observed planting and final harvesting days. The analysis was executed twice to allow an estimate of the variation due to sampling. For SUCROS87 the second sample was generated by randomly shuffling the first

sample to impose the desired correlation structure; for LINTUL two random samples were generated and the desired correlation structure imposed on each sample (cf. Section 5.1.3). The samples were generated with GENSTAT (Payne and Lane, 1993) using the multivariate parameter distribution (Chapter 4). The effect of parameter uncertainty on output variables was analysed for total aboveground dry matter (TADRW) and grain weight (WSO) at final harvest for SUCROS87 and LINTUL.

5.2.2 Results

The output distributions are presented in terms of their minimum and maximum values and their quartiles (25, 50, and 75%) for the simulated variables in Figures 5.1a and b (SUCROS87) and Figures 5.2a and b (LINTUL) for one sample. These figures allow to assess the output uncertainty and its variation over years. Summary characteristics of the distribution of TADRW and WSO are given in Table 5.2. The characteristics are averaged over thirteen years. The coefficient of variation (c.v., %) is calculated using the averages of standard deviation and production over years. The sampling variation turned out to be sufficiently small not to affect the conclusions.

Table 5.2 Characterization of the average distribution of total aboveground dry matter weight (TADRW) and grain weight (WSO) (both in $t.ha^{-1}$, average over 13 years, results for 2 samples) for SUCROS87 and LINTUL using the parameter distribution established in Chapter 4.

		SUCROS87		LINTUL	
TADRW	mean	6.80	6.76	5.04	4.84
	median	6.33	6.23	4.73	4.49
	lower quartile (25%)	3.05	3.13	1.67	1.55
	upper quartile (75%)	9.81	9.69	7.88	7.55
	c.v. (%)	68	68	75	75
WSO	mean	1.00	0.95	1.79	1.76
	median	0.20	0.15	0.59	0.63
	lower quartile (25%)	0.00	0.00	0.00	0.00
	upper quartile (75%)	1.28	1.34	3.56	3.57
	c.v. (%)	171	166	121	122

5.2.3 Discussion

The results show that the average prediction uncertainty is very large compared to the average value of the variable of interest (a c.v. of about 70% for TADRW; about 150% for WSO). The conclusion to be drawn from this simulation experiment is that the global variation of maize parameters (cf. Chapter 4) is much too large to use the resulting parameter distribution for predictions under specific meteorological conditions at a specific location. Accordingly, the use

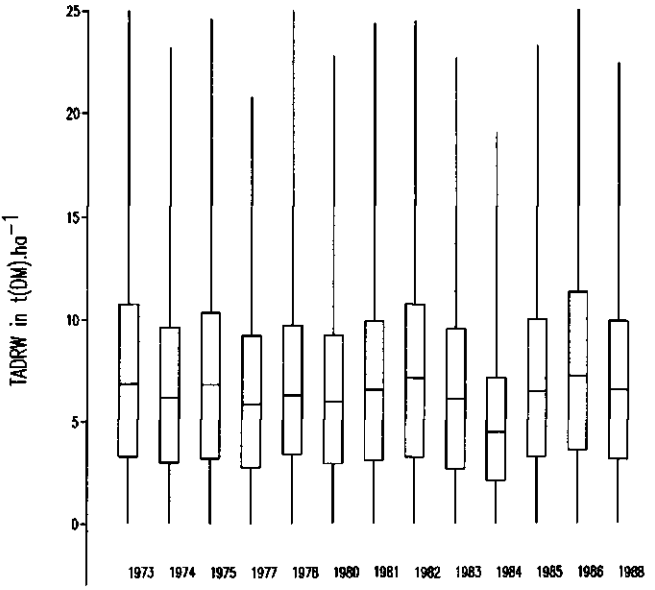

Figure 5.1a The uncertainty (minimum, lower quartile, median, upper quartile, and maximum) in the total aboveground dry matter weight (TADRW) in different years for SUCROS87 given the parameter uncertainty.

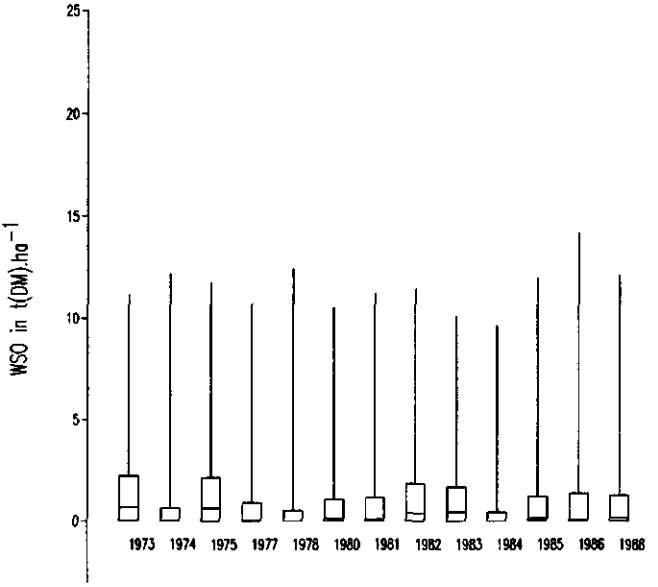

Figure 5.1b The uncertainty (minimum, lower quartile, median, upper quartile, and maximum) in the dry matter grain weight (WSO) in different years for SUCROS87 given the parameter uncertainty.

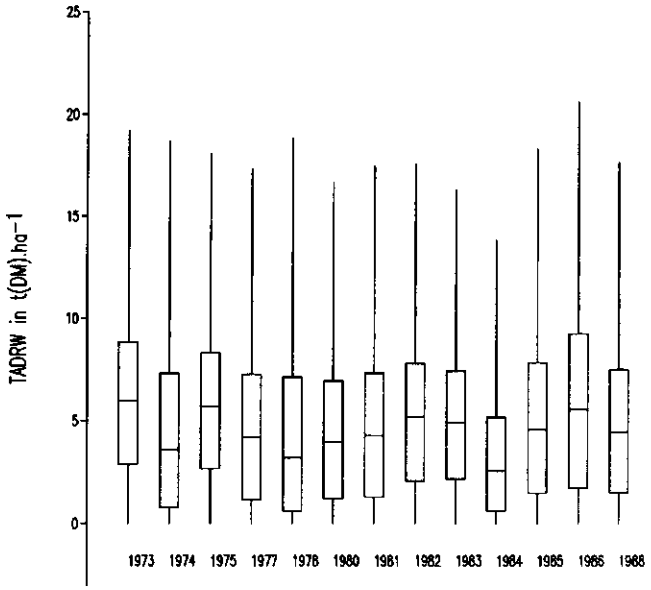

Figure 5.2a The uncertainty (minimum, lower quartile, median, upper quartile, and maximum) in total aboveground dry matter weight (TADRW) in different years for LINTUL given the parameter uncertainty.

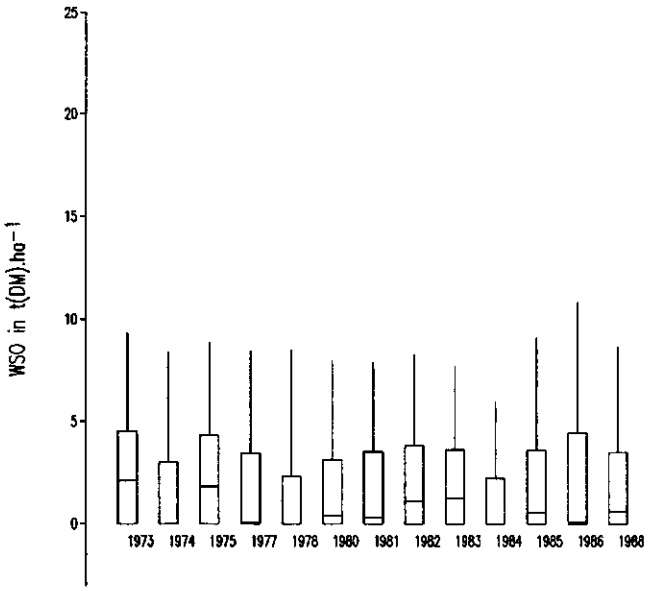

Figure 5.2b The uncertainty (minimum, lower quartile, median, upper quartile, and maximum) in the dry matter weight of grain (WSO) in different years for LINTUL given the parameter uncertainty.

of the models LINTUL and SUCROS87 is only justifiable if the parameter distributions are adapted to or derived under local conditions, and if through that process output uncertainty is reduced. The following conclusions can additionally be drawn:

- The distributions for TADRW are different for both models. The effect of parameter uncertainty on TADRW is of the same relative magnitude for both models and yields a c.v. of about 70%.
- The distribution of grain yield (WSO) differs between the models. The relative effect of parameter uncertainty on grain yield differs for both models. Grain yield in SUCROS87 is more uncertain than in LINTUL.
- For both models the mean upper quartile of the dry matter (potential) yield (TADRW) is lower than the 13 ± 2 t.ha^{-1} calculated as the average (actual) yield of the four best cultivars in variety trials over the period 1954-1981 (te Velde, 1984). This is a further reason to adapt the models to local conditions.

As parameter distributions were reviewed at least partly independent for the models, differences in yield distributions are to be expected. At this point, one can only speculate as to the causes of the differences between model simulations of potential yield and reported yield. The explanation is the non-specificity of the multivariate parameter distribution for Dutch conditions. However, attributing the difference to specific parameters or to specific aspects, such as correlations, of the multivariate distribution is not yet possible.

Uncertainty in meteorological data

A question remains: are other sources of uncertainty perhaps more important? One possible source of output uncertainty is the meteorological data. Given the meteorological data for a specific location, should we only adapt the parameters to the local conditions or should we - additionally, or initially - try to reduce the uncertainty of the meteorological data for that location?

In simulation studies one uses the meteorological data from a measurement site closest to the production instance of interest. Obviously this introduces an additional source of variation, that we will call uncertainty of the meteorological data. An inventory of the uncertainty in local meteorological data based on a varying distance of the measuring station to the experimental field was not executed in Chapter 3. Therefore, we can not compare the contributions of uncertainty in measured meteorological data and those of parameters.

Model behaviour suggests that for the uncertainty in meteorological data to be more important than parameter uncertainty, variation in temperature and radiation between experimental field and meteorological station should be large and systematic. To achieve the very small crop yields simulated, the required systematic differences in temperature should be huge. In that case temperatures should either be very high, to make maintenance respiration the dominant process, or very low, to reduce photosynthesis to a minimum. Systematic differences should also be large for radiation, since crop yield is proportional to cumulative intercepted radiation.

The study of Nonhebel (1993) in which climatic averages, monthly averages and decade averages and data from different meteorological stations were used as input on a daily basis, can

be interpreted as a study of the effects of variation between measurement instances (uncertainty) in weather data. She found that uncertainty in temperature and radiation lead to relative errors in the order of 10% in yield of springwheat simulated using SUCROS. This is much smaller than the relative output uncertainty established in the above simulation experiments (Table 5.2).

Adaptation of the two models through estimation of local parameter values, thus reducing parameter uncertainty, is, therefore, in this case study given priority over efforts to improve the quality of local meteorological data.

5.3 Estimating parameter contribution to prediction uncertainty

In the previous section, the output uncertainty in both models was shown to be large. Furthermore, it was argued that parameter uncertainty is a more important contributor to output uncertainty than the uncertainty in meteorological data. Decreasing parameter uncertainty is an important aspect in reducing prediction uncertainty.

To limit the number of parameters to be considered for local measurement or calibration, we would like to establish the relative importance of each parameter in explaining the output uncertainty. The ranking of the parameters on the basis of this criterion may serve as a guidance to set research priorities (e.g. the choice of parameters to be adapted to local conditions to improve quality of predictions).

An early example of parameter ranking based on uncertainty analysis is presented by Miller et al. (1976). Other examples are Reed et al. (1984) for a model of outdoor recreation; Blower et al. (1991) for HIV epidemiology; Saltelli et al. (1993) for nuclear waste disposal; Rossing et al. (1994b) for pest control in winter wheat. Aggarwal (1995) and Dagan and Bresler (1988) present analyses for parameters in models of water-limited crop yield.

The research question is how to establish the ranking of parameters. To answer this question a quantification of the uncertainty contribution relative to that of the other parameters suffices. The question can be rephrased as follows: What is the effect of an improvement in our knowledge regarding the parameter values on the output uncertainty? If this effect can be established for all parameters, we are able to assign a research priority to each parameter.

5.3.1 The mathematical frame

To establish the effect of a reduction of the uncertainty of parameters we have to analyse the relation between the uncertainty in the output, an univariate response y - or multiple univariate responses y_i - and the parameters, a multivariate argument. The parameters have a known distribution. Some parameter estimates are correlated. The relation between output y and the parameters is complicated and deterministic (in our case defined by the models SUCROS87 or LINTUL in a specific location and year).

Simulation, using a random sample from the joint parameter distribution yields the distribution of the response y, as presented in Section 5.2.2. We are interested in reducing the uncertainty in y. There are two possibilities to achieve a reduction in the uncertainty of y: either

the parameter distribution for the population becomes better known; or the parameter becomes known for a specific instance. In the latter case the parameter is no longer uncertain (in the terminology of Chapter 4): its value for the production instance is known. For our study the latter case is the most interesting.

In that case we have to investigate what happens to the variation in y if the uncertainty in different parameters is reduced to zero. Furthermore we want to know how this effect differs between individual parameters. If we characterize the uncertainty in terms of variances of the distributions, we are interested in characterizing the variance of the output if parameters become perfectly known (zero variance) for a specific production instance. Whereas in principle any number and any combination of parameters could simultaneously become perfectly known for a specific production instance, a systematic analysis is not feasible as the number of possible combinations increases very fast with the number of parameters.

Two problems have to be solved: 1) which ranking criterion to use and 2) how to keep the investigation manageable.

Given that we have expressed parameter uncertainty in terms of different distributions, a dimensionless ratio calculated as:

$$\frac{\textit{Change in output variance given that the parameter value(s) become(s) known}}{\textit{Output variance given that all parameters are uncertain}}$$

could be used as a ranking criterion. To circumvent the problem that the parameter values for a new instance are unknown, criteria should be formulated in terms of expected variances.

To keep the investigation manageable, research should be limited to a small number of options of learning about local values of the parameters. Apart from the reference scenario in which all parameters are uncertain, and a trivial scenario in which all parameter values for the production instance are known, two other cases are interesting because of their simplicity. In the first case the uncertainty of a single parameter is set to zero and all other parameters are uncertain. Executing this analysis for all parameters and ranking the results yields the answer to the question "What if the value of this parameter were the first to become known for a specific production instance?" In the second case all parameters except one are known for a specific production instance. The analysis would supply answers to the question: "What if the value of this parameter were the last to become known for a specific production instance?" The answer to this question allows us to establish the parameters we can 'safely' neglect. Based on these two options we will use two criteria to quantify the uncertainty contribution of a parameter, the top marginal uncertainty contribution and the bottom marginal uncertainty contribution.

The relative change in output variance, given the expected variance if a parameter (or a set of parameters) is the first to become perfectly known, is called the *top marginal uncertainty contribution*; if the parameter (or a set of parameters) is the last to become perfectly known the relative change in variance given the expected variance is called the *bottom marginal uncertainty contribution* (Jansen et al., 1994). We will use the symbol '$\top$' for the top marginal uncertainty contribution and '$\perp$' for the bottom marginal uncertainty contribution. A formal definition can be given in terms of conditional distributions (e.g. Janssen, 1994). In these definitions F is any simulation model, x is the vector of parameters, s is a subset of these parameters, and $-s$ the

complementary set of parameters. The top marginal uncertainty contribution of subset s is then defined as:

$$\top(s) \equiv \frac{Var[F(x)] - E[Var[F(x)|s]]}{Var[F(x)]} \tag{5.1}$$

which can be written more concisely as:

$$\top(s) = \frac{Var[E[F(x)|s]]}{Var[F(x)]} \tag{5.2}$$

using the variance decomposition rule for conditional distributions (e.g. Rao, 1973).
The bottom marginal uncertainty contribution of subset s is defined as:

$$\bot(s) \equiv \frac{E[Var[F(x)|-s]]}{Var[F(x)]} \tag{5.3}$$

If the model $F(x)$ is additive in the components of its argument x, and if the components of x are stochastically independent, the top marginal variance is equal to the bottom marginal variance (Jansen et al., 1994). If the estimates of $\top$ and $\bot$ differ more than may be ascribed to sampling error, one may conclude that there are dependencies, or interactions. As generally interactions and dependencies will occur, both $\top$ and $\bot$ have to be estimated.

Janssen (1994) refers to the top marginal contribution as the relative expected reduction in variance (RERV); whereas the absolute reduction in variance (or its square root) is also known as Importance Measure (Hora and Iman (1989), also Saltelli et al., 1995; Hora and Iman (1990), also Janssen, 1994). Other possible measures to quantify uncertainty contributions are reviewed by Janssen (1994), Saltelli et al. (1993), and Hamby (1994).

5.3.2 Straightforward simulation approach

Given these formal definitions the practical problem is how to estimate the different uncertainty contributions. A very straightforward 'what-if' approach to this problem is simulation: design a simulation experiment in which different parameters or groups of parameters are becoming known for a specific production instance. We will use this approach to illustrate the concepts of top and bottom marginal uncertainty contribution. The problem how to deal with correlated parameter values will be skipped in this section. For the practical execution a different approach (described in Section 5.3.3) will be used.

Assume that we would like to know the relative change in output uncertainty for a single parameter x_i, whereas the other, say k–1, parameters remain uncertain. To execute the simulation experiment, we need to specify the new value of the parameter in the new production instance. This introduces a problem as this value is as yet unknown, and as the output uncertainty may depend on its value. In order to find a general measure of the importance of a parameter, the

simulation study has to be repeated for a number of values, say L, of the parameter x_i. This can be achieved by using a random sample of parameter values $x_{i1}...x_{iL}$. Each of these L parameter values has to be combined with an N-sized sample of the other parameters.

Averaging the output variance over the L N-sized subsamples yields an estimate of the *expected* variance for a production instance in which the parameter is the first to become perfectly known. This can be used to estimate the top marginal contribution of the parameter as:

$$\hat{\tau}(x_i) = \frac{Var[F(x)] - \frac{1}{L}\sum_{j=1}^{L} Var[F(x)|x_{ij}]}{Var[F(x)]} \tag{5.4}$$

given that the output variance for a reference sample in which all parameters are uncertain is available. This approach would require $(L \times k + 1) \times N$ runs of the model.

A simulation study to calculate the bottom marginal uncertainty contribution would require to execute runs in which all but a single parameter are perfectly known, and a reference run in which all parameters are uncertain. However, as the exact value for the k–1 parameters for the new production instance are unknown, one could execute $L^{k-1} \times N$ runs for each of the k parameters, in analogy to the estimation of the top marginal contribution τ. As in the estimate of the top marginal contribution the output variance has to be averaged over the N-sized L^{k-1} subsamples to yield an estimation of the expected variance. This estimate of the bottom marginal contribution would require $(L^{k-1} \times k + 1) \times N$ runs, which raises questions as to its general feasibility.

5.3.3 Practical execution: efficient simulation approaches

A straightforward simulation approach is impractical as - even for the strongly simplified experiment as described in Section 5.3.2 - this would still require a large number of simulation runs. Notably the number of runs required for the bottom marginal contribution is high.

Fortunately, it is possible to further reduce the number of runs required. As will be explained below, the procedure that will be used is based on an approximation using a single sample in which all parameters are uncertain. An alternative to the increasingly complex analyses described in this section is a sampling method specifically designed for the problem of estimating variance contributions: winding stairs Monte Carlo (Jansen et al., 1994).

Approximation of the top marginal contribution

A hypothetical example of the straightforward simulation approach to estimate the top marginal uncertainty contribution is presented in Figure 5.3. It depicts a situation in which for each of the L values (in Figure 5.3 L = 10) of a single (uncorrelated) parameter for an unknown production instance an N-sized sample of the other parameter values is generated and the model output simulated. These results can be used to calculate an estimate of τ, as described in the previous section, but would have to be repeated for each parameter. A reduction of the number of runs can

be achieved by using an approximation which allows to skip the repetition for each parameter. This is possible by using a single sample in which all parameters are varied simultaneously. The result of a sample in which all parameters are varied simultaneously is shown in Figure 5.4. These results can be analysed for each parameter as if L values of a single parameter for an unknown production instance had been in some way generated. In this approximation, parameter values close to one of these generated parameter values are assigned to it. In Figure 5.5, Figure 5.4 is combined with an analysis in which we have assigned the different elements of the sample to L imaginary values of the parameter (L is again 10). Figure 5.6 shows another way of looking at this procedure: the sample is partitioned over bins of the parameter, where bins are subranges within each parameter range. Each of these bins contains a parameter value assumed to be *the* value in the production instance, to which all parameter values within the bin are assigned. On the basis of these results we calculate an estimate of the top marginal contribution characterizing a parameter. For parameter x_i, it is calculated as:

$$\hat{\tau}(x_i) = \frac{1}{b}\sum_{j=1}^{b}\left(1 - \frac{N-1}{N/b-1}\frac{ssy_j(i)}{ssY}\right) = 1 - \frac{N-1}{N-b}\sum_{j=1}^{b}\frac{ssy_j(i)}{ssY} \tag{5.5}$$

where b is the number of bins over which x_i is partitioned, $ssy_j(i)/(N/b-1)$ is the variance of the sample in bin j, and $ssY/(N-1)$ the output variance in the reference case in which all parameters are uncertain. The larger both the number of bins and the number of values contained in a bin, the better the estimate of τ. For a fixed sample size a trade-off has to be made between number of bins and bin size (number of values contained in a bin).

Approximation of the bottom marginal contribution

In analogy to the estimation of the top marginal uncertainty contribution, we may also partition a single sample over boxes for the $k-1$ parameters (where boxes are multidimensional bins). Given a fixed sample size (N) and a fixed number of boxes per parameter (b), the average number of values in a box decreases as $n = N/b^{k-1}$. Obviously, given that the variance within a box is (on average) calculated over a smaller number of values (as $N/b^{k-1} < N/b$ for $k > 1$) this approximation of the bottom marginal contribution will always be less in quality than the top marginal contribution. This extension of the above averaging procedure for the top marginal contribution to multiple dimensions (in terms of parameters) is not easily visualized and introduces the problem how to choose the boxes.

Whichever way the boxes are chosen, the average of the variance within each of the boxes for $k-1$ perfectly known parameters (excluding x_i) is calculated as:

$$MSE_{k-1}(i) = \frac{1}{b}\sum_{j=1}^{b}\frac{ssy_j(i)}{N/b-1} = \frac{1}{N-b}\sum_{j=1}^{b}ssy_j(i) \tag{5.6}$$

where b is the number of boxes and $ssy_j(i)/(N/b-1)$ is the variance of the sample in the box, given that parameter x_i is unknown. Note, however, that the index j now refers to a $k-1$ dimensional box coordinate.

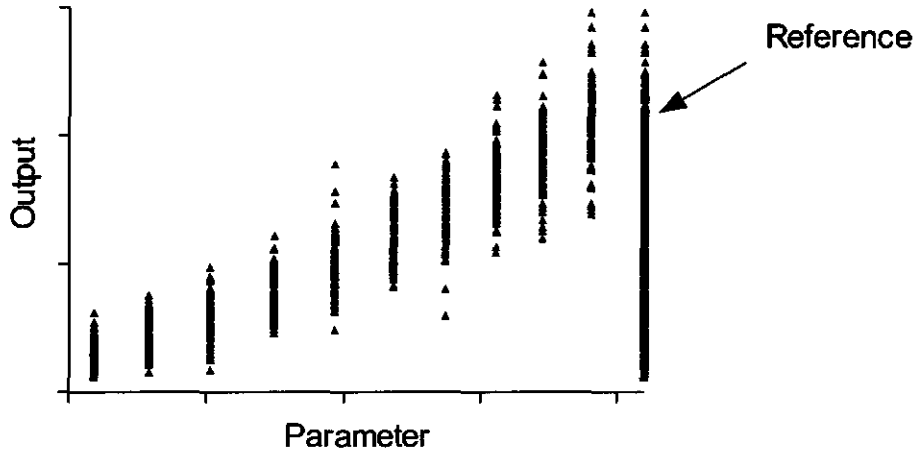

Figure 5.3 An example of a hypothetical simulation experiment in which L values of an individual parameter are chosen (in this Figure $L = 10$). Each value of the parameter is combined with a sample of the remaining uncertain parameters. The simulation is executed and the output is plotted against the values of the parameter. Together with the results of a reference run in which all parameters are uncertain, these results can be used to estimate the top marginal uncertainty contribution.

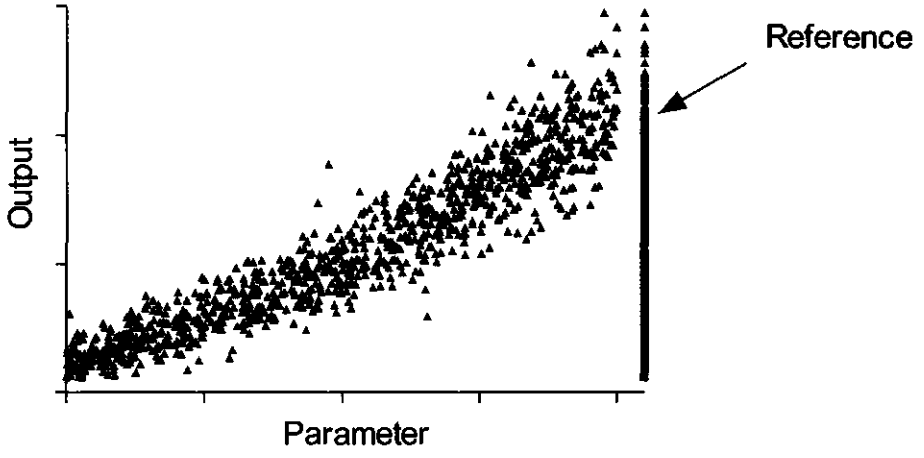

Figure 5.4 The outcome of a hypothetical simulation experiment in which all parameters are uncertain. A plot of the output against the values of a single parameter serves as the basis for an alternative estimate of the top marginal uncertainty contribution.

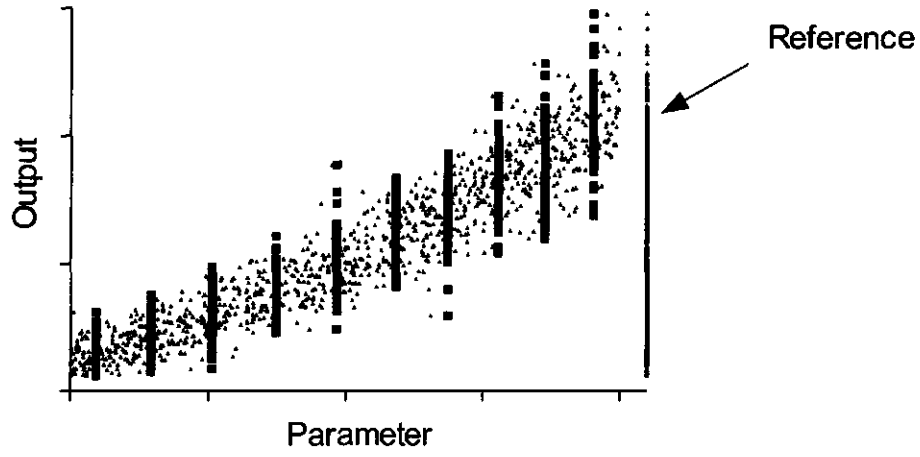

Figure 5.5 An analysis of Figure 5.4: the elements in the sample are assigned to the midpoint of a subrange of the parameter (a bin). These L midpoints (in this figure L=10) are interpreted as values of the parameter comparable to those in Figure 5.3 and can be used to estimate the top marginal uncertainty contribution.

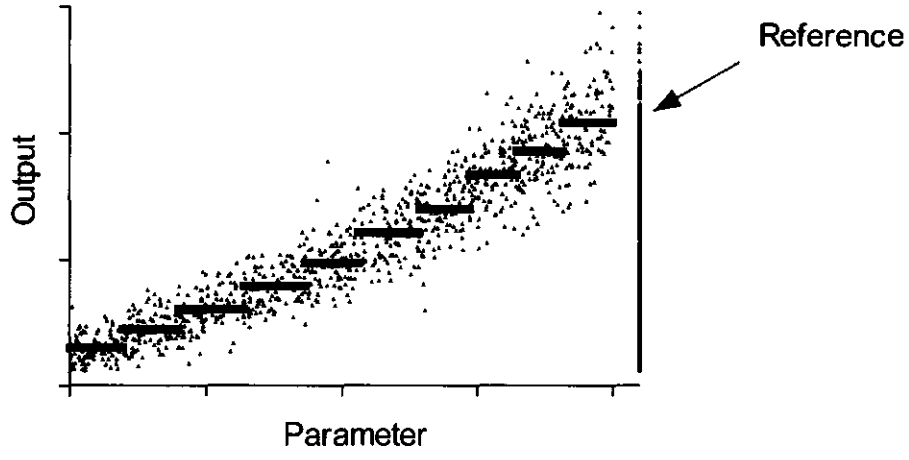

Figure 5.6 Another view of Figure 5.5: the range of the parameter is divided in L subranges (bins). The average output in the bin which is used in the estimate of the top marginal uncertainty contribution is presented. This figure suggests that other, smoother, functions may also be used to estimate the top marginal uncertainty contribution.

Given that calculation of $\perp$ is an approximation, the question is just how good it is. The quality can be estimated in terms of the above equation on the basis of the variance remaining when all parameters are perfectly known. If the approximation is perfect, the residual variance should be 0. The residual variance can be regarded as the 'blind spot' of the approximation: it reflects those effects we cannot 'see', e.g. for the above procedure effects smaller than box size. Its magnitude is a measure of the importance we should attach to the conclusions drawn from the analysis.

A residual variance not equal 0 also requires that we have to correct the calculation of the bottom marginal contribution. If MSE_k is the estimated output variance remaining if all parameters become perfectly known, $MSE_{k-1}(i)$ is the estimated variance remaining if k–1 parameters (except for parameter x_i) become perfectly known, both for a given partitioning of the sample over the boxes. The difference $MSE_{k-1}(i) - MSE_k$ is the estimated variance which is explained by the parameter to be the last to become better known. Dividing this difference by the full variance of the output sample (MSE_0) yields an estimate of $\perp$ for parameter x_i, calculated as:

$$\hat{\perp}(x_i) = \frac{MSE_{k-1}(i) - MSE_k}{MSE_0} \tag{5.7}$$

Estimating the top marginal contribution using regression

It is attractive to reformulate the above approach as a regression problem, because regression offers standard tools for analysis, and allows a more general view of the above approach. As we will show a regression approach is especially practical for the estimation of the bottom marginal contribution. In this regression approach one uses the model output as response and the parameter values as regressors.

In regression, the step function resulting from introducing bins with the associated average function values within each bin as the function value is known as a bin smoother (Hastie and Tibshirani, 1990). The use of a bin smoother can be (re-)interpreted as a regression problem in which a constant y value is fitted within each bin. Figure 5.3 suggests that one could also fit other, smoother, functions f with p parameters within each bin. The calculation of the top marginal contribution for parameter x_i could be generalized as:

$$\hat{\tau}(x_i) = \frac{1}{b}\sum_{j=1}^{b}\left(1 - \frac{N-1}{N/b-p}\frac{ss(y_j, f, i)}{ssY}\right) = 1 - \frac{N-1}{N-bp}\sum_{j=1}^{b}\frac{ss(y_j, f, i)}{ssY} \tag{5.8}$$

where b is the number of bins for x_i, $ss(y_j, f, i)/(N/b-p)$ is the residual mean square of the sample in bin j with respect to the regressed function, and $ssY/(N-1)$ the output variance if all parameters are uncertain. To estimate top marginal contributions any parametric or non-parametric function can be fitted, but one could also fit a linear function in a single bin ($b = 1, p = 2$). In the context of regression, the top marginal contribution τ is closely linked to quantities used in linear regression; τ equals the fraction of variance accounted for (R^2_{adj}, e.g. Montgomery and Peck, 1992). Given a small sample and a parameter-rich function f, this estimate of τ can be less than 0. For $b = 1$, and N approaching infinity, this measure becomes the squared linear correlation coefficient (e.g. Janssen, 1994).

As we have seen in Chapter 3, some of the relations between individual parameters and state variables are non-linear in SUCROS87. In this case relevant functions, other than the bin smoother, could be polynomials or splines, again - for simplicity's sake - assuming one bin. The most straightforward approach however, is to fit a linear function in a single bin. The mathematical simplicity of this option is attractive, if the quality of the approximation is sufficient i.e. if the percentage of variance accounted for by fitting all parameters is close to 100%.

Estimating the bottom marginal contribution using regression

In regression, the calculation of the bottom marginal contribution requires calculating the residual variance if all parameters are used as regressors (the full fit) and calculating the residual variance if all except one parameter are used as regressors. This is much more straightforward than the above approach in which the parameter sample is partitioned over boxes. In multiple linear regression $\perp$ becomes the relative partial sum of squares, another possible uncertainty measure (Dale et al., 1988; Janssen, 1994). As in the estimation of the top marginal contribution, the use of multiple linear regression to calculate $\perp$ is acceptable if it yields a good fit using all parameters as regressors.

Fitting a sum of bin smoothers, one for each parameter, allows to take non-linear relations into account in a linear regression context and yields a discontinuous and non-smooth description of the relation between output and parameters. Continuous and smooth non-linear relations can be fitted using additive regression models (Hastie and Tibshirani, 1990). Additive regression models are sums of non-parametric functions (such as splines) where a single function is fitted for each individual parameter.

5.3.4 Procedure and problems

To estimate top and bottom marginal parameter contributions to output uncertainty it is not necessary to start the analysis by immediately using the most complex approach. Both types of contributions can initially be estimated using linear regression. If the quality of the full linear fit is satisfactory, and top and bottom marginal contributions are more or less identical, the analysis can be concluded. However, one may also observe that the regression result of the full linear fit is insufficient, or that the ranking based on top and bottom marginal contributions yields ambiguous results, or that both fit and ranking are not satisfactory.

A general recipe for situations in which these problems occur is difficult to give. However, if the quality of the full linear fit is insufficient, non-linear functions (such as a bin smoother or splines) should be considered. If the regression result for the full additive fit does not improve, this indicates that parameters are interacting. Fitting flexible functions that allow for interaction was discussed by Morgan and Sonquist (1963) and has lead to different algorithms (e.g. Breiman et al., 1984; Friedman, 1991). These approaches require special software.

If the quality of the full additive fit is sufficient, but results of ranking based on top and bottom marginal contributions are strongly different, an alternative is to select that subset of parameters which jointly contribute most to prediction uncertainty as compared to subsets of the

same size. This leads to the experiment of fitting different combinations of parameters to optimize e.g. the percentage variance explained, while keeping the number of parameters constant. Executed in a more systematic fashion this is known as subset selection. The most straightforward procedure is to try all possible subsets of the same size. Because this is not always possible, other, more efficient methods have been developed. Reviews are given by Hocking (1976) and Miller (1990). Subset selection is also possible for additive regression models (Hastie and Tibshirani, 1990).

If the estimation of parameter contributions using functions allowing for interactions is not satisfactory, one should use sampling procedures which were developed to estimate uncertainty contributions for interacting parameters (e.g. Jansen et al., 1994), which generally require larger samples.

5.4 Uncertainty contributions in SUCROS87 and LINTUL

Based on the procedure sketched in the previous section, the results of the uncertainty analysis for SUCROS87 and LINTUL were analysed to estimate the contributions of individual parameters to output uncertainty.

5.4.1 Preliminary analysis

Using linear regression, the percentages variance explained fitting all parameters are given in Table 5.3. These results show that linear regression explains little of the total variation in both TADRW and WSO.

Table 5.3 Percentage of variance explained fitting a linear regression model of total aboveground dry matter weight (TADRW) and grain weight (WSO), with all parameters as regressors, for SUCROS87 and LINTUL. First number is average over 13 years, second number is standard error over years.

	TADRW	WSO
SUCROS87	59.8 ± 1.8	33.9 ± 2.5
LINTUL	43.3 ± 5.5	50.7 ± 2.7

As a next option we considered using the bin smoother described in Section 5.3.3. In this regression problem the number of levels (bins) is a free parameter, with a maximum value limited by sample size. To select this parameter the full regression model was fitted to the different outputs using different number of bins. Based on the results presented in Table 5.4a and b, the number of bins finally chosen was 15. The results in Table 5.5 were calculated over the two repetitions of the uncertainty analysis, each of a sample size of 1000, using 15 bins per parameter to fit the full regression model.

Table 5.4a Percentage variance explained (average over 13 years, single sample) for simultaneous fits of 27 parameters with x bins using 2 different responses of SUCROS87 (TADRW, WSO); for comparison the quality of the linear fit (lin) is given.

x	2	5	10	15	20	lin
TADRW	43.3	68.5	74.9	79.2	76.8	59.8
WSO	36.7	51.0	55.5	63.2	56.7	33.9

Table 5.4b Percentage variance explained (average over 13 years, single sample) for simultaneous fits of 8 parameters with x bins using 2 different responses of LINTUL (TADRW, WSO); for comparison the quality of the linear fit (lin) is given.

x	2	5	10	15	20	40	60	lin
TADRW	48.0	75.5	83.7	85.5	86.9	88.1	88.8	43.3
WSO	59.0	77.0	82.2	83.7	84.3	85.4	85.4	50.7

Table 5.5 Percentage variance explained for the full regression model using 15 bins per parameter. First number is average over 13 years and 2 samples, second number is average standard error over samples.

	TADRW	WSO
SUCROS87	78.0 ± 0.7	60.1 ± 4.4
LINTUL	85.0 ± 0.7	84.3 ± 0.7

These results allow to conclude that the use of a bin smoother allows a substantial improvement in percentage variance explained. This is due to the non-linear responses to the parameters in both simulation models.

5.4.2 The uncertainty contributions for the evaluated parameters in SUCROS87

The analysis of parameter uncertainty contributions is based on the two sample(s) described earlier. The output variable analysed is TADRW at final harvest. The top and bottom marginal contributions based on a bin smoother (15 bins) are presented in Table 5.6. They are presented as a summary over the 13 growing seasons, characterized by the overall average and the average sampling standard error. The parameters are ranked in order of decreasing top marginal contribution. The sampling error allows some impression as to the importance of the differences in ranking between the different parameters.

Comparison of $\top$ and $\perp$ allow to assess the effects of the correlations between parameters.

Differences between $\top$ and $\perp$ are only important for the temperature sum from silking to maturity (TSDVRR), due to its correlation with the temperature sum from emergence to silking (TSDVRV). It is interesting that TSDVRR has a smaller relative effect on $\perp$ than TSDVRV, probably due to the fact that it is effective over a shorter period of the growing season.

Negative estimates of $\top$ and $\perp$ may be regarded as an indication of the order of magnitude of sampling order. They are a result of the limited sample size N and the degrees of freedom required by the fitting function, and should be interpreted to mean that none of the output variance was explained by the model-parameter combination. The negative values can be removed if larger sample sizes, or a smaller number of bins are used.

Table 5.6 A summary of the top and bottom marginal uncertainty contribution (estimate ± standard error over samples in %) to the total aboveground dry matter weight (TADRW) for parameters in SUCROS87. For an explanation of the abreviations cf. Appendix 1.

	$\top$			$\perp$		
EFF	22.0	±	0.7	22.8	±	1.5
TSDVRV	17.0	±	1.6	11.9	±	1.0
BSH	11.9	±	0.6	11.6	±	1.6
TSDVRR	6.9	±	1.4	0.3	±	0.1
MAINRT	6.4	±	0.4	7.4	±	0.5
Q10	5.4	±	1.2	5.7	±	0.8
AMX	4.4	±	0.8	4.9	±	0.8
BSLA	4.2	±	0.5	4.4	±	0.2
LAINI	1.5	±	0.1	1.5	±	0.8
FRPAR	1.1	±	0.8	0.5	±	0.3
FRDFCB	0.9	±	1.0	−0.1	±	0.1
RGRL	0.8	±	0.5	0.2	±	0.1
BLV	0.6	±	0.9	0.2	±	0.3
TSEMER	0.5	±	0.9	0.6	±	0.2
ASRRT	0.3	±	0.2	−0.1	±	0.1
FRDFBA	0.2	±	0.5	0.0	±	0.2
FRDFBB	0.1	±	0.3	0.3	±	0.4
AAMDVS	0.1	±	0.7	0.1	±	0.6
ASRSO	0.1	±	0.2	−0.2	±	0.1
SCV	0.0	±	0.6	−0.1	±	0.3
KDIF	0.0	±	0.1	−0.1	±	0.2
DVSSEN	0.0	±	0.2	−0.1	±	0.1
ATRL	−0.1	±	0.2	−0.0	±	0.5
DLYATR	−0.1	±	0.4	−0.0	±	0.4
MAINLV	−0.1	±	0.1	0.6	±	0.2
FRDFCA	−0.4	±	0.3	−0.2	±	0.2
AAMTMP	−0.4	±	0.6	−0.2	±	0.1

5.4.3 The uncertainty contributions for the evaluated parameters in LINTUL

The analysis is based on the two samples described earlier (Section 5.2.1) and is identical to that described for SUCROS87. The output variable analysed is the total aboveground dry matter weight. The top and bottom marginal contributions as estimated using a bin smoother are presented in Table 5.7. They are presented as a summary over the 13 growing seasons, characterized by the overall average and the average sampling error. The parameters are ranked in order of decreasing top marginal contribution. There is a large difference for $\top$ and $\perp$ for FINTI and RI which is due to their correlation.

Table 5.7 A summary of the top and bottom marginal uncertainty contribution to total aboveground dry matter weight (TADRW) (estimate ± standard error over samples in %) for parameters in LINTUL. For an explanation of abbreviations cf. Appendix 2.

	$\top$			$\perp$		
TSDVR	43.1	±	2.5	42.7	±	1.0
FINTI	12.6	±	0.9	34.2	±	0.8
LUE	2.3	±	0.2	2.5	±	0.1
DVHALF	2.2	±	1.6	0.6	±	0.2
DVLGTH	2.0	±	1.1	0.3	±	0.0
RI	1.6	±	0.2	23.6	±	1.5
TSEMER	0.5	±	0.8	1.0	±	0.4
FRPAR	0.3	±	0.2	0.3	±	0.1

5.4.4 Selection of parameters for local adaptation

The different estimates of parameter contributions to output uncertainty can be used to select parameters with relatively large contributions to total output uncertainty for further research. Assuming that the number of parameters to be selected was free, a positive top or bottom contribution larger than twice the sampling error was used as a criterion for selecting parameters. Using this criterion some parameters for SUCROS87 are dropped from further consideration. The parameters retained on the basis of this criterion are presented in Table 5.8. All parameters for LINTUL meet this criterion, and therefore none are dropped from consideration. If the number of parameters to be selected is not free and less than 12, subset selection (cf. Section 5.3.4) would have been necessary, given the ambiguous results for both SUCROS87 and LINTUL.

5.4.5 Conclusion

Given the questions asked in the introduction of this chapter, it has now been established that the

uncertainty in selected output due to parameter uncertainty is large for both SUCROS87 and LINTUL. Parameter uncertainty has to be reduced to improve prediction quality. A ranking of contributions of the different parameters to the output uncertainty can be established on the basis of linear regression analysis. Two relevant measures - the top and bottom marginal uncertainty contribution of a parameter - have been used to do so. The results allow to select a number of parameters with non-negligible contributions to the total output uncertainty.

A classification of the selected parameters in terms of their associated processes shows that for SUCROS87 the processes leaf photosynthesis (EFF and AMX), crop development (TSDVRV and TSDVRR), partitioning (BSH), maintenance respiration (MAINRT, MAINLV and Q10), and leaf area development (BSLA, LAINI and RGRL) contribute most to output uncertainty, in roughly that order. In LINTUL most of the output uncertainty is caused by the uncertainty in crop development (TSDVR) and parameters used in the calculation of the reduction of the light use efficiency (FINTI and RI). These results could be used to support assessments of processes as reviewed e.g. by Penning de Vries and Spitters (1991). They state e.g. that simulation performance of growth respiration (parameters ASR'x') is adequate. In combination with the small contribution to yield uncertainty by growth respiration parameters one could formulate conclusions as to research priorities.

Table 5.8 Selection of parameters in SUCROS87 for further research on the basis of the top and bottom marginal uncertainty contribution (estimate ± standard error over samples in %) to total aboveground dry matter weight (TADRW). For an explanation of abbreviations cf. Appendix 1.

	⊤			⊥		
EFF	22.0	±	0.7	22.8	±	1.5
TSDVRV	17.0	±	1.6	11.9	±	1.0
BSH	11.9	±	0.6	11.6	±	1.6
TSDVRR	6.9	±	1.4	0.3	±	0.1
MAINRT	6.4	±	0.4	7.4	±	0.5
Q10	5.4	±	1.2	5.7	±	0.8
AMX	4.4	±	0.8	4.9	±	0.8
BSLA	4.2	±	0.5	4.4	±	0.2
LAINI	1.5	±	0.1	1.5	±	0.8
RGRL	0.8	±	0.5	0.2	±	0.1
TSEMER	0.5	±	0.9	0.6	±	0.2
MAINLV	−0.1	±	0.1	0.6	±	0.2

Chapter 6

Calibration

6.1 Introduction

Calibration in the context of this study

In the previous chapter, it was shown that reduction of parameter uncertainty is essential, as the output uncertainty is too large to allow yield predictions for specific (in this case Dutch) production instances.

In this chapter, we assume that no further direct information regarding the parameters of the model (as in Chapter 4) is available. We will assume that additional information is only available in terms of measurements of variables, e.g. as measured in field experiments. For the additional information to be relevant, the data should minimally fulfil the following conditions:

- The data should meet the conditions under which the model is assumed to be valid. Experiments may not meet these conditions or only to a certain extent.
- The measurements and the model output should refer to the same entity. In some cases this may be difficult to assess. E.g. in LINTUL the question whether the intermediary function FINT and the possibly measured system variable "fraction intercepted light" refer to the same entity is nontrivial to answer.

For the relatively complex simulation models discussed, parameters are commonly estimated by adjusting their values until simulation results and measurements correspond as closely as possible. This is known as calibration.

Very generally, calibration is any procedure to extract quantitative information regarding model parameters from observations using the complete simulation model. In view of the approach used in this thesis, the calibration procedure should at least result in parameter estimates for the production instance that yields the calibration measurements (π_i, cf. Chapter 4). If calibration also yields information regarding the quality of the estimate (η_i), calibration results additionally allow a more accurate specification of the multivariate parameter distribution 'all over the world'.

For our purpose, calibration procedures are best characterized in terms of the results they yield: calibration yielding means π_i, variances η_i and a covariance matrix V_i, calibration yielding only estimates of π_i, or other types of calibration. Other types of calibration are procedures in which a different description of parameter uncertainty is used, such as set-theoretic descriptions (e.g. Keesman, 1989). Procedures yielding π_i, η_i and V_i will be called distribution calibration; procedures yielding π_i will be called point calibration.

Distribution calibration surely is the most relevant option for both models, as the information obtained allows to (re-)establish parameter uncertainty after calibration. However, in the following section we will argue that distribution calibration is not a viable option for both models.

The possibility of distribution calibration

Theory and procedures for distribution calibration are well-developed for models that are linear functions of their parameters, or can be approximated using linear functions of their parameters. For these types of models, calibration is often analysed as a problem of parameter estimation

through optimization. Numerous monographs on parameter estimation and optimization have been written (e.g. Kowalik and Osborne, 1968; Bard, 1974; Gill et al., 1981; Tarantola, 1987). A formulation of the calibration problem as an optimization problem is the following: Given the observations, the model and constraints, determine parameter estimates by iteratively adjusting them in such a way that an objective function is optimized.

The objective function to be optimized is the discrepancy measure, which summarizes a function of all differences between measurements and simulated values. These differences are the residuals. Constraints are imposed to describe and restrict the domain in which solutions should be found. Prescribing parameter ranges, but also fixing a parameter at a specific value, are possible constraints.

In distribution calibration the residuals are used to estimate parameter variance and covariance. This is possible after analysis of the characteristics of the residuals. Residuals consist of a number of components: part of the residuals can be explained by measurement error, another part may be due to parameter fixing errors. The latter occurs if a number of parameters are assumed not to vary between production instances and are fixed at an erroneous value. Yet another part may be due to structural errors: the mathematical formulation of the model may not allow to reduce the residuals further.

The partitioning of the residuals over measurement error and other errors can be analysed, if replicated measurements are available. If residuals are dominated by measurement errors, the description of these residuals in terms of a statistical model and the translation of this statistical model in terms of a discrepancy measure allows the estimation of parameter variances and covariances in the calibration procedure. If residuals are dominated by the other errors, distribution calibration is not possible.

To determine parameter variances and covariances, efficient algorithms for distribution calibration use a methodological shortcut, local linearization. Local linearization allows to reduce the complex problem of estimating parameter uncertainty to the (relatively) simpler problem of analysing the locally linear model and evaluating these results (See the monographs cited, also e.g. Montgomery & Peck (1992)). The alternative to determining parameter variances and covariances by local linearization is a computer-intensive search of the parameter space (see e.g. Klepper and Hendrix, 1994; Aldenberg et al., 1995).

To summarize: Distribution calibration is possible if an error model can be formulated and justified on the basis of the residuals. Distribution calibration can be executed efficiently if the model can be approximated by linear functions of the parameters.

For SUCROS87 and LINTUL the formulation and justification of a model for the errors is problematical. An error model, and the associated discrepancy measure can only be used if measurement errors dominate the residuals. This dominance can be established if replicated measurements are available, which is often not the case.

The multivariate output of the models further complicates the formulation of an error model; the simple assumption that the individual errors are realizations of a single error distribution is generally not justified. Datasets are often small compared to the number of parameters to be estimated. Parameter selection, assuming parameters to be constant, is therefore necessary. This introduces parameter fixing errors of unknown magnitude. In these conditions the quality of the results of a distribution calibration can not be guaranteed.

Furthermore, distribution calibration can not be executed efficiently for SUCROS87 and LINTUL. The reason is that SUCROS87 and LINTUL can not be linearized in all their parameters (Chapter 3), and consequently local linearization as a methodological shortcut can not be used. The occurrence of discontinuities (Chapter 3) will obviously complicate the search of the parameter space which is then necessary to determine parameter variances and covariances.

Given the requirements for distribution calibration and the difficulties to meet these requirements, distribution calibration is problematic, and in any case not practical, for both SUCROS87 and LINTUL. A point estimate based on calibration of a number of selected parameters in both models is, therefore, the best we can hope for.

General issues in calibration

Any type of calibration may yield non-satisfactory estimates for individual parameters. In distribution calibration, parameter estimates are evaluated in terms of the variance of the estimate and its covariance. A parameter estimate with a large variance and a large covariance is not satisfactory. This result may be due to lack of information contained in the data (data are not sufficiently rich) or to the formulation of relations in the model (parameters can not be estimated individually, but only in combination). Very often both aspects determine the problem, especially in the case of large, complex models and small datasets that contain little information.

The problem that parameters can not be estimated individually, or more precisely that different combinations of parameter values yield optimal values of the discrepancy measure, is known as parameter non-identifiability.

To reduce the problem of identifiability, and increase the quality of the estimates (in terms of their variance) the number of parameters to be estimated from the data should be reduced by fixing a larger number of parameters (e.g. Sorooshian and Gupta, 1985). This conditions the estimates of the calibrated parameters on the values of the fixed parameters and consequently increases the fixing error. These conditioned parameter estimates are biased as compared to those estimated from experiments based on individual submodels. The issue in calibrating relatively parameter-rich models like SUCROS87 and LINTUL is to minimize problems of parameter identifiability, while keeping the fixing error as low as possible.

Another issue in any type of calibration is the dependence of its result on the initial parameter guess, the parameter values from which the search is started. The sensitivity of the solution to changes in these conditions is often investigated by repeating the calibration for different initial guesses. If these solutions are different, the problem has multiple, so-called locally optimal, solutions. Ideally all initial conditions result in an identical solution. If this identical solution is also the lowest achievable value of the discrepancy measure given the constraints, the solution is referred to as the global minimum.

The problems which may be encountered in the calibration of SUCROS87 and LINTUL strongly resemble problems encountered in the so-called conceptual catchment (or rainfall-runoff) models (Johnston and Pilgrim, 1976; Sorooshian, 1991; Duan et al., 1992). According to these authors the calibration of these models is characterized by a number of problems: problems of parameter identifiability, indifference of the discrepancy measure to values of 'inactive' (threshold-type) parameters, discontinuities in the discrepancy measure as a function of the

parameters, and presence of local optima.

Given these issues and the difficulties in executing distribution calibration, research in this chapter will be restricted to pragmatic approaches in three situations: 1) What if little or no information for a specific production situation is available? (Section 6.2), 2) What if parameter estimation algorithms have not (yet) been implemented? (Section 6.3), and 3) What if parameter estimation algorithms can be chosen and used? (Section 6.4).

6.2 Calibration using very limited information

If data from detailed field experiments are not available, information of a very general type, such as production characteristics of a crop in the region of interest, may be of value. One can think of yield ranges, time of flowering, or time of crop canopy closure (for maize under Dutch conditions e.g. Sibma, 1987). In crop monographs for a crop in a particular region one often finds rules of the thumb. Output constraints for application of a maize model under Dutch conditions could be derived e.g. from the following statements (Sibma, 1987):

> "Generally formation of a closed canopy and a leaf area index larger than 3 is not achieved until a few weeks after the longest day."
>
> "The light interception by a maize canopy never reaches 100% and is maximally of the order of magnitude of 95%."

Sometimes general rules of thumb are available, such as a condition for potential production that is formulated in terms of bounds on the growth rate (Rabbinge et al., 1989):

> "The growth rate of the crop in these conditions (of potential production) is determined by weather conditions, and in terms of dry matter amounts to 150-350 $kg.ha^{-1}.d^{-1}$ when the canopy fully covers the soil."

These and other statements are best interpreted in terms of acceptable and non-acceptable values for specific model outputs. Alternatives, such as translating the statements in terms of probability distributions require more assumptions. Assuming that the available information defines acceptable and non-acceptable model outcomes, it is straightforward to analyse this type of information in a set-theoretic context (cf. Chapter 4). We can simulate repeatedly with randomly generated parameter values to generate different outcomes and select those parameters for which the simulated outcome lies within the acceptable range. The approach used in this section is similar to Regionalized Sensitivity Analysis (RSA) (Hornberger and Spear, 1980; Spear and Hornberger, 1980). The method will be illustrated for a simple example. The procedure can be sketched as follows:

1) Formulate output constraints on the model output for the population of production instances for which you would like to calibrate your model. The output constraints are hard - set theoretic - boundaries.
2) Take parameter values at random, execute simulations, and discard those parameter combinations which violate the output constraints. All available information regarding the parameters should be used. In our case the simulations are necessarily executed using the prior parameter distributions derived in Chapter 4.
3) Analyse the remaining parameters and use the results. The result is a sample from the

restricted parameter domain, that can be used in different ways, e.g. for prediction (but see conclusions at the end of the example in Section 6.2.). If the parameter values are generated from distributions, the formal characterization of the new domain (as a set is combined with a distribution) is not straightforward.

Restricting the parameter domain: an example

An example of the procedure to restrict the parameter domain given output constraints will be given for the model LINTUL. In this example the maximum growth rate under potential conditions was used as the criterion. The basic data for this example were the parameter sample and the simulation results from the uncertainty analysis in Chapter 4. In that uncertainty analysis Monte Carlo simulation was executed for a sample of 1000 parameter vectors, over thirteen years using Dutch weather, known planting and harvesting dates. The parameter vectors and the maximum growth rate during the growing season were retained for the analysis.

Analysis and results

The analysis was based on the criterion that the maximum growth rate for potential production should be within 150-350 $kg.ha^{-1}d^{-1}$. In view of the difficulty of quantifying full cover, this condition in the earlier quote was dropped.

The results are presented in terms of the changes in the averages, variances and the correlations, with respect to the characteristics of the sample originally used (Table 6.1a-b). Out of an initial sample of 1000 parameter vectors 293 sets of parameter values were retained that met the constraints in all thirteen years of simulation. The averages and variance of the total aboveground dry matter weight for the different samples are presented in Table 6.2. It is important to note that these characteristics are used to summarize the sample and should not be interpreted as characteristics of a post-calibration parameter distribution.

Table 6.1a Effect of domain restriction on means and variances of parameter distributions. Parameter values are expressed relative to their nominal value (cf. Appendix 6); the % change is expressed relative to the mean of the initial sample. For the parameter FINTI* also the median and relative change of the median are given (between brackets).

	mean of initial sample	% change w.r. to initial sample	variance of initial sample	% change w.r. to initial sample
TSEMER	0.683	–2.8	0.0828	3.5
TSDVR	1.425	–33.8	0.4053	–73.4
FRPAR	0.935	–0.2	0.0024	–4.2
FINTI*	0.919 (0.190)	25.4 (152.9)	2.3500	8.2
RI	1.000	0.5	0.0995	–7.3
DVHALF	0.985	3.6	0.0488	3.3
DVLGTH	0.994	15.4	0.6938	12.1
LUE	1.62	–2.0	0.0397	–20.4

Table 6.1b Effect of domain restriction on correlations between parameters. Changes are expressed relative to correlations in the initial sample (correlations to be induced are based on Table 4.8).

		TSEMER	TSDVR	FRPAR	FINTI	RI	DVHALF	DVLGTH	LUE
TSEMER	initial	1							
	% change	0							
TSDVR	initial	–0.004	1						
	% change	2300	0						
FRPAR	initial	–0.003	0.008	1					
	% change	–2533	863	0					
FINTI	initial	–0.013	0.027	–0.002	1				
	% change	–215	307	–2250	0				
RI	initial	0.009	0.001	–0.002	–0.525	1			
	% change	733	19100	3750	5	0			
DVHALF	initial	0.001	–0.006	–0.002	0.041	0.012	1		
	% change	400	550	–2600	107	–258	0		
DVLGTH	initial	0.005	–0.012	0.001	0.035	0.007	0.892	1	
	% change	560	217	5700	189	–829	2	0	
LUE	initial	0.003	0	0.002	–0.013	–0.008	0.006	0.022	1
	% change	–1133	n.d.	–5500	277	1088	483	282	0

Table 6.2 Effect of domain restriction on means and standard error of yield distributions in each of the 13 years (TADRW, $kg.ha^{-1}$) with respect to the initial values.

year	mean of the initial sample	% change w.r.t. initial sample	standard error of initial sample	% change w.r.t initial sample
1973	6099	36.4	3901	–24.4
1974	4435	77.2	3870	–43.3
1975	5708	36.6	3678	–25.9
1977	4536	63.6	3581	–38.7
1978	4162	82.6	3798	–46.3
1980	4333	62.0	3353	–40.2
1981	4617	62.6	3596	–39.2
1982	5222	48.1	3637	–33.1
1983	5006	44.2	3381	–28.8
1984	3199	75.5	2786	–43.3
1985	4940	61.7	3780	–39.3
1986	5835	57.7	4443	–33.9
1988	4780	60.8	3635	–38.1
mean	4836	57.4	3649	–36.3

Imposing bounds on the maximum growth rate results in a consistent increase in the yield of aboveground dry matter and a decrease in the yield variation due to parameter uncertainty as compared to the results for the original sample (Table 6.2). The average yield is more in keeping with values expected under Dutch conditions. This allows to conclude that very general information regarding output variables of the system, in combination with local planting and harvesting dates, and meteorological data, can be used to adapt a model to those local conditions by excluding parameter values and parameter combinations which do not satisfy the information available. It is possible to adapt all parameters simultaneously.

If the parameter values retained are summarized in terms of a distribution, the results as compared to the initial distribution show changes in parameter means and in parameter variability (both increases and reductions) and show that a number of generally small correlations between parameters are introduced. The effects differ for each parameter. The change in the value of the mean is relatively large for TSDVR, FINTI and DVLGTH, which suggests that the values for the 'world' are on average too large (TSDVR) resp. too small (FINTI and DVLGTH) for Dutch conditions.

The extension of this procedure to a full-fledged calibration procedure using more extensive datasets is possible. Instead of defining one single set which model outcomes have to meet, each observation is reformulated in terms of a set, and has to be met by the simulation results using a specific parameter vector. Executing a set calibration procedure using the results of an uncertainty analysis (as in the example) may have advantages in those cases where the model uses a sensitive numerical procedure, which gives rise to unexpected crashes of the model. If executed as separate runs, a model crash for a parameter combination means that output for that specific parameter combination is not available, whereas in iterative adjustment procedures the procedure crashes in its entirety.

However, for larger datasets, the procedure followed in the example, discarding parameter vectors from a random sample, is inefficient if the probability of generating an acceptable parameter vector is small, and if parameter sets have to be determined in more detail. In these instances the procedure used by Klepper and Rouse (1991), and Scholten and van der Tol (1994) based on a controlled random search method, is more efficient. A further refinement of this type of calibration has been developed and used by Keesman (1989). As in distribution calibration, problems should be expected in the justification of the set bounds for individual observations.

6.3 Manual point calibration

In the case that experimental data are available, but a calibration algorithm is not (easily) available, calibration is often executed manually. Manual point calibration is a trial and error calibration method. It is characterized by little formalization and is executed by manually changing parameter values and visually evaluating the discrepancy between measurements and simulated data.

The following publications explicitly refer to manual point calibration (as trial and error

calibration): Karnieli et al. (1994); Holmes et al. (1989), Hopmans and Guttierez-Rave (1988); Patoine and Fortin (1992). For other examples of trial and error calibration, cf. Schouwenaars (1990); van Lanen (1991); Querner, (1993); Stutterheim, (1995). The workshop studies of Groot et al. (1991), van Grinsven (1993), Kabat et al. (1995), and McVoy et al. (1995) give an impression of calibration strategies and predictive performance for different simulation models using the same data set.

Manual point calibration is said to be successful for a small number of parameters (Janssen and Heuberger, 1995; Tarantola, 1987). Generally however, results are thought to be difficult to reproduce, and given that manual point calibration is little formalized, a discussion of possible room for improvement or of possible drawbacks is difficult. Aspects which might be considered an advantage are the informal incorporation of expert knowledge, the time gained if the model does not have to be incorporated in an optimization routine and the 'feel' for the model achieved by 'hands on' execution. Whether manual point calibration yields better predictions has to be tested in practice. Comparisons between the predictive performance of models calibrated by hand and those calibrated using an automatic method are rare. One example is the analysis by Holmes et al. (1989), who compare the prediction results of manual point calibration and an automatic method; they conclude that the best strategy is to use both methods conjunctively, as it allows to incorporate expert knowledge. A first attempt to evaluate the effect of the model user on the calibration result for a single hydrological model is presented by Botterweg (1995). He concludes that calibration results strongly depend on background and knowledge of the model user.

The above strongly suggests that manual point calibration procedures may benefit from limited formalization, e.g. in terms of a script, describing steps in a possible procedure. Formalization of calibration procedures has the same advantages as formalization of system models, i.e. ease of presentation, ease of transfer to other users, ease of discussion and assessment, and ease of result reproduction (Chapter 1). In this section, an attempt will be made to formalize manual point calibration in terms of a script.

Hand calibration for a single output variable

The starting point for a possible script is a calibration problem in which one output variable is calibrated using a number of parameters. To keep the script as simple as possible, the calibration will be executed as a series of sequential line searches: once a first parameter has been calibrated, its value remains fixed during the following steps of the calibration. In a following step the second parameter is calibrated to improve the fit. Once the fit can not be improved, both parameters are kept fixed at their calibrated value during the calibration of the following parameters. Whereas the above defines the procedure to adjust parameter values, the choice of a visual discrepancy measure is less straightforward. Visual comparison of the differences between non-linear functions is a difficult task and may lead to misjudgements (Cleveland and McGill, 1984). Literature suggests that the simpler and more to the point the representation of the information, the better the evaluation results (Cleveland and McGill, 1984; Flowers and Hauer, 1993; Wainer, 1984). A comparison between visual fitting of straight lines and linear least squares regression (Mosteller et al., 1981) showed that correspondence between linear regression results and visual

regression results was good, and that the individual-to-individual variability in slope and intercept estimates was near the standard error provided by least squares. Given these results, the results of manual point calibration are therefore best evaluated in terms of a plot of observed values vs. simulated values and analysed in terms of the visual fit of a single straight line. A perfect calibration should result in a plot of observed vs. simulated values with all data on the 1:1 line. Visual evaluation of plots of simulated against observed values introduces the problem how to document choices. To document the different choices, one may summarize the different plots in terms of slope, intercept and correlation using linear regression between observations and simulated values. The fit is then evaluated in terms of three discrepancy measures: the numerical values of slope of observed vs. simulated, which should be as close to 1 as possible; the intercept, which should be as close to 0 as possible, and correlation of the relation between artificial data and simulated values, which should be as close to 1 as possible.

Hand calibration of multiple output variables

The above refers to the calibration of a single output variable. The models discussed require a procedure for manual point calibration if information on several output variables is available. To keep the evaluation task as simple as possible (cf. the previous discussion regarding visual evaluation) the result of a manual point calibration of multiple output variables is preferably evaluated in terms of a plot of observed vs. simulated values. This will require expressing the output variables on the same scale, e.g. as relative values.

A different approach is possible if the model is hierarchically structured as defined in Chapter 3. In that case a sequential calibration of individual output variables may be possible. A sequential calibration procedure based on the structural analysis can be briefly recapitulated as follows:

- Determine relations between the measured variables and the parameters and present them as a structural matrix.
- Exchange columns and exchange rows until the structural matrix is closest to a triangular matrix.
- Calibrate the specific parameters and variables one at a time in the order prescribed by the structural matrix.

In both cases the calibration script for multiple output variables is similar to the calibration of a single output variable, either because a single discrepancy measure is used, or because the problem can be reduced to a sequential calibration of the individual output variables.

An example of manual point calibration: one output variable

In the previous section a possibility to execute hand calibration in a more formal way was discussed. To illustrate this procedure and investigate what should be documented to be able to repeat the analysis, we present the following example in which a single output variable has been measured, and in which we try to adhere as strictly as possible to the procedure sketched. The example also serves to illustrate the problems one may encounter in hand calibration.

Method

SUCROS87 was used to generate error-free artificial observations of TADRW at 10-day intervals, using the management data for an experiment in 1985, and the median values for two parameters (RGRL (0.9499) and EFF (0.7346), relative to their reference value), which were regarded as unknowns. The rest of the parameters were kept at their nominal value (=1) and assumed known. The procedure executed was the following: In the first line search (at the nominal value of RGRL (= 1)), the parameter EFF was calibrated. In the second line search EFF was then kept at the calibrated value and the second parameter RGRL calibrated. To test the calibration results, the calibration was repeated changing the order in which the parameters are calibrated: RGRL was calibrated at fixed, nominal EFF in a first line search; in the second line search EFF was calibrated at the estimated value of RGRL. The line searches were executed using 200 values of the parameter to be calibrated within the range 0.5-1.5 (range relative to nominal parameter value). The best parameter value in terms of the discrepancy measure used was chosen as the parameter estimate.

The fit was evaluated in terms of a linear regression between observations and simulated values as argument. Fit was documented in terms of three discrepancy measures: slope of observed vs. simulated, the intercept, which should be as close to 0 as possible, and the correlation of the fit between artificial data and simulated values, which should be as close to 1 as possible.

Results

The results of the different calibrations and the intermediate steps are presented in the Table 6.3. The different discrepancy measures (slope, intercept and correlation) do not jointly define a single best solution: in each step a decision has to be made which single discrepancy measure defines the subjectively chosen 'best' solution. An overview of the criteria values selected in each of the calibration steps is given in Table 6.4.

Understanding the choice of the different discrepancy measures

To understand why the three different discrepancy measures do not yield a single best parameter value, the surface describing the calibration criteria as a function of the two parameters to be calibrated was established using a systematical grid (40×40) over the ranges 0.5-1.5 (relative to their nominal value) for both parameters. The calibration discrepancy measure surfaces are presented in the Figures 6.1-6.3. In Figure 6.1, which represents the value of the slope of the plot of measurements vs. simulations (M-S), which ideally equals 1, the arrangement of optimal values suggest a non-linear relation between the estimates of EFF and RGRL, along which the criterion is 'best'. This is also true for the value of the intercept of the regression M-S (optimal value 0; Figure 6.2). In contrast to the other criteria, the contour lines using the coefficient of determination (Figure 6.3) as a criterion are less ambiguous. One may show that the coefficient of determination used as a criterion is a transformed residual sum of squares criterion. Minimizing the residual sum of squares corresponds to maximizing the coefficient of determination. It should

Table 6.3 The effect of the sequence in which parameters (RGRL, the relative leaf area growth rate during juvenile leaf area growth, and EFF, the initial light use efficiency for individual leaves) are calibrated on the results of manual point calibration (true values in brackets, parameter values relative to default).

	Step 1		Step 2	
Calibration 1	RGRL fixed	EFF calibrated	EFF fixed	RGRL calibrated
	1	0.71 (0.7346)	0.71	0.99 (0.9499)
Calibration 2	EFF fixed	RGRL calibrated	RGRL fixed	EFF calibrated
	1	0.97 (0.9499)	0.97	0.72 (0.7346)

Table 6.4 Correlation, slope, and intercept of the linear fit between observed and simulated values in manual point calibration. The criterion used is in bold type face.

	Calibration 1		Calibration 2	
	Step 1	Step 2	Step 1	Step 2
Correlation	1.000	1.000	**1.000**	1.000
Slope	**0.9858**	**0.995**	0.6542	**1.0023**
Intercept	−47.71	−41.1	**1.7**	−15.66

be noted however, that the sensitivity of R^2 near the optimal parameter values is considerably less than that of the - non-transformed - residual sum of squares.

If the different line searches are analysed using the results of the grid search, as presented in the Figures 6.1-6.3, the choice of the discrepancy measure and the final results can be understood. In the line search at which RGRL = 1, the criterion slope yields a single best value for EFF in the first step. Using intercept as calibration discrepancy measure would yield optimal values of EFF outside the assumed range. In the second line search (EFF fixed at its estimated value) varying RGRL, while improving the value of the intercept, decreases the value of the slope. The small change of RGRL as compared to its initial value can now be understood. In the second calibration (for the line search at which EFF = 1) the use of either slope or intercept as calibration criterion would yield two completely different 'best' values of RGRL (about 0.6 or about 0.9), and the necessity to use correlation as additional discrepancy measure becomes understandable. The second line search along the estimated value of RGRL to calibrate EFF is unproblematic, using slope as the evaluation criterion.

From Figures 6.1-6.3 it is quite clear that in this manual point calibration procedure the parameter values used to generate the data can only be found after a larger number of iterations. It is inefficient to do this by hand, and would require a multiple of the 400 runs executed.

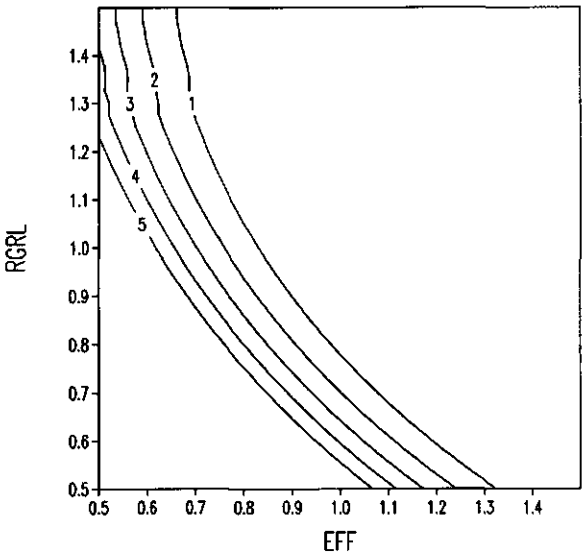

Figure 6.1 Lines of equal 'slope' as criterion in an example of manual point calibration (see text) in the parameter space of the parameters EFF and RGRL. The optimum value of this criterion equals 1 and has coding 3 in the figure. Levels 1, 2, 4 and 5 refer to slopes of 0.8, 0.9, 1.1, and 1.2, respectively. Parameter values are relative to their default values.

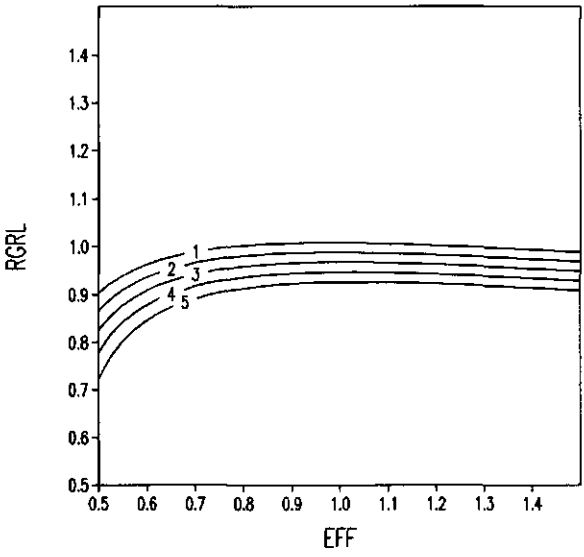

Figure 6.2 Lines of equal 'intercept' as criterion in an example of manual point calibration in the parameter space of the parameters EFF and RGRL. The optimum value of this criterion equals 0 and has coding 3 in the figure. Levels 1,2, 4 and 5 refer to intercepts of −40, −20, 20 and 40 kg ha^{-1}, respectively. Parameter values are relative to their default value.

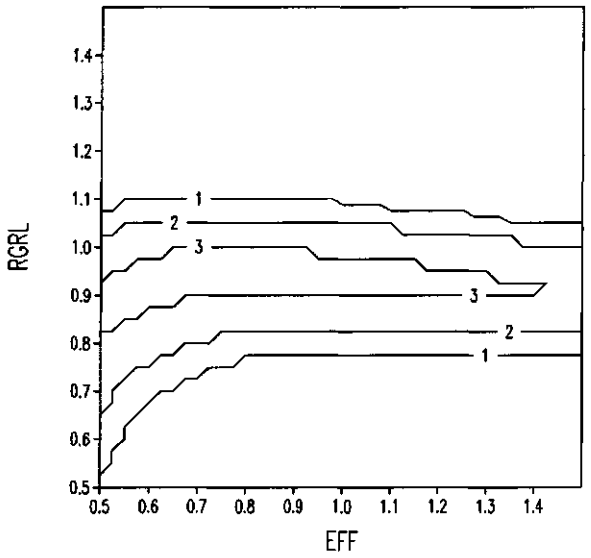

Figure 6.3 Lines of equal 'coefficient of determination' as criterion in an example of manual point calibration in the parameter space of the parameters EFF and RGRL. The optimum value of this criterion equals 1 and has coding 3 in the figure. Levels 1 and 2 refer to values of the correlation coefficient of 0.9996 and 0.9998, respectively. Parameter values are relative to their default value.

Discussion regarding the example and conclusions

The example presents a possible script for a specific type of manual point calibration. To execute the calibration in a reproducible manner, the order in which parameters are calibrated has to be documented. If the order in which the parameters are not documented the example shows that the procedure is irreproducible as long as the calibration procedure is not iterated. This result is not new and is to be expected in any situation in which parameters have effects on the calibration variable that are to some extent exchangeable. A discussion is presented by e.g. Ibbitt and O'Donnell (1971).

The documentation of a qualitative visual evaluation criterion in terms of several quantitative criteria also introduced a need for further documentation. The decision to adhere to visual interpretability was shown to be a disadvantage. Choices between criteria have to be made and documented from calibration step to calibration step. To reduce these criteria to a single criterion, it would be preferable to use the residual sum of squares, or its root. In a plot of observations vs. simulations the root criterion can be visualized as the sidelength of a square that represents the sum of all residual squares (cf. Figure 6.4).

In contrast to the above line search, a grid search results in a calibration procedure, in which less decisions have to be documented, and in which the discrepancy measures can be evaluated

and combined. However, as the number of calibration parameters increases, the number of simulation runs becomes very large.

It is possible to calibrate a number of parameters by hand using data from one output variable in a procedure which is executed in a number of reproducible simple steps. However, for the procedure to be reproducible, the number of - seemingly trivial - decisions to be documented is high and is expected to increase with the number of parameters and output variables calibrated. If the decisions are not documented, the results of manual point calibration may become irreproducible.

6.4 Point calibration using optimization algorithms

6.4.1 Introduction

Point calibration is a calibration method that yields a single vector of parameter values using a parameter estimation algorithm. In point calibration, the following questions have to be addressed:

- which discrepancy measure should be chosen,
- which parameters are to be estimated, and
- which optimization algorithm may or should be used.

In connection with the last problem the question of stopping the procedure - the termination criterion - has to be discussed.

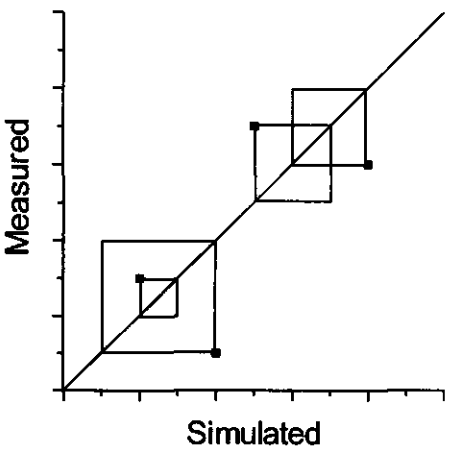

Figure 6.4 Given the graphical presentation of the results of a calibration in terms of a plot of measurements against simulated values, this figure presents the graphical interpretation of the L_2-norm as 'sum of squares'

6.4.2 Selecting the optimization algorithm

For overviews over possible optimization methods the reader is referred to e.g. Gill et al. (1981), Kowalik and Osborne (1968), Tarantola (1987), and Bard (1974). Several optimization methods require that the objective function, the discrepancy measure should be (at least) twice differentiable with respect to the parameters. Methods not requiring derivatives are known as direct search procedures. The continuity of the relation between model output and parameters was analysed in Chapter 3, and it was clearly shown that it does not hold for all parameters of SUCROS87. As a consequence the first derivative of the discrepancy measure is not defined for specific values of different parameters. Formally, derivative-based algorithms should not work. Practically, optimization methods based on numerical approximations of the derivatives may work sometimes. This is indicated by different comparisons between derivative based optimization methods and direct search methods for conceptual catchment models (CC-models) presented by Ibbitt and O'Donnell (1971), Johnston and Pilgrim (1976), Hendrickson et al. (1988), and Sorooshian et al. (1993). The comparisons are relevant for the calibration of SUCROS87 and LINTUL, as the CC-models also contain discontinuities in the relation between output variables and parameters (cf. Hendrickson et al. (1988) for an example). The comparisons allow the common conclusion that direct search algorithms require more time, but are at least as reliable in finding the true values of parameters in artificial calibration experiments as derivative based optimization methods. In other words they are slow, but robust.

It was therefore decided to use robust parameter estimation methods that do not need the use of partial derivatives of output variables with respect to parameters. This limits the possibilities for both models to direct search algorithms. Two direct search methods will be used here, a controlled random search procedure according to Price (1979) and a multiple restart version of the Simplex method (Nelder and Mead, 1965). The method used in the calibrations executed in this thesis is based on a direct search algorithm by Price (1979) and its implementation in software (Stol et al., 1992). The Simplex procedure will be used as a check on the results of the Price method.

Description of the optimization algorithm

The algorithm chosen for calibration is a controlled random search procedure to determine the global minimum of a function with constraints on the parameters. It aims at gradually improving an initial set of parameter vectors by iteratively replacing the worst parameter vector in this set by a better one. The parameter vectors contained in the set are used to generate the new trial vector. When discussing the calibrations actually executed we will describe the modifications implemented for this study.

In an initial step the algorithm generates a set of size N_g of the n parameters to be calibrated. The initial set is randomly generated within a hyperbox defined by constraints on the parameter values. The discrepancy measure is calculated for each of the N_g parameter vectors. In the following steps a new parameter vector P is generated from the existing ones. Its discrepancy measure is compared to the worst of the N_g parameter vectors. If the new vector yields a better discrepancy measure, the new vector replaces the worst; if not, nothing changes. This procedure

is repeated as long as a stop criterion or the maximum allowable number of simulations N_s is not met. The hope is that with each successful step the set of vectors will cover a smaller parameter domain.

The generation of the new parameter vector P is based on the following procedure: $n+1$ parameter vectors are randomly chosen from the N_g stored. The $n+1^{th}$ vector R_{n+1} is used to generate the next trial vector with respect to the vector of averages G of the other n vectors:

$$P_{new} = 2G - R_{n+1} \tag{6.1}$$

The new point is mirrored with respect to the average G.

The algorithm is defined by four parameters, one fixed, three free. The number of parameter vectors selected to generate a new trial vector (n) is fixed at the number of parameters to be calibrated; the size of the initial set N_g; the maximum allowable number of function evaluations N_s; and the termination criterion. The choice of a multiplication factor (2) and the choice of $n+1$ vectors are reminiscent of the Simplex algorithm. In the implementation the $n+1$ vectors are either resampled, if the generated new parameter value is outside the hypercube defined by the parameter constraints, or the new parameter value is set to the parameter bounds crossed.

Although the method is intended to yield point estimates, the final result is a cloud of parameter vectors, the differences between which are determined by the termination criterion. For a description of the Simplex algorithm cf. e.g. Press et al. (1992).

6.4.3 Discrepancy measures

In the introduction of Chapter 6, the concept 'discrepancy measure' was introduced. In case of distribution calibration the choice of the most appropriate measure is dictated by the statistical model for the measurement errors. In manual calibration the discrepancy measure was defined in terms of a visual evaluation of the fit. Criteria for the selection of a visual criterion are accuracy and reproducibility. Apart from the fact that discrepancy measures should be zero for a perfect fit, the choice of a discrepancy measure for point calibration is not subject to any condition.

Discrepancy measures summarize the quality of the fit in a single value, and are a scalar function of the vector of simulated data S and measured data M. An often used class of discrepancy measures, called L_p norms, is defined as:

$$L_p = \left(\sum_{i=1}^{n} |M_i - S_i|^p \right)^{\frac{1}{p}} \tag{6.2}$$

where p is a real number greater or equal 1 and M_i and S_i refer to the individual measured and simulated quantities respectively. Values $p = 1$ and $p = 2$ are often used.

The discrepancy measure with value $p = 2$ will increase the importance of relatively large errors. For larger values of p this tendency increases and culminates in a discrepancy measure equivalent to the limit case for $p \rightarrow \infty$. This limit is dominated by the largest individual

difference and is determined by a single difference value. This discrepancy measure is implemented as:

$$L_\infty = \max\left(\left|M_i - S_i\right|\right) \text{ for } i = 1...n \tag{6.3}$$

and represents the absolute value of the maximum individual difference. The optimization procedure is used to minimize this maximum distance, and is therefore sometimes called a minimax procedure. Extensions of these criteria allow weights for the individual observations, or - in the case of a criterion for multiple output variables - allow weights for the individual output variables (see e.g. Montgomery and Peck, 1992). The general rule is that weighing data or output variables only makes sense if information on the accuracy of the different data is available.

The discrepancy measure chosen

In the algorithm used the discrepancy measure for multiple output variables consists of the L_∞ norm combined with a L_2 norm of the relative deviations for individual output variables.
This discrepancy measure C_∞ is calculated from the measurements M and simulation results S as:

$$L_{2j} = \left(\sum_{i=1}^{n_j}\left(\frac{M_{ij} - S_{ij}}{M_{ij}}\right)^2\right)^{\frac{1}{2}}$$

$$C_\infty = \max(L_{2j}) \tag{6.4}$$

where the i is the index of n_j data per output variable j, and j is the index of the m different output variables. This discrepancy measure is suitable if measured values are strictly positive.

The use of relative errors gives the same weight to output variables measured on different scales; the use of the L_2 norm is a standard, and the use of the L_∞ norm over the output variables links the relative error to a single output variable, which may be of diagnostic value. In both optimization algorithms (Price and Simplex) the same discrepancy measure was used.

6.4.4 Selecting termination criteria

After choosing a discrepancy measure, the next question is when to stop the calibration. Calibration should stop when the discrepancy measure has reached the global minimum, the smallest value possible within the search domain. In that case the optimization procedure is said to have converged. A necessary, but not sufficient, practical condition for convergence is that the value of the first derivative of the discrepancy measure with respect to the calibrated parameters should have become smaller than a user-defined constant ε, the termination criterion. As this derivative may not be defined for SUCROS87 or LINTUL, this is not a satisfactory option.

A loosely formulated alternative is to stop if the parameter estimates from one iteration of the optimization procedure to the next do not change any more. Evaluating the stopping criterion in terms of changes from one iteration to the next is less preferable, as this is not only a function of the convergence of a procedure, but also of its search efficiency. As some algorithms, in this case the Price algorithm, are relatively inefficient, the probability that nothing changes from one iteration to the next is high. For an overview of other termination criteria cf. Gill et al. (1981).

Apart from what could be called criterion for 'normal' termination of a calibration, one also has to impose another termination criterion for abnormal conditions. This is necessary if the calibration job does not converge, and would keep on adjusting parameters, or simply because of the amount of time available. The abnormal termination criterion is often a maximum number of iterations. A rule of the thumb is difficult to formulate, as it depends on the optimization problem and on the optimization algorithm.

Choice of termination criterion

The stopping criterion implemented in the optimization procedure chosen is based on the relative difference between the values of the discrepancy measures of the worst (C_w) and the best parameter vector (C_b) in the set, as:

$$\varepsilon = \frac{C_w - C_b}{C_w} \tag{6.5}$$

As C_w is larger than C_b, and C_b is generally larger than 0, there is no danger that this criterion yields non-defined values. Interpreting the results in terms of a set calibration, all parameter vectors contained within the final cloud (defined by the required ε) are regarded as acceptable. For a point calibration the parameter vector associated with the best discrepancy measure is regarded as the final result. The termination criterion in the Simplex procedure is identical to that used in the Price algorithm. In contrast to Price the Simplex procedure is restarted from the then optimal point until C_b does not improve.

6.4.5 Parameter selection

Both the size of the dataset itself and the algorithms used for parameter estimation pose a limit to the number of parameters that can be simultaneously estimated from any dataset. Parameter fixing error and parameter non-identifiability which were briefly discussed in the introduction may force one to select a limited number of parameters for calibration.

Parameter fixing error

The issue of the parameter fixing error, the effect of keeping some parameters constant on the value of the estimated parameters, was introduced earlier. The only satisfactory solution to keep this problem small is to calibrate all parameters. This is not possible. However, the uncertainty

contributions of the different parameters to the output uncertainty (cf. Chapter 5) suggest to select those parameters for calibration that have the largest effect on output uncertainty. Fixing the parameters with the smallest contribution to output uncertainty at their nominal value will hopefully introduce the smallest systematic error in the model output. The analysis of parameter contributions presented in Chapter 5 was used to select parameters for calibration.

It should be noted however, that the selection was based on an earlier analysis of parameter contributions for SUCROS87 using a different selection criterion (position among the highest dozen values of τ, analysis using a bin smoother) and based on values of LAINI which were too low. This yielded a slightly different selection, which is used in calibration. In the selection used in calibration the parameters FRDFBB, DVSSEN and ASRRT were chosen instead of LAINI, TSEMER and MAINLV.

Parameter identifiability

The uncertainty analysis executed in Chapter 5 allows to define priorities as to which parameters should preferably be estimated from the data, given their contribution to prediction uncertainty. If parameter uncertainty is viewed in terms of a joint probability distribution, calibration can be seen as an attempt to reduce prediction uncertainty by reducing parameter uncertainty.

As described in Section 6.1.1, it is possible that, after calibration, the uncertainty of some of the calibrated parameters is not reduced in terms of their variance. This is possible if the data do not contain any information regarding the parameter; another possibility is that the reduction of uncertainty appears in the covariance terms rather than in the variance of the individual parameters. In the first case - the data do not contain any information regarding the parameter - the wrong parameter has been selected for calibration.

The second case - the reduction of uncertainty does not show up in terms of the variance of the individual parameters - is more complex. As an example, assume that two parameters A and B only occur in a model as their sum $A+B$. In an artificial calibration experiment executed as a grid search to estimate A and B results will show that the parameter estimates of A and B are strongly correlated. In fact a plot of A vs. B will show that A is a linear function of B with a slope of -1. In terms of the joint probability distribution of A and B their covariance (correlation) has become better known, whereas individually the distributions of A and B will have changed very little. Fixing one parameter allows to redirect the effect of the information to a decrease in the uncertainty of either parameter A or B.

Instances as this example for A and B become problematical if point estimates are made and aspects of joint parameter distributions such as correlation are not considered. In the context of point estimation this problem shows up as a problem of non-unique estimates: equal model output for different parameter vectors does not imply that these parameter vectors are equal. Plotting the discrepancy measure as a function of both parameters will show results similar to those in Figures 6.1-6.3: the value of one parameter is a function of the other for similar values of the discrepancy measure.

The example suggests that careful selection allows to reduce this problem, as long as the non-selected parameters are not fixed at values which are impossible for the production instance. The alternative is to refrain from estimating both parameters using the entire model, or to

reparameterize the model in terms of a new identifiable parameter ($A+B$). These options are not considered in this case study.

The question is how to detect, circumvent or counter these type of problems in a point calibration. As point calibration generally does not yield information regarding the relations between parameters for similar values of the discrepancy measure, the options are:

- to analyse identifiability prior to calibration,
- to select a point calibration procedure which additionally yields the required information for input in a post-calibration analysis,
- to use a point calibration procedure in such a way that the problem is circumvented.

Pre-calibration analysis of parameter identifiability

Analysis of identifiability for non-linear models that yields results that hold independent of parameter values (known as analysis of global identifiability) is complicated (e.g. Walter, 1982). In contrast, for distribution calibration in its simplest form, linear regression, there are several methods to analyse whether problems in estimation should be expected (Montgomery and Peck, 1992). These methods can be applied for non-linear models if the model can be locally linearized in the parameters. Local linearization is based on the use of the derivatives of the calibration variables with respect to the parameters for a given parameter vector. For an example cf. Mous (1993). This type of analysis is known as local identifiability analysis, and can be executed for the nominal parameter values (pre-calibration) or parameter estimates (post-calibration). For the models under discussion linearization is not possible for a number of parameters (cf. Chapter 3). An alternative, simple linear regression of model output on the full range of a single parameter for all parameters using the results of uncertainty analysis, is used by Klepper (1989). After clustering parameters with similar linear effects on the output, Klepper selects parameters from each cluster for calibration. This approach would in principle also be possible for SUCROS87 and LINTUL, but the low percentage of variance accounted for by linear regression (cf. Chapter 5), and the occurrence of non-linear relations with internal extremes (Chapter 3) suggests limited usefulness of the results. Given the problems sketched, an analysis of identifiability problems prior to calibration is not considered in this thesis.

Post-calibration analysis of parameter identifiability

As indicated in the example, the analysis of identifiability requires an analysis of the relations between parameters for similar values of the discrepancy measure. In calibration algorithms that provide information about the values of the parameter for similar values of the discrepancy measure, analysis of the results can yield indications for limited identifiability. Limited identifiability is indicated by correlations between parameter estimates, and by parameter estimates, the uncertainty of which has not been reduced at the end of calibration.

Obviously a definition of the above-mentioned 'similar' values of the discrepancy measures is called for, before both parameter uncertainty and correlations can be calculated. Once this similarity has been defined and an optimization algorithm yielding information regarding these two characteristics is used, parameters can be further selected on the basis of these results.

Circumventing analysis of parameter identifiability

In both pre- and post calibration analysis, the analysis of identifiability is only possible if information about parameter values for similar values of the discrepancy measure is available. Some optimization procedures, such as the Simplex procedure, in principle only yield a single point estimate. These procedures have to be extended to make the required information available.

One may also altogether circumvent the extension of the optimization procedure to analyse parameter identifiability. To illustrate the approach, assume that two parameters (A and B) only occur in the model as their product. If parameter A has already been calibrated, adding parameter B will not allow to improve the discrepancy measure further. Adding C as a second parameter may improve the calibration result as compared to adding parameter B. Systematic execution for all possible combinations of two parameters will yield the pair of parameters which allows the best fit to the data. Systematically calibrating the model, trying all possible different combinations of parameters for all possible numbers of parameters, allows to select that combination of a specific number of parameters for which the calibration result is best.

There are two variants of this procedure: an approach in which the fit (the descriptive performance) is optimized, and an approach in which the predictive performance is optimized. For a prescribed number of parameters to be calibrated, one would select that parameter set for calibration which yields the best fit or prediction. To compare the performance of parameter sets of different sizes one would have to use a measure which balances fit or prediction error against the number of parameters calibrated. This approach is known in regression literature as subset selection. Criteria and methodology is discussed by e.g. Montgomery and Peck (1992), Hocking (1976), and Miller (1990).

Detecting identifiability problems using Price

Given the random search character of the Price algorithm, it is too slow to allow subset selection procedures to circumvent analyses of parameter identifiability. Pre-calibration analysis is not an option, because of the differentiability problems. Post-calibration analysis is the only remaining possibility. Post-calibration analysis of parameter identifiability requires a notion of parameter uncertainty after calibration, and an estimate of correlation between parameters after calibration. It can be argued that the Price algorithm yields the required information.

To recapitulate, the Price algorithm starts with a cloud of parameter vectors. The size of this cloud should decrease in the long run as parameter vectors yielding badly fitting results are discarded. Assume that a parameter which has little effect on the discrepancy measure and is uniformly distributed over an initial range has been selected for calibration. As the parameter has little effect on the discrepancy measure, parameter values discarded will be nearly uniformly distributed over the range. Characterizations of the final set (mean, variance) will not change very much compared to the initial range. Alternatively if a parameter has a large effect on the discrepancy measure, discarded parameter values will not be randomly distributed over the range. In case of a single optimal value for that parameter, characteristics of the set (mean, variance) will change relative to the initial set. If at the end of the calibration the discrepancy measures of the parameter vectors in the remaining set are sufficiently similar, calculation of a correlation

between parameter values contained in the set is possible and meaningful. This discussion suggests that the Price algorithm yields information that allows a further selection of parameters given a set of observations and given problems of identifiability.

6.5 A point calibration procedure based on the Price algorithm

As each optimization procedure has its own problems and pitfalls, the question is how to take the different issues discussed in the previous sections into account if the Price algorithm is used.

Decisions regarding the discrepancy measure and the convergence criterion to be used were already made in Section 6.4. As in any optimization, there has to be a check on the sensitivity of the final parameter estimates for the initial settings. To check this the calibration has to be repeated for different initial starting settings. A procedure to take the possibly limited identifiability of parameters into account and to keep the parameter fixing error small has to be formulated.

Decision rules

To allow selection of parameters, decision rules have to be formulated. As a check on the reproducibility of the calibration, the consistency of the solution found should be checked by evaluating the variance of the discrepancy measure over a number of repetitions using different initial parameter sets. A rule for the number of repetitions needed to check this can not be given. If the variance in the best discrepancy measure is relatively large, the procedure is still sensitive to the initial set, and local minima may have been found. A coefficient of variation of 5% of the best discrepancy measure over the repetitions is subjectively considered to reflect a consistent solution. To meet this constraint one either has to increase the size of the initial cloud or one has to increase the total number of iterations of the Price algorithm. The values have to be established in a trial and error procedure. Another check is the convergence of the parameter values. Differences between the best parameter vectors for each repetition should generally be small, unless the estimated parameters are non-identifiable.

To diagnose possible problems of identifiability, results of the calibration in terms of parameter sets will be discussed in terms of relative shrinkage per parameter. Relative shrinkage (ζ) is the sum of squared differences of the final set F (containing N_g vectors of k calibrated parameters x_j (j=1...k)) divided by sum of the squared differences for the initial set I (of the same size as F). The differences are calculated with respect to the best parameter value in the final set (x_{Bj}):

$$\zeta_j = \left(\frac{\sum_{i=1}^{N_g} \left(x_{ij}^{(F)} - x_{Bj} \right)^2}{\sum_{i=1}^{N_g} \left(x_{ij}^{(I)} - x_{Bj} \right)^2} \right)^{\frac{1}{2}} \tag{6.6}$$

If the relative shrinkage is smaller than 1 the data contain information regarding the parameter; if it is larger than 1, the procedure is widening the search for a parameter. Search widening may occur if a number of parameter values are outside the initial set, but still within parameter bounds. The decision rule used to retain a parameter in the set of parameters to be calibrated is that the value of the relative shrinkage should be smaller or equal to 1 for all the repetitions.

The Spearman rank correlation between parameters will be used as a secondary criterion. These correlations are established on the basis of the information contained in the final sets. The decision rule which we will use is based on the sign of the correlation over the repetitions with different initial sets. The correlation will be regarded as consistent if its sign is the same over all repetitions. The decision to select rank correlation as a criterion of secondary importance is based on the argument that the Price-algorithm is a controlled random search procedure, which may generate search patterns in specific directions of the parameter domain. This may cause correlations between parameters which are an artefact of the search strategy. Analyzing consistency of the sign of the correlation will to some unknown extent compensate this effect.

Based on the above discussion the following calibration procedure is proposed and will be executed:

1) Select the group of parameters to be considered for calibration on the basis of their contribution to output uncertainty. To keep the parameter fixing error small, start with a relatively large number of parameters.
2) Repeatedly calibrate the selected number of parameters using different initial sets.
3) Analyse the consistency of the algorithm behaviour: if the coefficient of variation in the final best discrepancy measure over the repetitions is sufficiently small, and if the best parameter vectors are sufficiently close, go on with the next steps, else choose different settings for the Price algorithm (larger initial set, increase maximum number of runs).
4) Determine the relative shrinkage of each parameter; test the consistency of the shrinkage. If the shrinkage is inconsistent, and the parameter is not consistently correlated to another parameter, drop the parameter.
5) Determine the correlations between parameters; test the consistency of the correlation. If two parameters are consistently correlated, drop the parameter which contributes least to the prediction uncertainty, as established in the uncertainty analysis prior to calibration (Chapter 5).
6) Execute the calibration, fixing the dropped parameters at their nominal value and start again at step 1 of the procedure. We will refer to steps 1-5 as a *dropcycle*. In principle it should be sufficient to repeat the analysis until all selected parameters are shown to shrink consistently for a given data set.

An additionally necessary decision rule is the following:

7) If the calibration is stopped because the maximum number of runs is reached and the termination criterion is not met or not consistently met over the repetitions, discard a number of parameters on the basis of the (ranking of) shrinkage and correlation criteria and go on with calibration, to see whether the termination criterion is consistently met at the reduced number of parameters.

In this procedure two subjective decisions have to be made: the number of repetitions and the acceptable coefficient of variation in the best criterion value over the repetitions. Another, less

obviously, subjective choice is the priority of shrinking parameter ranges over correlation, and finally one has to decide whether to accept the calibration results.

6.6 Calibrating SUCROS87 and LINTUL

6.6.1 Data

In this thesis the conditions for crop production are those which yield potential production: ample water, ample nutrients and no weeds or diseases. Experiments described in literature may meet these conditions to various degrees:

- it is stated that the model assumptions are met, but the method used to check this is not described,
- it is experimentally guaranteed that the model assumptions (e.g. non-limiting moisture conditions) are met (e.g. by sprinkling), but this was not checked,
- it is experimentally guaranteed that the model assumptions are met and check analyses were executed (e.g. sprinkling combined with pF measurements).

Data were selected from published experiments where production was not moisture limited, either because of sprinkling, or because of sufficient moisture availability in that year. In the selected datasets nutrient conditions and control of pests were based on standard management procedures.

A total of six experiments for three varieties was available. An overview is given in Table 6.5. Two experiments (H1 and H5) have a very small number of observations; the experiment E3 differs from the other experiments in that yield only measurements after flowering were available. From these experiments output variables common to or similar in both models were selected for calibration. For SUCROS87 the variables development stage (DVS), temperature sum to emergence (EMERG), leaf area index (LAI) and total aboveground dry matter weight (TADRW) were calibration variables; for LINTUL DVS, EMERG, fraction of intercepted light (FINT) and TADRW were used in calibration. If not measured, FINT was estimated from LAI. FINT was not used to estimate LAI.

Table 6.5 Overview over the different datasets available for calibration. H: Wageningen Hoog, variety *Dorina*, reference I; E: Wageningse Eng, variety *LG11*, ref. II; D: Droevendaal, variety *LG11*, ref. III; S: Sinderhoeve; variety *Vivia*, refs. IV,V.

	Location			
Year	H	E	D	S
1981	H1			
1983		E3		
1985	H5		D5	S5
1986				S6

Table 6.5 continued References to experiments used in the calibration

I : Scholte, K., 1987. Relationship between cropping frequency, root rot and yield of silage maize on sandy soils. Netherlands Journal of Agricultural Science 35(1987) 473-486.

II : Sibma, L. and W. Louwerse, (unpublished report) Changes in quantity and quality in a maturing maize silage crop.

III : Louwerse, W., L. Sibma, and J. van Kleef, 1990. Crop photosynthesis, respiration and dry matter production of maize. Netherlands Journal of Agricultural Science 38 (1990) 95-108.

IV : Ouwerkerk, I. and H. Drenth, 1986. Verslag beregeningsonderzoek naar het gewas snijmais 1985. Interne mededeling nr. 439. Lelystad: PAGV.

V : Ouwerkerk, I., 1987. Resultaten van het beregeningsonderzoek in het gewas snijmais in 1986. Interne mededeling nr. 461. Lelystad: PAGV.

6.6.2 Adapting the calibration algorithm

Modifications implemented

In contrast to the uniform sampling procedure originally implemented (Stol et al., 1992) the algorithm was modified to allow to use N_g vectors from the prior distribution as determined in Chapter 4 as the initial input. This allows to incorporate all available information regarding the parameters and also to use the bounds on the parameters as hard bounds. The hard bounds on the initial sample were selected as the 1 and 99% probability bounds. For SUCROS87 the initial sets were generated from a single sample, which was then 'shuffled' (cf. Section 5.1.3) to impose the desired correlation structure; for LINTUL the initial sets were based on different random samples, and the desired correlation structure imposed on each sample (ibid.).

Choosing settings for the Price algorithm

As described in the calibration procedure the size of the initial set N_g, the normal termination criterion ε and the maximum number of runs N_s have to be chosen in such a way that the discrepancy measure value of the best parameter vector varies very little over a number of repetitions. To determine the parameters N_s and N_g a test calibration was executed. We arbitrarily chose a fixed normal termination criterion ($\varepsilon = 0.001$) and varied initial set size and maximum number of runs over the six datasets. The duration of the calibration, given these settings, was another, practical, constraint. The settings of the algorithm were considered to be acceptable if the coefficient of variation of the final best discrepancy measure values was less than 5% over six repetitions. In this procedure it was noticed that for one dataset (H1) an almost perfect fit (discrepancy measure about 0) could be achieved for the model LINTUL (cf. Table 6.6c). In that case, the coefficient of variation is not a relevant measure. The settings for the Price algorithm

were based on the remaining datasets. Considering the results in Tables 6.6a-c, and the predefined requirement to achieve a coefficient of variation of less than 5% over repetitions, an initial set size of 80 and a maximal number of runs of 7500 was chosen for LINTUL. The latter number is higher than necessary, as a maximal number of 6400 runs already gave acceptable results for most of the datasets except H1 and H5. For SUCROS87 (cf. Table 6.7) the initial set size was chosen as 130 and calibration was stopped after maximally 13000 runs. The coefficient of variation for the dataset H5 could not be reduced below 5% for a feasible number of simulations.

Table 6.6a Test calibration for LINTUL for all datasets: Coefficient of variation (%) of the discrepancy measure over 6 repetitions for a set size N_g of 40 and different values of the maximum number of simulations N_s. See Table 6.5 for references to datasets.

N_s	D5	E3	H1	H5	S5	S6
400	14.2	19.4	37.9	15.8	34.7	27.1
800	20.2	27.7	50.5	26.8	13.4	29.4
1600	20.8	44.4	54.0	35.7	9.4	26.6
3200	40.3	44.1	43.1	40.2	4.8	7.1
6400	19.7	42.9	44.3	42.6	4.2	7.2
12800	18.9	42.7	46.1	42.8	4.2	7.4

Table 6.6b Test calibration for LINTUL for all datasets: Coefficient of variation (%) of the discrepancy measure over 6 repetitions for a set size N_g of 80 and different values of the maximum number of simulations N_s.

N_s	D5	E3	H1	H5	S5	S6
400	11.2	14.7	20.9	11.1	26.0	18.8
800	12.9	9.3	18.9	11.0	22.7	15.2
1600	13.3	16.4	28.6	14.8	2.6	12.5
3200	14.3	2.2	41.7	15.5	0.2	4.4
6400	5.2	1.9	46.6	36.4	0.1	1.0
12800			42.5	16.1		
25600			74.4	12.4		

Table 6.6c Test calibration for LINTUL: Calibration criterion and coefficient of variation (%, between brackets) of the discrepancy measure over 6 repetitions for a set size N_g of 120 and different values of the maximum number of simulations N_s.

N_s	D5	E3	H1	H5	S5	S6
3200			0.3478	0.7460		
			(43.5)	(11.9)		
6400			0.1147	0.6301		
			(41.5)	(16.9)		
12800			0.0414	0.4308		
			(38.6)	(7.6)		
25600			0.0125	0.4094		
			(22.7)	(4.1)		
51200			0.0026	0.3998		
			(38.1)	(2.1)		

Table 6.7 Test calibration for SUCROS87 for all datasets: Coefficient of variation (%) of the discrepancy measure over 6 repetitions for a set size N_g of 130 and different values of the maximum number of simulations N_s. See Table 6.5 for references to datasets

N_s	D5	E3	H1	H5	S5	S6
1300	8.3	2.9	4.0	14.8	4.2	3.6
2600	9.4	2.9	3.1	18.4	6.0	2.3
5200	5.2	2.1	1.4	22.2	4.2	0.0
10400	4.3	0.1	1.1	22.7	3.2	0.0

6.6.3 Calibration of SUCROS87

The calibration of the different datasets was executed according to the procedure described above. Six initial sets with a size (N_g) of 130 parameter vectors each were generated on the basis of the distributions determined from literature.

The values of the discrepancy measure and its c.v. are presented in Table 6.8. The calibration results were checked using a Simplex algorithm (last column Table 6.8) for the parameters retained in the final dropcycle. This was used as an additional test of the convergence of the Price algorithm. The Simplex algorithm (Nelder and Mead, 1965; Press et al., 1992; Stol et al., 1992) was initialized independent from the results already achieved. No or very little improvement of the discrepancy measure for calibration using the Simplex algorithm was achieved compared to the results of the Price algorithm. The Simplex result of exactly 1 for dataset H5 requires an explanation. For the selected discrepancy measure, the worst fitting output

variable determines the discrepancy measure. In the case of emergence the observation either has the value emerged (1) - or not emerged (0). If the crop has not emerged a day later or earlier, the discrepancy measure yields a criterion value of exactly 1. This suggests that calibration of emergence determines the criterion value. The test calibration had already shown that it was necessary to include the temperature sum to emergence (TSEMER) as an additional parameter to allow Price to reduce the criterion value from 1 to smaller values; obviously the SIMPLEX procedure is not sufficiently robust.

Evaluating the results in terms of the maximum value of the relative shrinkage for the individual parameters yields the results - using the results for dataset D5 as an example - presented in Table 6.9. As an additional safety, the decision to stop was taken for all datasets simultaneously. This means that the dataset which was most difficult to fit determined the number of dropcycles (a step in the procedure described in Section 6.5). For SUCROS87 5 dropcycles were necessary.

Visual evaluation

As stated in the presentation of the decision rules, the decision to accept a calibration result remains a subjective choice. The calibration results were evaluated in terms of time courses of the calibrated output variables. The results are presented in Figures 6.5a and b. The calibration results achieved in experiments D5 and E3 are not very satisfactory.

Table 6.8 The value of the discrepancy measure and its coefficient of variation (%) for the iterative calibration of SUCROS87. For references to the datasets cf. Table 6.5.

Dataset	Dropcycle					Simplex (check)
	1	2	3	4	5	
D5	1.621	1.604	1.605	1.604	1.604	2.206
	(1.9)	(0.1)	(0.1)	(0.1)	(0.1)	
E3	2.035	2.033	2.033	2.033	2.033	2.451
	(0.4)	(0.0)	(0.0)	(0.0)	(0.0)	
H1	0.193	0.200	0.206	0.205	0.193	0.183
	(6.2)	(3.3)	(1.3)	(2.0)	(2.8)	
H5	0.398	0.546	0.546	0.546	0.546	1.000
	(22.1)	(0.0)	(0.0)	(0.0)	(0.0)	
S5	0.949	1.012	1.013	1.138	1.134	1.156
	(2.5)	(2.1)	(2.8)	(0.5)	(1.0)	
S6	1.013	1.008	1.008	1.008	1.008	1.065
	(0.2)	(0.0)	(0.0)	(0.0)	(0.0)	

Table 6.9 Maximum value of the relative shrinkage of selected parameters and selection process in the calibration of SUCROS87, for the dataset D5. Parameters for which no shrinkage is given in later dropcycles have been dropped and are fixed at their nominal value. For an explanation of the abbreviations cf. Appendix 1.

Parameter	Dropcycle				
	1	2	3	4	5
TSDVRV	0.1	0.0	0.0	0.0	0.0
EFF	1.0	0.5	0.8	0.6	0.6
BSH	1.3				
BSLA	0.6	0.3	0.6	0.6	0.5
AMX	1.1				
MAINRT	1.1				
Q10	1.9				
RGRL	0.6	0.0	0.1	0.1	0.0
TSDVRR	1.1				
FRDFBB	1.6				
DVSSEN	0.3	0.1	0.2	0.2	0.3
ASRRT	1.6				
TSEMER	0.0	0.0	0.0	0.0	0.0

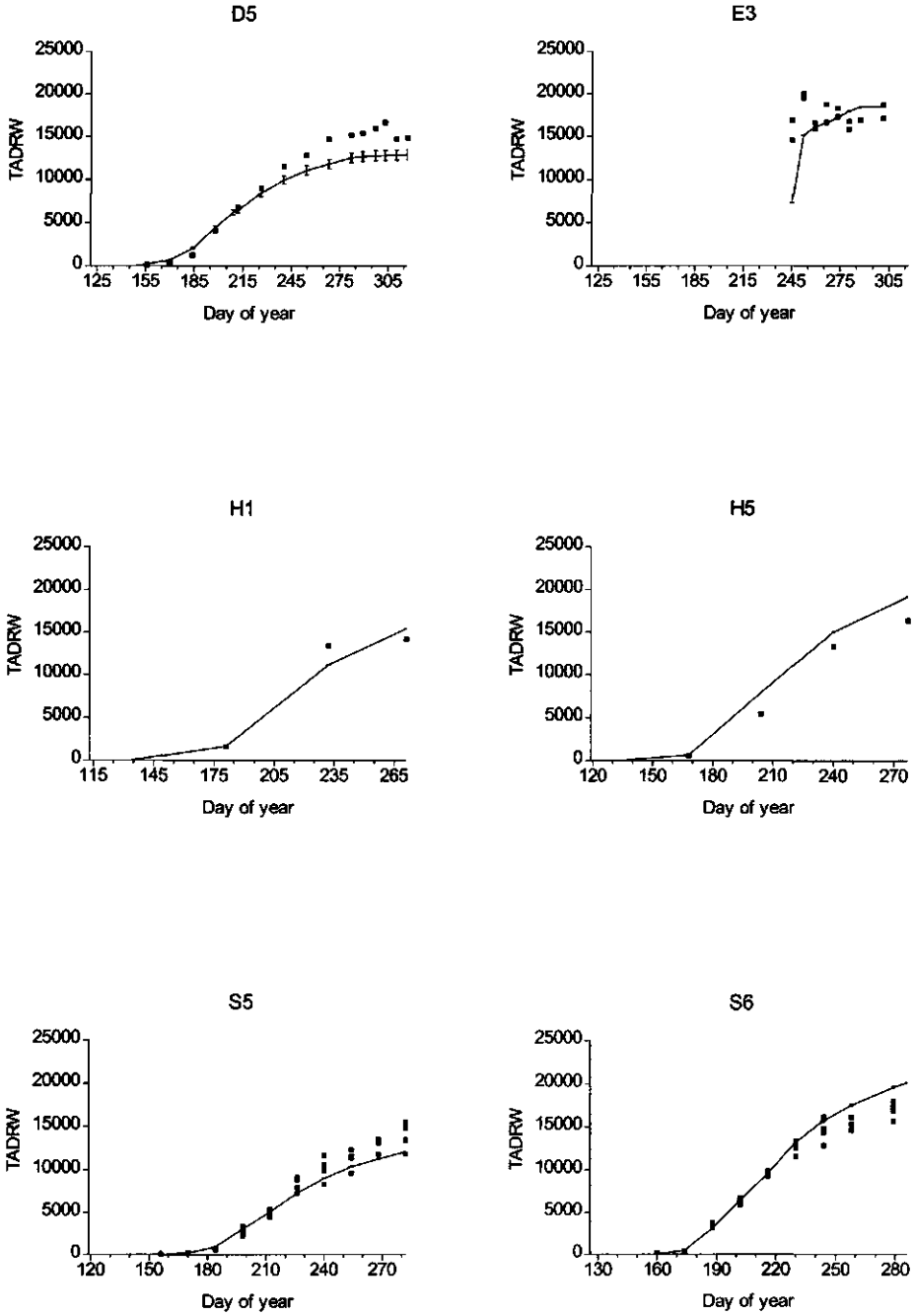

Figure 6.5a The results of the calibration of SUCROS87 (cf. Table 6.5 for references to the datasets. The fitted model output (drawn line) and measurements of total aboveground dry matter (TADRW, kg ha^{-1}) as a function of day of year for the different datasets.

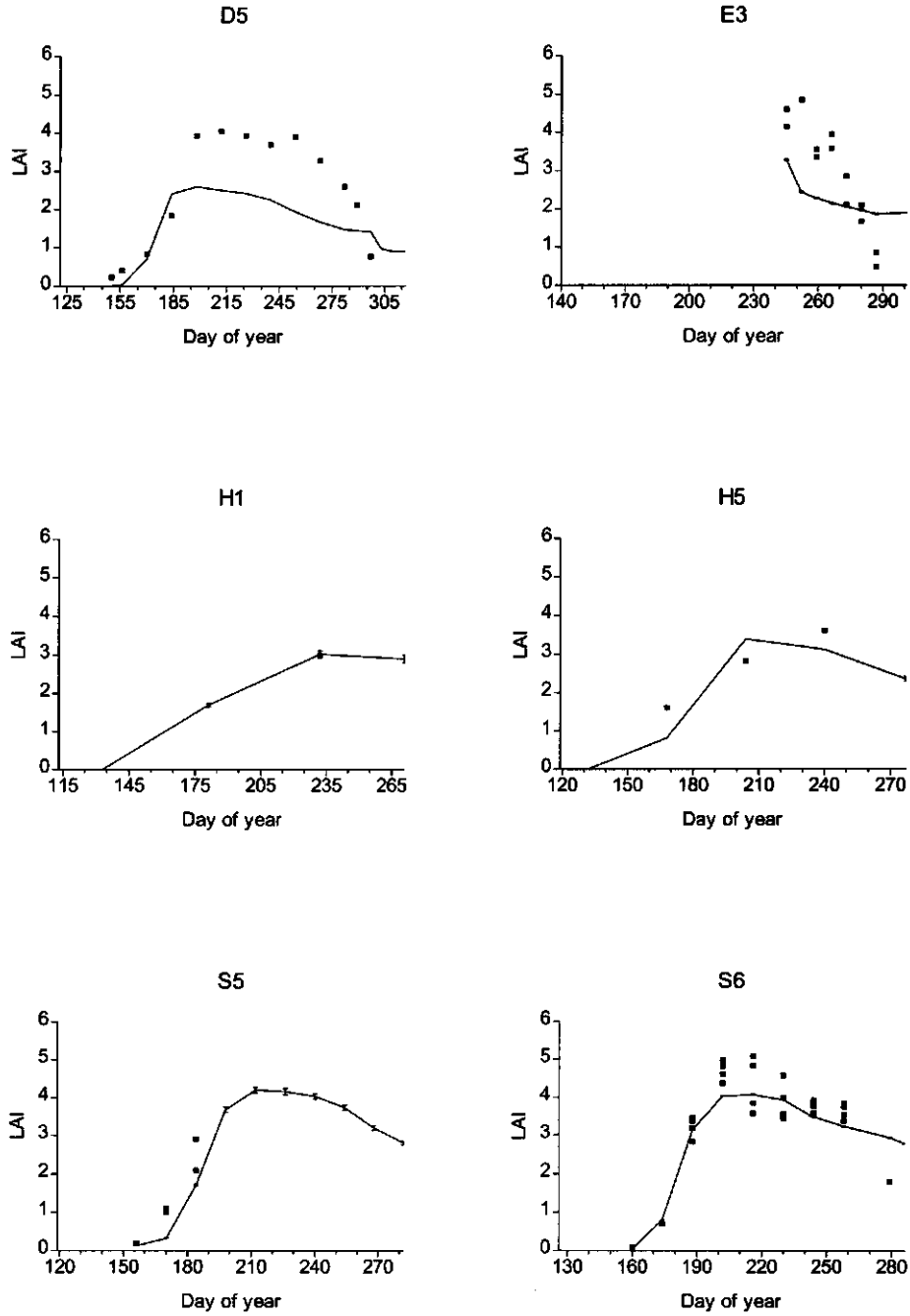

Figure 6.5b The results of the calibration of SUCROS87 for the different datasets. The fitted model output (drawn line) and measurements of leaf area index (LAI, m^2 leaf m^{-2} ground) as a function of day of year.

6.6.4 Calibration of LINTUL

The calibration of the different datasets was executed according to the procedure described in Section 6.4. Six initial sets of 80 parameter vectors each were generated on the basis of the distributions determined from literature. The values of the discrepancy measure and the c.v. over six repetitions are presented in Table 6.10. The calibration was checked repeating the calibration for the finally selected parameters using the Simplex algorithm, independently parameterized (Table 6.10). No improvement of the fit could be achieved. In fact the results show a discrepancy measure of exactly 1, which suggests that the Simplex algorithm (again) can not find a better value for TSEMER.

Evaluating the results in terms of the maximum value of the relative shrinkage for the individual parameters yields the results presented in Table 6.11, using dataset D5 as an example. Three dropcycles were necessary to finish the selection process for the datasets; in the example presented here, dataset D5, no parameters were dropped, although parameter FRPAR is close to being dropped.

Visual evaluation

The calibration results were evaluated in terms of time courses of the calibrated output variables. In the experiments discussed here, all results achieved were acceptable. Results are presented in Figure 6.7a and b.

Table 6.10 The value of the discrepancy measure and its coefficient of variation (%) for the iterative calibration of LINTUL. For references to the datasets cf. Table 6.5.

	Dropcycle			
Dataset	1	2	3	Simplex (check)
D5	0.773	0.775	0.774	0.837
	(0.5)	(0.3)	(0.4)	
E3	0.626	0.590	0.559	0.558
	(3.2)	(0.079)	(0.0)	
H1	0.006	0.002	0.001	1.000
	(191)	(127)	(59)	
H5	0.416	0.465	0.407	1.000
	(12.4)	(30.3)	(5.2)	
S5	0.776	0.777	0.776	0.780
	(0.2)	(0.2)	(0.2)	
S6	0.772	0.767	0.767	0.788
	(0.7)	(0.0)	(0.0)	

Table 6.11 Maximum value of the relative shrinkage of selected parameters and selection process in the calibration of LINTUL, for dataset D5. Parameters for which no shrinkage is given in later dropcycles have been dropped and are fixed at their nominal value (in this example none). For an explanation of the abbreviations used cf. Appendix 2.

	Dropcycle		
Parameter	1	2	3
TSDVR	0.1	0.1	0.1
FINTI	0.0	0.0	0.0
RI	0.1	0.1	0.1
LUE	0.3	0.3	0.2
DVHALF	0.1	0.0	0.1
TSEMER	0.0	0.0	0.0
DVLGTH	0.0	0.0	0.0
FRPAR	1.0	0.9	1.0

6.6.5 Parameter point estimates

As the decision to stop the calibration was taken for all datasets simultaneously, the dataset which was most difficult to fit determined the number of dropcycles (a step in the calibration procedure described in Section 6.5). After 5 dropcycles the calibration was regarded as finished for SUCROS87. During the different dropcycles about half of the parameters were dropped (Table 6.12). For LINTUL the calibration was stopped after 3 dropcycles. The number of parameters dropped from calibration was small: In 4 out of the 6 calibrations of LINTUL none of the parameters is dropped (Table 6.13).

The mean values of the parameter point estimates for the different datasets are presented in the Table 6.12 for SUCROS87 and in Table 6.13 for LINTUL. They are presented relative to their nominal values. In general, the standard errors over the repetitions of the calibration are in the order of 10% of the average for both models. Variation over datasets is much larger.

The correlations between the parameters were calculated pooling all parameter estimates from the 6 final sets N_g for each dataset. The results are presented in Table 6.14 for SUCROS87 and in Table 6.15 for LINTUL.

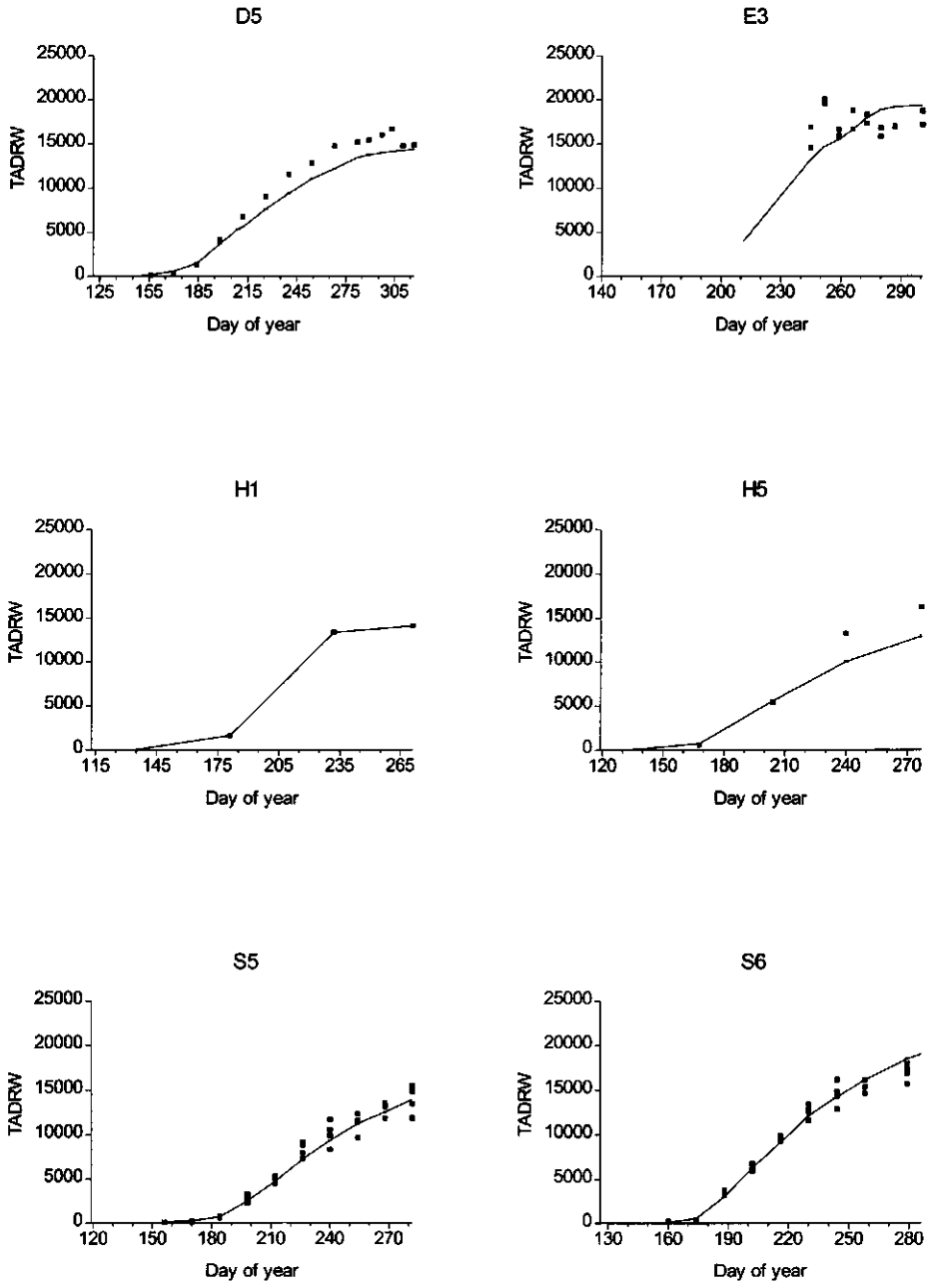

Figure 6.6a The results of the calibration of LINTUL. For references to the different datasets cf. Table 6.5. The fitted model output (drawn line) and measurements of total aboveground dry matter (TADRW, kg ha^{-1}) as a function of day of year for the different datasets.

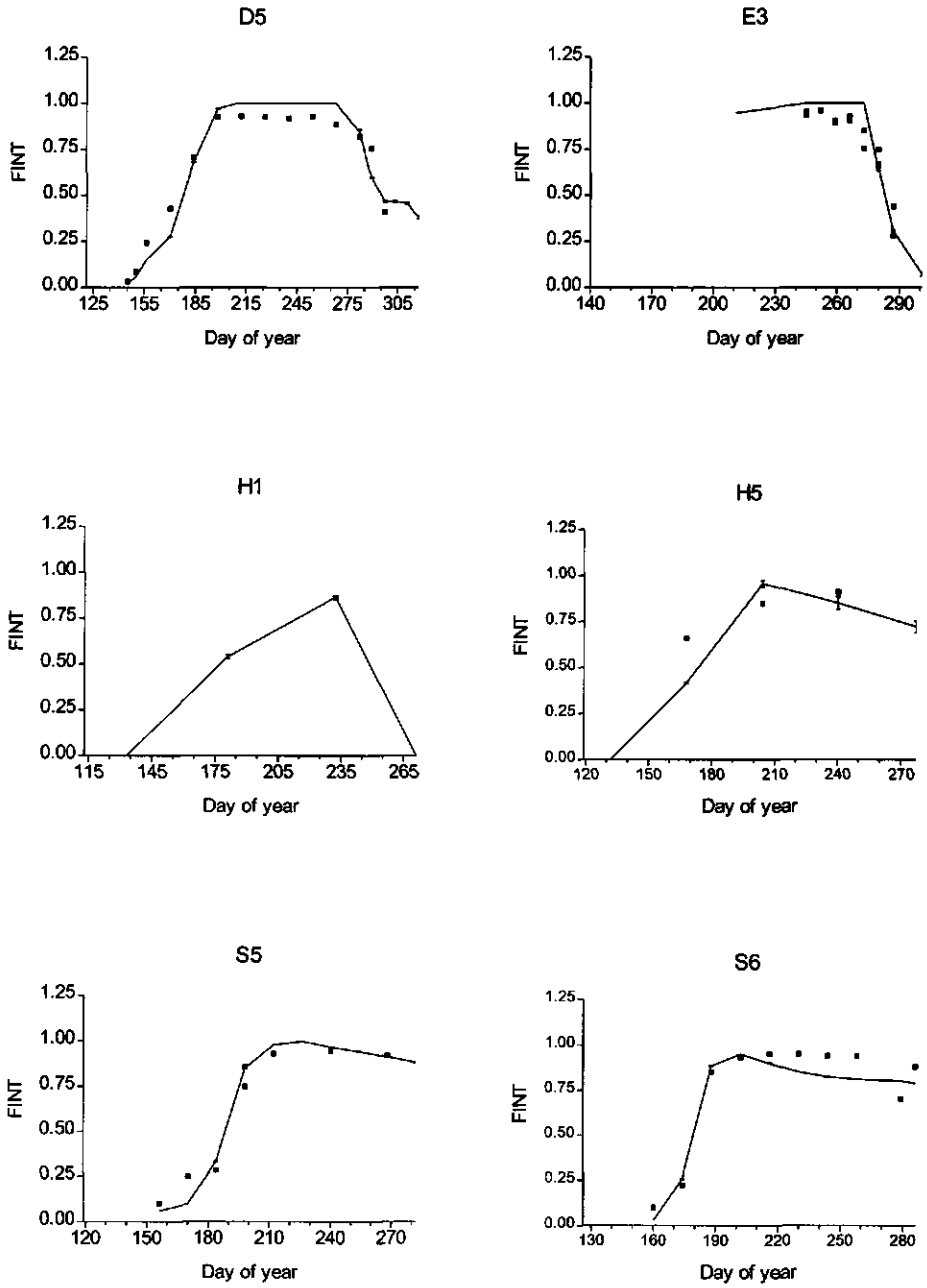

Figure 6.6b The results of the calibration of LINTUL. The fitted model output (drawn line) and measurements of the fraction intercepted light (FINT) as a function of day of year for the 6 datasets.

Table 6.12 Calibration results for SUCROS87: average of the parameter vectors yielding the best criterion value and their standard error (in brackets) over six initializations of the procedure. Estimates close or equal to the assumed bound are indicated by ∧ (high value) and ∨ (low value). Empty table cells indicate parameters dropped during the procedure. For an explanation of the abbreviations cf. Appendix 1. Parameter values are relative to nominal values as presented in Appendix 5.

	Dataset					
	D5	E3	H1	H5	S5	S6
TSDVRV	0.82	0.91	0.96	0.82	0.80	1.12
	(0.00)	(0.00)	(0.01)	(0.00)	(0.09)	(0.00)
EFF	0.73	1.09∧	1.06	1.09∧		0.88
	(0.11)	(0.00)	(0.03)	(0.00)		(0.02)
BSH		3.21∧				
		(0.14)				
BSLA	1.09		0.90		1.51	1.13
	(0.13)		(0.06)		(0.07)	(0.02)
AMX		1.48	1.19		0.49	
		(0.03)	(0.22)		(0.09)	
MAINRT			2.37		0.41	
			(0.93)		(0.01)	
Q10						2.46
						(0.24)
RGRL	1.03∧		0.89			0.91
	(0.00)		(0.01)			(0.00)
TSDVRR				1.78	1.00	0.68
				(0.21)	(0.01)	(0.00)
FRDFBB						
DVSSEN	1.29		1.69			0.80
	(0.03)		(0.01)			(0.00)
ASRRT						
TSEMER	0.16∨		0.45	0.22	0.43	0.16∨
	(0.02)		(0.03)	(0.01)	(0.03)	(0.00)

Table 6.13 Calibration results for LINTUL: average of the parameter vectors yielding the best criterion value and their standard error (in brackets) over six initializations of the procedure. Estimates close or equal to the assumed bound are indicated by ∧ (high value) and ∨ (low value). Empty table cells indicate parameters dropped during the procedure. For an explanation of the abbrevations cf. Appendix 2. Parameter values are expressed relative to nominal values as presented in Appendix 6.

	Dataset					
	D5	E3	H1	H5	S5	S6
TSDVR	1.35	0.98	1.13	0.93	1.49	2.02
	(0.15)	(0.00)	(0.01)	(0.05)	(0.13)	(0.04)
FINTI	5.88		3.7	1.91	5.76	6.94∧
	(0.50)		(1.07)	(0.33)	(1.06)	(0.01)
RI	1.48		1.27	1.53	1.41	1.82
	(0.17)		(0.09)	(0.06)	(0.12)	(0.04)
LUE	1.15	1.94	2.15	1.15	1.39	1.56
	(0.04)	(0.00)	(0.02)	(0.06)	(0.04)	(0.01)
DVHALF	0.71	1.04	0.61	1.42	1.37	0.77
	(0.08)	(0.00)	(0.01)	(0.20)	(0.09)	(0.01)
TSEMER	0.22		0.45	0.22	0.44	0.16∨
	(0.03)		(0.01)	(0.01)	(0.08)	(0.01)
DVLGTH	0.03∨	0.03∨	0.03∨	1.65	1.64	
	(0.00)	(0.00)	(0.00)	(0.44)	(0.11)	
FRPAR	0.92		0.89	0.93	0.94	
	(0.03)		(0.01)	(0.04)	(0.02)	

Table 6.14 Correlation between parameter vectors pooled over final sets for each dataset: Results for SUCROS87, presented in the order D5-E3-H1-H5-S5-S6. * indicate parameter(s) dropped from calibration. For an explanation of the abbreviations cf. Appendix 1.

	TSDVRV	EFF	BSH	BSLA	AMX	MAINRT	Q10
TSDVRV	1						
EFF	0.38	1					
	0.12						
	−0.04						
	−0.25						
	*						
	0.12						
BSH	*	*	1				
	−0.14	−0.36					
	*	*					
	*	*					
	*	*					
	*	*					
BSLA	−0.47	−0.99	*	1			
	*	*	*				
	−0.41	−0.07	*				
	*	*	*				
	−0.93	*	*				
	−0.28	−0.93	*				
AMX	*	*	*	*	1		
	−0.08	−0.68	−0.25	*			
	0.51	−0.02	*	−0.85			
	*	*	*	*			
	−0.98	*	*	0.88			
	*	*	*	*			
MAINRT	*	*	*	*	*	1	
	*	*	*	*	*		
	0.38	0.57	*	−0.49	0.67		
	*	*	*	*	*		
	0.02	*	*	−0.01	0.05		
	*	*	*	*	*		
Q10	*	*	*	*	*	*	1
	*	*	*	*	*	*	
	*	*	*	*	*	*	
	*	*	*	*	*	*	
	*	*	*	*	*	*	
	0.06	−0.62	*	0.41	*	*	

Table 6.14 continued Correlation between parameter vectors pooled over final sets for each dataset: Results for SUCROS87, presented in the order D5-E3-H1-H5-S5-S6. * indicate parameter(s) dropped from calibration.

	TSDVRV	EFF	BSH	BSLA	AMX	MAINRT	Q10
RGRL	0.41	0.01	*	−0.07	*	*	*
	*	*	*	*	*	*	*
	−0.20	−0.41	*	0.31	−0.61	−0.85	*
	*	*	*	*	*	*	*
	*	*	*	*	*	*	*
	0.10	−0.15	*	−0.11	*	*	0.16
TSDVRR	*	*	*	*	*	*	*
	*	*	*	*	*	*	*
	*	*	*	*	*	*	*
	−0.10	−0.01	*	*	*	*	*
	0.96	*	*	−0.98	−0.95	−0.05	*
	−0.34	0.08	*	−0.03	*	*	−0.10
FRDFBB	*	*	*	*	*	*	*
	*	*	*	*	*	*	*
	*	*	*	*	*	*	*
	*	*	*	*	*	*	*
	*	*	*	*	*	*	*
	*	*	*	*	*	*	*
DVSSEN	−0.20	−0.11	*	0.12	*	*	*
	*	*	*	*	*	*	*
	−0.23	−0.20	*	0.07	−0.07	−0.15	*
	*	*	*	*	*	*	*
	*	*	*	*	*	*	*
	0.06	−0.16	*	0.15	*	*	−0.02
ASRRT	*	*	*	*	*	*	*
	*	*	*	*	*	*	*
	*	*	*	*	*	*	*
	*	*	*	*	*	*	*
	*	*	*	*	*	*	*
	*	*	*	*	*	*	*
TSEMER	−0.04	0.01	*	−0.01	*	*	*
	*	*	*	*	*	*	*
	−0.01	0.13	*	−0.09	−0.01	0.03	*
	−0.02	0.01	*	*	*	*	*
	0.02	*	*	−0.01	−0.02	−0.06	*
	0.05	−0.01	*	−0.01	*	*	−0.01

Table 6.14 continued Correlation between parameter vectors pooled over final sets for each dataset: Results for SUCROS87, presented in the order D5-E3-H1-H5-S5-S6. * indicate parameter(s) dropped from calibration.

	RGRL	TSDVRR	FRDFBB	DVSSEN	ASRRT	TSEMER
RGRL	1					
TSDVRR	*	1				
	*					
	*					
	*					
	*					
	−0.01					
FRDFBB	*	*	1			
	*	*				
	*	*				
	*	*				
	*	*				
	*	*				
DVSSEN	−0.17	*	*	1		
	*	*	*			
	−0.06	*	*			
	*	*	*			
	*	*	*			
	0.08	−0.07	*			
ASRRT	*	*	*	*	1	
	*	*	*	*		
	*	*	*	*		
	*	*	*	*		
	*	*	*	*		
	*	*	*	*		
TSEMER	0.05	*	*	0.12	*	1
	*	*	*	*	*	
	0.03	*	*	−0.01	*	
	*	0.06	*	*	*	
	*	0.01	*	*	*	
	0.09	−0.06	*	0.03	*	

Table 6.15 Correlation between parameter vectors pooled over final sets for each dataset: Results for LINTUL, presented in the order D5-E3-H1-H5-S5-S6. * indicate parameter(s) dropped from calibration. For an explnation of the abbreviations cf. Appendix 2.

	TSDVR	FINTI	RI	LUE	DVHALF	TSEMER	DVLGTH	FRPAR
TSDVR	1							
FINTI	−0.65							
	*							
	−0.25							
	0.42							
	0.19							
	−0.22							
RI	0.99	−0.71	1					
	*	*						
	0.52	−0.94						
	0.94	0.14						
	1.00	0.15						
	1.00	−0.24						
LUE	0.14	−0.05	0.13	1				
	0.26	*	*					
	−0.02	0.52	−0.46					
	0.42	0.12	0.45					
	0.02	0.03	0.02					
	0.33	−0.10	0.28					
DVHALF	−1.00	0.64	−0.99	−0.14				
	−0.99	*	*	−0.27				
	−1.00	0.25	0.25	0.02				
	−0.46	−0.13	−0.13	−0.39				
	−0.55	−0.20	−0.20	0.01				
	−0.99	0.21	0.21	−0.42				
TSEMER	−0.69	0.90	−0.71	−0.10	0.68	1		
	*	*	*	*	*			
	0.11	−0.07	0.06	−0.08	−0.11			
	0.03	0.15	−0.01	−0.03	−0.00			
	0.10	0.87	0.08	0.03	−0.13			
	−0.09	0.06	−0.09	0.03	0.08			
DVLGTH	−0.47	0.20	−0.45	−0.30	0.47	0.22	1	
	0.07	*	*	0.15	−0.16	*		
	−0.42	0.03	−0.20	0.03	0.43	−0.05		
	0.20	−0.09	0.22	−0.13	0.65	0.01		
	−0.30	−0.12	−0.30	0.04	0.95	−0.06		
	*	*	*	*	*	*		
FRPAR	−0.17	0.08	−0.16	−0.99	0.16	0.12	0.30	1
	*	*	*	*	*	*	*	
	−0.20	0.36	−0.36	−0.47	0.20	0.04	−0.03	
	0.19	0.06	0.17	−0.62	0.20	0.10	0.43	
	0.00	0.05	0.00	−0.99	−0.01	0.03	−0.03	
	*	*	*	*	*	*	*	

6.6.6 Discussion

Calibration of SUCROS87

The results presented in Table 6.12 show the parameter estimates (as averages over 6 repetitions) for SUCROS87. We will refer to the parameter estimates using subscript E. Tables 6.12 and 6.14, and Figures 6.6a and b are used in the discussion for SUCROS87. The parameters which have been estimated in multiple datasets differ strongly. However, there are no characteristics in the calibration results which are common to at least some datasets. The calibration results are therefore discussed separately for each calibration set.

The calibration of SUCROS87 using dataset D5. The results are not satisfactory, as LAI remains too low. $RGRL_E$ is at its upper bound; $TSEMER_E$ is at its lower bound, the estimates of $BSLA_E$ and EFF_E are negatively correlated, and are relatively imprecise compared to the other estimates. This suggests that a problem of identifiability exists for BSLA and EFF.

To improve the fit the LAI has to become higher. A higher LAI can not be achieved by increasing RGRL, as $RGRL_E$ has reached its limit. If we accept the ranges for $RGRL_E$, this implies that other parameters used in the calculation of juvenile leaf area index have to be modified. There are two options: increase the duration of the juvenile growth phase, or lower the base temperature for leaf area index growth, which in this calibration equals 10 °C. Additionally, parameters determining leaf senescence will have to be adapted, to allow a steeper decrease.

The calibration of SUCROS87 using dataset E3. This dataset is problematical, because there is no information prior to flowering. There is a systematic error in fitting LAI in the senescing phase. Production is high compared to the other datasets which may explain why both EFF_E and BSH_E are at their upper bound, while AMX_E is 50% higher than nominal. Correlations are not exceptionally high.

If we assume the ranges of the parameters to be correct, a possible procedure to improve the fit is similar to the one proposed for dataset D5. Given that RGRL is dropped from the calibration, leaf growth should probably be increased using partitioning to leaves. Furthermore the senescence parameters should be adapted, which should allow lower values for AMX, EFF, and perhaps allows to drop BSH from the calibration, as its inclusion is exceptional.

The calibration of SUCROS87 using dataset H1. This is a very small dataset, which given the large number of parameters, is difficult to fit. It is curious that the number of parameters retained is relatively high, whereas the number of high correlations is limited to the pairs $BSLA_E$-AMX_E, and $RGRL_E$-$MAINRT_E$. The high correlation between $RGRL_E$ and $MAINRT_E$ is unexpected.

The calibration of SUCROS87 using dataset H5. This is a very small dataset, which given the large number of parameters, is difficult to fit. In comparison to dataset H1 more parameters are dropped. Both $TSDVRR_E$ and EFF_E are high, but the fit does not allow to hypothesize why.

The calibration of SUCROS87 using dataset S5. The fit underestimates LAI and is in the low ranges of TADRW, as for dataset D5. The suggestion to improve the fit is therefore similar:

include parameters which would allow to increase LAI in early stages of growth.

A number of parameter estimates are strongly correlated: $TSDVRV_E$-$BSLA_E$-AMX_E-$TSDVRR_E$ with absolute values of the correlations not below 0.88. In this dataset no phenological observations are available. This leaves more possibilities for the values of TSDVRV and TSDVRR. They clearly play a role in the fit, given the correlations with the other parameters. However, it is difficult to understand their effects given the non-monotonous dependence of yield on these two parameters. The high value of $BSLA_E$ (1.5) is positively correlated with the low value of AMX_E (0.5), in contrast to the negative correlation in dataset H1.

The calibration of SUCROS87 using dataset S6. The fit is satisfactory. The results show that $BSLA_E$ and EFF_E are negatively correlated. The results of the calibrations show that combining a parameter determining photosynthesis (AMX or EFF) and the parameter determining partitioning to leaf area yields correlated estimates.

The high value of $Q10_E$ suggests that maintenance respiration is decreased (for average temperatures lower than 25 °C). The negative correlation with EFF_E shows that Q10 and EFF_E are difficult to estimate as a pair.

Calibration of LINTUL

Whereas the quality of the fit to the data is subject to discussion for SUCROS87, the fit is in general satisfactory for LINTUL. In contrast to SUCROS87, the overview of the parameter estimates and their correlations shows results which are common to a number of datasets. The discussion will therefore focus on groups of parameters, and not on individual datasets as in the discussion for SUCROS87. We will refer to the parameter estimates using subscript E. Tables 6.13 and 6.15, and Figures 6.7a and b are used in the discussion of the calibration results for LINTUL. As in the calibration of SUCROS87, parameter estimates common to different datasets vary widely.

The estimates of TSDVR and kI are strongly correlated in datasets D5, H5, S5 and S6. The parameter TSDVR is used in the calculation of the development stage (DVS); RI is used in the calculation of the reduction function of the light use efficiency (FINT, a logistic function of DVS). DVS can be calculated as:

$$DVS = \frac{1}{TSDVR}\sum_{i=1}^{n}\max(0, TAVG_i - TBASE) \qquad (6.7)$$

for n integration steps. The product RI×DVS used in the calculation of FINT can be rewritten as C×Σ( ), where C=RI/TSDVR. As a consequence estimates of RI and TSDVR are linearly related for the same value of the temperature sum. In the calibration this shows as a positive correlation. These identifiability problems explain the relatively high values of $TSDVR_E$ in dataset S5, for which phenology data are not available. The high value of $TSDVR_E$ in dataset S6 (twice nominal) and in dataset D5 could at least partly be a consequence of this part of the model formulation.

LUE_E and $FRPAR_E$ are negatively correlated in datasets D5, H5 and S5. These parameters

are only used as the product FRPAR×LUE, which is used in the calculation of the daily growth rate. Consequently, for the same growth rate, the product of these two parameters is constant, and its estimates are therefore negatively correlated, and ill-determined individually.

A strong negative correlation between $DVHALF_E$ and $TSDVR_E$ also exists in datasets D5, E3, H1 and S6. $DVLGTH_E$ and $DVHALF_E$ are positively correlated in datasets H5 and S5. The auxiliary function in which these parameters are used, is:

$$FINTS = 0.5 - \frac{(DVS - DVHALF)}{DVLGTH} \tag{6.8}$$

which can rewritten as:

$$FINTS = 0.5 + \frac{DVHALF}{DVLGTH}\left(1 - \frac{\sum_{i=1}^{n}\max(0, TAVG_i - TBASE)}{DVHALF \times TSDVR}\right) \tag{6.9}$$

In this form the equation shows that for a constant ratio DVHALF/DVLGTH and a constant product DVHALF×TSDVR, FINTS remains the same. In Table 6.15 this shows as a positive correlation between $DVLGTH_E$ and $DVHALF_E$, and as a negative correlation between $DVHALF_E$ and $TSDVR_E$. Values for $DVLGTH_E$ are at their lower bound in datasets D5, E3 and H1; this means that leaf senescence is almost instantaneous on the DVS-scale.

Another positive correlation, that between $TSEMER_E$, the temperature sum to emergence, and $FINTI_E$, the light interception at emergence, can be explained qualitatively. If the crop emerges too late, because TSEMER is too high, the fit can be improved by increasing the initial light interception. This correlation is high in dataset D5 and dataset S5. This correlation is to be expected if emergence itself has not been observed as in these datasets. Other effects may dominate: the result in dataset S6, where both $TSEMER_E$ and $FINTI_E$ have stopped close to their lower ($TSEMER_E$) resp. upper bound ($FINTI_E$) may be the effect of a functional correlation. If total dry matter is to be kept high, TSEMER should be low, and FINTI should be high.

6.6.7 Conclusion

On the basis of a qualitative discussion of options and issues in calibration (Sections 6.2-6.5) a calibration procedure was formulated. In this procedure calibration was repeated to ensure that the parameter estimates are reproducible for different starting points (restarts). The restarts were regarded as reproducible if the coefficient of variation of the best value of the discrepancy measure was less than 5%. To achieve this target, the initial set size for the algorithm used had to be chosen as 10 times the number of parameters; the maximum number of runs had to be chosen as 1000 times the number of parameters (Section 6.5.2). In the actual execution of the calibration the target was achieved most of the times.

The restarts of the calibration procedure were also used as a method to drop parameters from

the calibration. The criterion to drop parameters from calibration was based on the decrease in their variation after calibration compared to that prior to calibration, with correlation between parameter estimates as a secondary criterion. This procedure lead to dropping about half the parameters originally selected for SUCROS87; for LINTUL in 4 out of 6 cases none of the parameters were dropped.

The results suggest that the quality of the calibration of SUCROS87 may be an effect of non-available information. The fit of simulated leaf area index is not satisfactory. A number of parameters determining LAI, juvenile leaf area index growth and senescence parameters could not be retrieved from literature and were therefore not considered for calibration. The quality of the fit was not a problem for LINTUL. The calibration results for SUCROS87 show problems of identifiability in the combination of a photosynthesis parameter (AMX or EFF) and the parameter BSLA, which determines partitioning to leaf area. The results show LINTUL to contain a relatively large number of correlated parameter pairs. Analysis of the source code shows that these correlations are due to the formulations chosen.

The procedure was formulated with the intent to exclude non-identifiable parameters from the calibration. Given this intent, the conclusion from the previous discussions is that it failed in doing so. This is due to the decision rule that pairs of correlated parameters were retained if shrinkage was consistent. However, for both models problems of parameter identifiability are recognized using the procedure proposed. A more straightforward calibration procedure can be based on the condition that a point estimate is only possible for parameters which are not correlated. In that case, relative shrinkage can be discarded as a criterion, leaving rank correlation as the only criterion on which to base parameter selection.

Chapter 7

Validation

7.1 Introduction

In this chapter the model performance will be re-evaluated for the calibrated models. Generally, the type of evaluation to be executed depends on the objective for which the model is developed or used. This dependence on objective may explain why there is no standard terminology and no standard protocol (Hamilton, 1991; Landry and Oral, 1993; Rykiel Jr., 1996): the number of possible options for an evaluation procedure is large.

In this case study, in which the models are to be used for prediction, models are evaluated in terms of predictive quality. If historical data are available, model predictions using the relevant inputs can be compared to these data. Instead of forecasting, one tries to reproduce (hindcasts) known data. This comparison allows to establish an empirical estimate of the prediction error. This prediction error reflects the joint effects of different sources of error: parameter errors (fixing parameters at possibly incorrect values), structural errors, programming errors and input errors (cf. Chapter 1).

The process of comparing model predictions with observations is known as validation (e.g. Sage, 1987), although terminology is certainly not standardized, and the connotation 'valid' is rather unhappy (Konikow and Bredehoeft, 1992; Bredehoeft and Konikow, 1993; Bair, 1994). Validation results allow to characterize the usefulness of a model for a specific application, where the objective is to predict (Sargent, 1984). As discussed (cf. Chapter 1), validation results may lead to a decision whether or not to use a model for the particular application.

Not all models of systems are easily validated. Systems can be characterized in terms of their timescale, their complexity and their spatial extension (de Wit, 1993). In relation to human capabilities, e.g. the degree of organization required to analyse a specific system, these characteristics determine the ease of model validation.

Systems with a small timescale, a small spatial extension and of small complexity can be isolated and repeated: models of these systems are accessible to validation. For systems with increasing time- and spatial scales and of increasing complexity, validation of system models requires increasingly more effort. This increase in effort can be illustrated if one considers validation requirements in terms of data, manpower, and organizational continuity for a model of an annual crop, for a model describing a crop rotation experiment, and validation requirements for a model to analyse sustainable agricultural practices. At some point the effort required to gather validation data is no longer feasible, and validation becomes dependent on natural recurrences (de Wit, 1993). For a unique (non-recurring) system finally, validation can only be based on data-splitting.

The discussion of validation in this chapter focuses on a repeatable system: an annual crop in a field. We will attempt to establish the usefulness of the simulation models for the prediction of final yield.

7.2 Concepts in validation

Validation will be defined as a procedure to establish the usefulness of a model for an application. We will define usefulness in terms of predictive quality relative to alternative predictors.

Validation covers only one of the many aspects which should be input to model evaluation; other aspects are e.g. documentation quality, software quality.

To establish predictive quality, one has to assume that model output and the observation to be predicted are in some way related. In a non-predictive context these links between experiment and observation on one hand and model output on the other hand are known as correspondence rules (Nagel, 1971). In the context of prediction (and in this thesis), the anticipated correspondence rule will be called a link hypothesis.

Given observations and model predictions of these observations, a link hypothesis serves to establish predictive quality. Predictive quality relative to that of alternative predictors, or relative to stated prediction objectives, determines model usefulness.

In this section we will analyse 1) possible link hypotheses, 2) possible measures of link quality, given the link hypothesis, and 3) discuss possible alternative predictors which are to be used for reference purposes. We assume that a number of validation data are available, giving rise to n pairs of measurements M and corresponding predictions P. We will refer to these measurements as the validation set.

7.2.1 Link hypotheses

In prediction the link hypothesis, the hypothesis regarding the connection between measurements and the output of a model, is often the identity relation: measured values (M) equal the value of the corresponding model predictions (P), or $M = P$, where both M and P refer to the same entity and the correspondence is exact within the margins allowed by measurement errors. Nothing subtle about that, and moreover a natural choice as the potential equality of measurements and simulations is the starting point in calibration. As in calibration the potential equality between M and P has to be ensured by the experimental set-up, guaranteeing - in our case - potential crop production.

However, as we discussed, model evaluation depends on the objective. The question we should ask is whether, for the same objective (prediction), all applications require that $M = P$ in the above sense. For some applications prediction of absolute values is not essential. A remark that is made in simulation studies (e.g. Smit, 1996 (page 192)) is the following: "Even if the model does not predict the absolute values very well, relative results are still useful". This indicates the possibility that the model is not useful for one type of application (prediction of absolute values), whereas it is still acceptable for another application (prediction of relative values).

Moreover, we may also ask ourselves whether we really expect the model predictions to be equal to the observations in all applications. It is entirely possible that a model is used for predictions out of its application domain, e.g. a model for potential production is used to predict yields achieved under non-potential conditions. This use of a model may be simply dismissed as wrong, but a more constructive analysis is to assume that in such applications the identity relation between pairs of M and P can not be valid and to consider to what extent the model predictions are still useful. In these cases the identity link hypothesis should be excluded, and other link hypotheses have to be discussed. We assume therefore that link hypotheses are application-

dependent. Given that different applications are possible, different link hypotheses must be formulated.

The identity link hypothesis is attractive, because of its simplicity and because no additional arguments and parameters are used. Ideally, alternative identity link hypotheses retain these properties. Some possible identity link hypotheses are presented in Table 7.1. The first three link hypotheses (ratio, differences and relative differences) are based on the assumption that measurements and predictions are linearly related (as $M_i = BP_i$ (ratio), $M_i = A+P_i$ (difference) and $M_i = A+BP_i$ (relative difference)).

Kleijnen (1995) suggests that for some applications it may not even be necessary that measurements and predictions are linearly related; it may be sufficient if each prediction is uniquely associated with a measurement and v.v., allowing e.g. for monotonous non-linear relationships. In that case the associated link hypothesis is that the ranking of measurements and predictions are identical. An example in which such a link hypothesis is used is presented by Palanisamy et al. (1993). In their simulation study the ranking of cultivars on the basis of yield had to be predicted correctly.

Table 7.1 Overview over different link hypotheses, where M_i is the measurement in instance i and P_i the associated prediction. M_x and M_n are the minimum and maximum measurement; P_x and P_n the associated predictions; *Rank* is the rank transformation of the measurements and the predictions.

Standard link hypothesis	
identity link hypothesis	$M_i = P_i$
Alternative identity link hypotheses	
identity of ratios	e.g. $M_i / M_x = P_i / P_x$
identity of differences	e.g. $M_i - M_n = P_i - P_n$
identity of relative differences	e.g. $(M_i - M_n)/(M_x - M_n) = (P_i - P_n)/(P_x - P_n)$
identity link hypothesis based on monotonicity	$Rank\,(M_i) = Rank\,(P_i)$

7.2.2 Link quality

Given a link hypothesis, the quality of the predictions is preferably summarized in a single measure. We will refer to this measure as the link quality. The link quality is the predictive quality given a link hypothesis. The measures for link quality are comparable to discrepancy measures used in calibration. There is no lack of proposals for measures of link quality (e.g. Aitken, 1973; Green and Stephenson, 1986; Loague and Green, 1991); the problem is to decide whether the proposed measure is relevant for an application considered.

Absolute measures of link quality

In evaluating predictive quality one measure which should be presented is the overall prediction error. Measures characterizing overall prediction error are also known as prediction accuracy. Given the identity link hypothesis, the norms used in calibration and their variants (L_1, L_2 and L_∞, cf. Chapter 6) can be used to summarize these aspects of predictive quality. The sum of squared prediction errors (L_2) is often used to characterize the prediction quality. For the identity link hypothesis it is calculated as:

$$L_2 = \sum_{i=1}^{n} (M_i - P_i)^2 \tag{7.1}$$

for the n measurement-prediction pairs of the validation set, which may consist of e.g. measurements of final crop yield.

Dividing this sum (the sum of squared prediction errors) by n yields the mean squared prediction error (*MSEP*); taking the root yields the root mean squared prediction error (*RMSEP*). Modifications of this criterion to accommodate alternative link hypotheses are not entirely straightforward: whereas the number of prediction-observation pairs is n for the identity link hypothesis, and also n for the link hypotheses as presented (using M_x and M_n), using coefficients estimated from the data will modify this number (to $n-1$ for the difference and the ratio link hypothesis, and $n-2$ for the relative difference).

The measurement-prediction pairs may also refer to data within a growing season. In these cases in which the different data are not comparable, calculating the sum of squared relative prediction errors, as e.g. proposed as calibration criterion in Chapter 6, is more relevant. Extensions to the validation of multivariable predictions may be based on criteria proposed for calibration (Chapter 6).

Model usefulness: reference values for link quality

To assess whether a model is more or less useful for prediction, it is necessary to have reference values to compare the link quality to. Reference values (or benchmarks) are the prediction errors of easily available alternative predictors (e.g. Colson et al., 1995). One has to decide which alternative predictors are the most relevant for the application considered. We will discuss possible benchmarks for two cases: the prediction of final crop yield over a number of seasons, and the prediction of the growth of a crop over the growing season.

For the prediction of final yield for a validation set of n seasons, a possible reference value is the 'average leave-one-out prediction error'. To calculate this error one value of the validation set is regarded as not available. The mean calculated over the remaining $n-1$ data is used to 'predict' the non-available value. The average error made when repeating this n times is the benchmark prediction error to which the model prediction error can be compared. The comparison describes the efficiency of the model as a method to replace missing values as compared to taking the mean of the available ($n-1$) values. As in our case study the simulation models incorporate the effect of weather and the benchmark predictor does not, a simulation

model that yields a larger prediction error than the leave-one-out prediction error is not very successful.

Other reference prediction errors may also be relevant. Yield in one year at one location may be used as a predictor for yield in the next year at the same location. This is a zero order forecast (Sorooshian, 1991), and one would certainly expect the model to be better. Yield in the same year at a different location (not a measurement replicate) may be used as a predictor for the location of interest. This reference site predictor may in some cases be better than the model prediction as a crop is a more complete 'model' of the same crop in another field than the simulation model.

In the case of analysing model performance within the growing season based on time series, similar benchmark predictors are possible. After the growing season one may determine an average relative leave-one-out prediction error, calculated e.g. from a linear interpolation based on the two nearest measurements. This allows to evaluate the model as a method to replace missing values within a growing season in comparison to some other, simpler interpolation method. On-line forecasts within a growing season are also possible, but are outside the scope of this thesis.

Finally, without measurement error estimates based on replicated measurements, prediction error can not be partitioned into model error and measurement error (e.g. Vereecken et al., 1991; ten Berge et al., 1995). If measurement errors are not available, assuming a negligible measurement error will result in the conclusion that the model is faulty, whereas the assumption of a negligible model error will suggest the prediction error to be entirely due to measurement error. The measurement error (which the model does not allow for by definition) then is a natural reference prediction error. Measurement error in e.g. yield depends on the spatial scale, introducing questions of representativity and sample size. This should be taken into account whenever prediction errors are compared to measurement errors.

An alternative to the on-site measurement error calculated from on-site replicates is the variation in measured yield between different locations. Simulation results do not differ between production instances if the inputs (for potential production the associated meterological station and management data) are identical.

In analogy to parameter uncertainty (cf. Chapter 4) the variation between yield measured in production instances which are identical to the model, but are geographically distinct, will be called yield uncertainty. Yet, it may at the same time still be interpreted as the performance of a reference predictor. If interpreted as a reference predictor, one may conclude that the model predicts better or worse; if interpreted as yield uncertainty, it indicates a bound past which predictions can be improved if meterological data are available on a more detailed scale.

Aspects of link quality

If sufficient validation data are available, a plot of measurements versus predictions may suggest that part of the prediction error is systematic. The systematic error is also known as prediction bias. Non-systematic errors, random prediction errors, are known as prediction precision.

Three types of measures then characterize aspects of predictive quality: measures characterizing the overall prediction error (accuracy), measures characterizing the systematic prediction errors (bias) and measures characterizing the random prediction errors (precision).

The overall prediction error, the sum of squared prediction errors (the L_2 norm), can be decomposed in different ways to yield measures characterizing bias and precision (Theil, 1961; Power, 1993; Smith and Rose, 1995). A decomposition closely associated to the evaluation of prediction results in terms of plots of measurements M vs. predictions P (Mesplé et al., 1996; Meinke, 1996; Heuvelink, 1996; Habekotté, 1996; Greenwood et al., 1985) is given by Theil (1961). The decomposition partitions link quality in terms of a least squares linear regression of M on P (when $M = A+BP+\varepsilon$). The decomposition can be written as:

$$\sum_{i=1}^{n}(M_i - P_i)^2 = n\left(A - (1-B)\overline{P}\right)^2 + (B-1)^2\sum(P_i - \overline{P})^2 + (1-R^2)\sum(M_i - \overline{M})^2 \quad (7.2)$$

The parameters A, B and the coefficient of determination R^2 are the result of regression of M on P. The three components in this sum jointly characterize the overall squared prediction error; individually they characterize specific aspects of the sum of squared prediction errors. The first two components indicate different types of systematic errors and jointly characterize prediction bias, whereas the third component indicates the importance of the random error (precision). For the L_2 norm to be 0, the values of A and B should be those of the identity link hypothesis (A = 0 and B = 1), and R^2 should equal 1.

An alternative decomposition of the sum of squared prediction errors (references as above) is the following:

$$\sum_{i=1}^{n}(M_i - P_i)^2 = n(\overline{M} - \overline{P})^2 + n(s_M - s_P)^2 + 2n(1-R)s_M s_P$$

where (7.3)

$$s_M^2 = \frac{1}{n}\sum_{i=1}^{n}(M_i - \overline{M})^2 \quad \text{and} \quad s_P^2 = \frac{1}{n}\sum_{i=1}^{n}(P_i - \overline{P})^2$$

This decomposition can be interpreted as a comparison of descriptive aspects of both data series (M and P): the first term compares the means, the second term compares the standard deviations, and the third compares the actual correlation (between M and P) to the ideal correlation (R = 1).

These decompositions are of use as a diagnostic for a limited class of systematic errors (those that can be removed using a linear transformation of the original data). Obviously there is no simple relation between this systematic error and its cause.

The coefficient of determination in the linear regression decomposition is sometimes presented as a separate measure to quantify prediction quality (Aitken, 1973; Mayer and Butler, 1993; Green and Stephenson, 1986). To illustrate its interpretation it can also be written as:

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(M - (A + BP_i))^2}{\sum_{i=1}^{n}(M_i - \overline{M})^2} \quad (7.4)$$

This measure characterizes the sums of squares explained by model predictions modified by the

relation ($M = A+BP$). As shown above, this correlation reflects only part of the total prediction error. It is more useful to characterize the prediction error in terms of all its components (for discussion cf. Harrison, 1990; Mayer et al., 1994), unless the relation $M = A+BP$ is actually used in prediction.

An alternative, the correlation ratio between M and P, is calculated as:

$$Y^2 = 1 - \frac{\sum_{i=1}^{n}(M_i - \overline{M})^2}{\sum_{i=1}^{n}(M_i - \overline{M})^2} \tag{7.5}$$

This measure is known as the model efficiency (Aitken, 1973; Mayer and Butler, 1993; Green and Stephenson, 1986). It characterizes the relative accuracy of the predictor.

Both measures characterize the sum of squared prediction errors relative to the sum of squares over the validation set. The sum of squares over the validation set is sometimes interpreted as the prediction error of an alternative predictor, the mean of the values in the validation set.

7.2.3 Uncertainty of the link quality

A single average prediction error is often sufficient to yield an impression of the predictive quality of a model. However, it does not allow to distinguish between models that have been validated in a single instance, and models that have been validated in a large number of instances.

If we have to choose between models with prediction errors determined over a different number of validations, one would like a measure of the reliability of the average prediction error. Repeated validation allows to establish the distribution of the prediction errors.

The variance of the squared prediction errors is calculated as (Wallach and Goffinet, 1989):

$$Var((M_i - P_i)^2) = \frac{1}{n-1}\sum_{i=1}^{n}((M_i - P_i)^2 - MSEP)^2$$

where (7.6)

$$MSEP = \frac{1}{n}\sum_{i=1}^{n}(M_i - P_i)^2$$

This is the standard variance estimator. The variance of the mean prediction error, var($MSEP$), then decreases with the number of validations (n) as $\mathrm{var}((M_i - P_i)^2)/n$.

7.3 Unacceptable validation results

A model is not very useful if all link hypotheses that are relevant for possible applications lead to predictive errors larger than those of alternative or benchmark predictors. However, this still

need not lead to final rejection of the model. This depends to a large extent on the possibility to use the alternative predictors. This was already stated by de Wit et al. (1978, page 68): "..., a model resembles a car in use: components may be malfunctioning, but the user is reluctant to discard the vehicle until something better turns up." Of course it should be possible to revise these malfunctioning components, once the car is temporarily parked.

Model revisions

If the model predictions are not acceptable, an analysis of the differences between predictions and measurements should be undertaken. In that case detailed measurements become relevant, e.g. time series of state variables measured within a single growing season.

Differences between the model outputs and measurements become the subject of investigation. The simplest way to proceed is to plot differences against some interesting variable, e.g. time, or the simulated values. This is known as residual analysis in statistics. In such an analysis the systematic (i.e non-random) part of the prediction error becomes the focus of the analysis.

Finding hypotheses to explain and correct for systematic differences is a creative process for which no guidelines can be formulated. An uncertainty analysis for groups of parameters may aid in establishing likely candidates to explain the systematic differences. If used to reformulate the model it should be clear that these differences are indeed systematic. Modifying models on the basis of one validation yielding a large prediction error introduces the risk of irrelevant ad-hoc hypotheses, as the established prediction error has an unknown uncertainty (cf. Section 7.2.3). A case study of revisions of a simulation model given validation results is e.g. presented by Beck (1987).

Instead of revising the model after an unsuccessful validation, one may already formulate alternative models prior to validation, calibrate them and use them for prediction. The results of the validation are used to choose between models. This approach removes the necessity to interpret the residuals, and allows to use validation as a method of model selection, for which it is often used in statistics. This approach is used by e.g. de Wit et al. (1970), Sinclair and de Wit (1976), Rossing (1991), Querner (1993), Habekotté (1996). As in model revision, the decision to accept an alternative model should be made on the basis of multiple validation instances. A single validation instance is no basis for model revision.

Black box model revisions

The prediction of absolute values may be that important and the time available for model modification that short, that - instead of revising the model - the model output is embedded in a black box model containing additional arguments, such as time, or other model outputs. These arguments may account for (part of) the systematic prediction error as found in validation. Predictions may then be made e.g. on the basis of a linear relationship between measurements M and prediction results P, ($M = A+BP$). Other arguments, which are not incorporated in the model, such as a time trend in yield due to better varieties or changing crop management, lead to a further extension of this linear model. Such an extension is e.g. used by de Koning et al. (1993),

for the prediction of yield in the European community.

A disadvantage of this approach is that additional regression coefficients have to be estimated. Once established, these output transformations may allow to reduce differences between predictions and measurements, but have to be validated themselves. Once validated, the simulation model is a part of a black box model and the identity link hypothesis can again be used to establish the prediction error.

7.4 Validation of SUCROS87 and LINTUL

7.4.1 A literature review

In literature it is often stressed that validation is not a matter of 'once and for all', but is an ongoing process that incrementally increases 'confidence' in a model. A limited review of validations for both models is therefore included in this chapter. Validation results reviewed refer to versions of the model for potential production.

The literature search on which the review was based was furthermore limited to those sources that explicitly refer to SUCROS87 or LINTUL. The validation results are reviewed in terms of simulated potential and measured final yield. This is a rather limited scope, and this review does not do justice to the full historical and conceptual range of publicized validations, as especially SUCROS87 is an offspring within an extended genealogical tree of often re-named models. A full review of the simulation models to which SUCROS87 and LINTUL are related is presented by Bouman et al. (1996), on which a broader review of validation results may be based. A short review without these restrictions is presented by Kropff et al. (1993).

The results of the validations are presented in Table 7.2. Not all comparisons between predicted and observed yield are based on experiments in which potential yield was achieved. As the authors (Table 7.3: [3], [4]) suggest the prediction errors for onions and crambe are large, due to the fact that actual, probably water-limited, production data were used for comparison. The results for rice are based on predictions using a model that was partly calibrated on the data to be predicted. Excluding these results, the relative prediction error is in the order of 20% for both models.

7.4.2 Validation of SUCROS87 and LINTUL for maize

Studying validation of SUCROS87 and LINTUL is interesting from a methodological point of view, and is interesting in terms of the predictive quality of the models, as the final step before application. To establish the predictive quality of the calibrated models, potential final yield was predicted for forage maize under Dutch conditions.

For prediction, final yield data of total aboveground dry matter weight (TADRW) for two crop rotation experiments with maize executed in the neighbourhood of Wageningen (Scholte, 1987) were available. The dataset met conditions for potential production in that the crop was sprinkled to reduce the effect of moisture shortage, sufficient fertilizer was applied, and the crop

was sprayed for weeds. Sowing and harvesting dates are similar for both experiments. The datasets available for both calibration and validation are presented in Table 7.4. The number of validation instances is too small to analyse systematic errors, which is anyway not very interesting to do on the basis of final yields alone.

We will limit the analysis to the use of a measure characterizing the overall prediction error, the root mean squared error of prediction (*RMSEP*), calculated as:

$$RMSEP = \left(\frac{1}{n} \sum_{i=1}^{n} (P_i - M_i)^2 \right)^{\frac{1}{2}} \tag{7.7}$$

where P_i are the predictions defined as below, M_i the measurements, and n is the number of *M-P* pairs.

Table 7.2 The root mean squared prediction error for SUCROS87 and LINTUL based on published results; the relative error (%) is given between brackets, n denotes the number of validation instances. References are presented in Table 7.3 using the corresponding numbers.

Crop		Average observed yield (kg.ha^{-1})	*RMSEP*		n	Table 7.3
			SUCROS87	LINTUL		
carrot		12242	2858 (23)	2198 (18)	13	[1]
cassava		7561	1355 (18)		4	[2]
crambe		2145		1308 (61)	6	[3]
onions		3207	4913 (153)	4985 (155)	18	[4]
potato		12764	1821 (14)	2955 (23)	2	[5]
rice	biomass	17750	1456 (8)		4	[6]
	yield	9500	843 (9)		4	
sugarbeet		65400	16485 (25)	11315 (17)	20	[7]
sugarbeet		74333	7277 (10)		12	[8]
winter oilseed rape		4700	300 (6)		1	[9]
winter wheat		8700	1783 (21)		16	[10]

Table 7.3 Limited overview of references presenting validations of SUCROS87, or SUCROS-based models, and LINTUL, or LINTUL-based models for different crops.

[1] Visser C.L.M. de, J.A. Schoneveld, and M.H. Zwart-Roodzant, 1995. Development of a crop growth model for carrots based on SUCROS87. Ontwikkeling van een gewasgroeimodel voor peen op basis van SUCROS87. (Verslag No. 199). Lelystad: Proefstation voor de Akkerbouw en de Groenteteelt in de Vollegrond, 124 pp.

[2] Gijzen, H., H.J. Veltkamp, J. Goudriaan, and G.H. de Bruijn, 1990. Simulation of dry matter production and distribution in cassava. Netherlands Journal Agricultural Science, 38: 159-173.

[3] Mathijssen, E.W.J.M. and W.J.M. Meijer, 1995. Development and application of a crop growth model for crambe. Ontwikkeling en toepassing van een gewasgroei-model voor Crambe. DLO-Instituut voor Agrobiologisch en Bodemvruchtbaarheidsonderzoek,. Rapport No. 38, 33 pp.

[4] Visser, C.L.M. de, 1992. Studying the course of growth in seed-raised onions and the construction of a growth model. (Bestudering van het groeiverloop van zaaiuien en bouw van een groeimodel). (Verslag No. 142). Lelystad: Proefstation voor de Akkerbouw en de Groenteteelt in de Vollegrond, 264 pp.

[5] Kooman, P.L. and C.J.T. Spitters, 1995. Pages 253-274 in: Modelling and parameterization of the soil-plant-atmosphere system - a comparison of potato growth models. Kabat, P., B. Marshall, B.J. van den Broek, J. Vos, and H. van Keulen (Eds). Wageningen: Wageningen Pers.

[6] Kropff, M.J., K.G. Cassman, H.H. van Laar, R.O. Torres, S.P. Liboon, R.L. Williams, and T. Horie, 1994. Evaluation of the model ORYZA1. Chapter 4 in: Kropff, M.J., H.H. van Laar, and R.B. Matthews (Eds). ORYZA1: an ecophysiological model for irrigated rice production. (SARP Research proceedings). Wageningen: DLO-Research institute for Agrobiology and Soil Fertility.

[7] Smit, A.B. and P.C. Struik, 1995. The first step towards a decision-support system for sugar-beet growing: selection of a basic growth model. Journal of Agronomy and Crop Science 175: 4, 213-220.

[8] Bouman, B.A.M., 1992. SBFLEVO and WWFLEVO: growth models to simulate crop growth, optical reflectance and radar backscatter of sugar beet and winter wheat, calibrated for Flevoland. (Report No. 163) Wageningen: CABO-DLO [62 p.]

[9] Backx, M., J. van Duivenvoorden, and J. Goudriaan, 1984. Simulation of the production pattern of rape seed on the basis of a field experiment. Netherlands Journal Agricultural Science 32: 247-250.

[10] Bouman, B.A.M., 1992. SBFLEVO and WWFLEVO: growth models to simulate crop growth, optical reflectance and radar backscatter of sugar beet and winter wheat, calibrated for Flevoland. (Report No. 163) Wageningen: CABO-DLO [62 p.]

Table 7.4 Overview over the datasets available for validation and those used for calibration. H: Wageningen Hoog, variety *Dorina*, reference I; A: Achterberg, variety *Dorina*, ref. I; E: Wageningse Eng, variety *LG11*, ref. II; D: Droevendaal, variety *LG11*, ref. III; S: Sinderhoeve, variety *Vivia*, refs. IV,V.

Year	Location				
	H	A	E	D	S
81	H1*	A1			
82	H2	A2			
83	H3	A3	E3*!		
84	H4	A4			
85	H5*	A5		D5*!	S5*
86	H6	A6			S6*
87		A7			

* : Datasets used for calibration, not used in validation

! : Datasets probably used in model development, not used in validation

References to experiments used in calibration and validation

I : Scholte, K., 1987. Relationship between cropping frequency, root rot and yield of silage maize on sandy soils. Netherlands Journal of Agricultural Science 35(1987) 473-486.

II : Sibma, L. and W. Louwerse, (unpublished report) Changes in quantity and quality in a maturing maize silage crop.

III : Louwerse, W., L. Sibma, and J. van Kleef, 1990. Crop photosynthesis, respiration and dry matter production of maize. Netherlands Journal of Agricultural Science 38 (1990) 95-108.

IV : Ouwerkerk, I. and H. Drenth, 1986. Verslag beregeningsonderzoek naar het gewas snijmais 1985. Interne mededeling nr. 439. Lelystad: PAGV.

V : Ouwerkerk, I., 1987. Resultaten van het beregeningsonderzoek in het gewas snijmais in 1986. Interne mededeling nr. 461. Lelystad: PAGV.

The methodological questions are the following:

1) Is calibration useful in reducing the prediction error?
2) What strategy should be used if calibration results for more than one set are available: average the parameters and predict using average parameter values, or use each calibration result for a separate prediction and average the predictions?
3) Is there a difference in prediction error between a relatively simple model (LINTUL) and a relatively complex model (SUCROS87)?

These questions require several prediction runs of the model. The structure of the validation is

presented in Table 7.5, and is identical for SUCROS87 and LINTUL.

Predictions were made using the parameter values calibrated for individual datasets (Table 7.5, columns 1 and 2). For each individual calibration set 6 best parameter vectors were available (cf. Chapter 6). Instead of using the best parameter vector out of 6, the average of 6 predicted final yields was regarded as the yield prediction.

Predictions were also made using parameter values averaged over two datasets, in which the same variety was grown ((H1 and H5), (D5 and E3), (S5 and S6)). In case the set of selected and calibrated parameters did not correspond between different datasets, the average of the nominal and the average calibrated value was used for prediction (Table 7.5, column 3).

Alternatively, predicted yield was calculated as the average of the yield prediction for datasets (H1 and H5), (D5 and E3) and datasets (S5 and S6) (Table 7.5, column 4).

Predictions were also made using the nominal parameter value (Table 7.5, column 5). The nominal version of LINTUL was parameterized on the basis of the literature research. Phenological parameters were based on the parameters used in SUCROS87. The original parameter values in SUCROS87 were retained, except for the emergence parameters TSEMER and TBSEM, that were based on Sibma (1987).

Table 7.5 The validation scheme for both SUCROS87 and LINTUL as executed in this chapter. Six datasets (H1, H5, D5, E3, S5 and S6) have been used for calibration. Two dataseries, Wageningen-Hoog (H) and Achterberg (A) are used for validation. These are predicted using different parameter values, different averages of parameter values, and different averages of predictions.

Data series predicted	1) Parameters from	2) Parameters from	3) Average of parameters (1,2)	4) Average of predicted yield (1,2)	5) Nominal
H	H5	H1	H1 H5	H1 H5	
A	H5	H1	H1 H5	H1 H5	
H	S5	S6	S5 S6	S5 S6	
A	S5	S6	S5 S6	S5 S6	
H	D5	E3	D5 E3	D5 E3	
A	D5	E3	D5 E3	D5 E3	

7.4.3 Results and methodological conclusions

The variation around the predicted yield due to the 6 calibration results was generally small, which suggests that choosing the best parameter vector would not have yielded very different results. The prediction errors are summarized in terms of the root mean squared prediction error (*RMSEP*), as defined earlier, over the validation sets. The *RMSEP* was calculated from data from which the years used in calibration were excluded. If e.g. parametersets derived from datasets D5

and E3 were used, predictions in 1983 and 1985 were excluded from the calculation of the *RMSEP*. Results are presented in Table 7.6 (SUCROS87) and Table 7.7 (LINTUL).

Table 7.6 Root mean squared prediction error (*RMSEP*, kg DM ha^{-1}) for predictions using SUCROS87 with different parameter sets based on calibration of the different datasets. Layout of this table is based on Table 7.5.

Data series predicted	1	2	3	4	5
H	1911	2972	3318	2095	1967
A	3420	2532	2549	1929	2840
H	4542	3097	1335	1379	1967
A	3991	4831	2362	2373	2840
H	6671	5177	2690	1659	1967
A	6369	6437	2068	761	2840

Table 7.7 Root mean squared prediction error (*RMSEP*, kg DM ha^{-1}) for predictions using LINTUL with different parameter sets based on calibration of the different datasets. Layout of this table is based on Table 7.5.

Data series predicted	1	2	3	4	5
H	3089	2088	1255	2697	3814
A	2919	2653	2575	2729	2899
H	1555	1131	608	443	3814
A	1656	2556	1729	1670	2899
H	2837	1618	606	752	3814
A	2833	3699	1977	1893	2899

Interpretation of validation results in terms of methodology

The results were analysed in terms of relevant root mean squared errors to formulate some tentative answers to the methodological questions asked in Section 7.4.2. It is emphasized that the analysis is not a formal statistical analysis: the results presented are based on a relatively small set of validation results. We have refrained from presenting estimation errors associated with the *RMSEP*.

The first and third question were whether calibration is useful and whether there is a difference in prediction error between LINTUL and SUCROS87. To answer this question we calculated the average *RMSEP* for each of the methods to generate predictions. The results are

presented in Table 7.8 (in kg DM ha^{-1}). They indicate that the effect of calibration is not clearcut. The effect of calibration on prediction error depends on the methods used to generate the predicted value, and on the model. For LINTUL calibration allows a decrease of the prediction error in comparison with the non-calibrated model. For SUCROS87 a clear improvement in comparison to the non-calibrated model is only achieved for the case in which the prediction is based on averages over predictions (Table 7.8, third column). As to differences in prediction error between SUCROS87 and LINTUL: the results suggest that neither model is systematically better than the other.

Another problem (question 2, section 7.4.2) is which strategy should be used if more than one calibration set is available for calibration: average the parameters, or average the predictions; or base the predictions on calibration results using a single data set? The results in Table 7.8 indicate that for SUCROS87 any type of averaging is to be preferred over the single calibration set prediction results. The advantage of averaging is also clear for LINTUL. Whether to use average parameter vectors for prediction or use average predicted yield as a predictor is not clear: the difference between both is relatively small, and the results for SUCROS87 and LINTUL show an ambiguous effect. The procedure most convenient for practical use is to average parameters over calibrations and over calibration sets

Returning to Tables 7.6 and 7.7 one notes that the prediction error varies quite substantially within the individual columns, that is over prediction procedures. This variation is probably caused by the use of different varieties to predict, and is also an effect of the individual calibration sets (e.g. location, experimental design). The data do not allow to assess the relative importance of possible factors.

Table 7.8 The prediction error averaged over all datasets for different methods. In brackets relative error with respect to overall average yield (%).

	Average prediction error using parameter values from calibration results of individual sets (Table 7.5, column 1, 2)	Average prediction error using average parameter values from 2 sets (Table 7.5, column 3)	Average prediction error using yield predictions averaged over two sets (Table 7.5, column 4)	Average prediction error using nominal parameter values
SUCROS87	4616 (31)	2448 (17)	1783 (12)	2476 (17)
LINTUL	2524 (17)	1688 (11)	1936 (13)	3353 (23)

7.4.4 Link quality and usefulness

The available data also allowed to establish the following benchmark prediction errors, all calculated as *RMSEP*, for comparison:

1) the prediction error if the yield in one year is used to predict the yield in the next year (Zero-order forecast error),

2) the prediction error if the average from the available yield data is used to predict a 'missing' value (Average leave-one-out prediction errors),
3) the prediction error if the yield measured in Wageningen-Hoog is used to predict the yield in Achterberg (Reference site prediction error, or given specific conditions, yield uncertainty).

The selection of 'the' model prediction error for comparison with these three benchmark predictors is difficult and to a certain extent subjective, given the procedures possible to generate predictions, and the calibration sets available. The prediction error selected for comparison is the average prediction error using average parameter values from two sets (Table 7.5, column 3). This prediction error is selected because we would like to evaluate the predictive capacity of the models for the range of calibration sets available. In this validation experiment the prediction error is unavoidably also averaged over varieties. If these varieties differ in terms of the parameters calibrated one would expect the selected prediction error to be larger than that determined using the same variety in all instances. The selection of averaged parameter values is suggested by the advantages in application.

For both models this prediction error is compared to the three reference measures defined above; the results are presented in Table 7.9. The reference site prediction error may be interpreted as a rough estimate of yield uncertainty, as both experiments in Achterberg and Wageningen-Hoog are identical in experimental design, but only roughly identical as to planting and harvesting days.

In all cases, the model predictions offer an improvement over a zero-order forecast (Column 1 in Table 7.9), a very simple prediction rule. Model predictions are also the better alternative if missing values in time series of final yield have to be hindcast for known planting and harvesting dates (Column 2 in Table 7.9). Apart from the calibrated version of LINTUL, model hindcasts are not an alternative if yield has been measured at a different location under roughly the same conditions (Column 3 in Table 7.9). If the reference site prediction error is (re-)interpreted as a rough estimate of yield uncertainty, the results show that the selected (calibrated) prediction error is 21% smaller for LINTUL, 12% larger for SUCROS87 (Column 3 in Table 7.9). This suggests that further improvement of predictive quality for LINTUL may depend on the availability of local weather data, whereas for SUCROS87 improvement is still possible (assuming that measured yield is indeed potential).

Model prediction error can be compared with measurement error within a production instance. In this comparison the effect of measurement plot size within a field should be taken into account. Measurement variance of yield per unit area (V_x) determined from replicated plots decreases with increasing plot size x as:

$$V_x = V_1 x^{-b} \tag{7.8}$$

where the constants V_1 and b are estimated using linear regression (Fairfield Smith, 1938). This equation can also be written as:

$$\frac{V_{x1}}{V_{x2}} = \left(\frac{x_1}{x_2}\right)^{-b} \tag{7.9}$$

where x_1 and x_2 are two different plot sizes. In terms of a ratio of standard errors (s_1 and s_2) for two different plot sizes (x_1 and x_2) this can be written as:

$$\frac{s_1}{s_2} = \left(\frac{x_1}{x_2}\right)^{-\frac{b}{2}} \tag{7.10}$$

If $b = 1$, standard errors decrease as if the data consist of independent contributions of smaller plots within the original plot. The standard error is independent of plot size for $b = 0$. Values of b presented in literature (Fairfield Smith, 1938; Hallauer, 1964) determined from a regression of yield variance on plot size reflect actual conditions in a production instance, and are therefore difficult to generalize.

The experimental replicates available for Achterberg allow to estimate the standard error over replicates to be 425 kg ha^{-1} for a sample size of 30 m^2 (about 330 plants). For assumed values of b, and for s_2 = 425 kg ha^{-1} and x_2 = 330 plants, the equation can be used to calculate a rough estimate of the experimental plot size (in number of plants) which would yield a measurement error equal to the selected prediction error (x_1) for both models, assuming that the model prediction error is random. These equivalent plot sizes are presented in Table 7.10. For the present study (potential yield) the equivalent plot size serves to visualize the magnitude of the prediction error. For other simulation studies equivalent plot sizes may be compared to e.g. the (net) sampling plot sizes, used in selection experiments (15-20 plants; proposed under conditions of optimality of experimental design, Chaves and de Miranda Filho (1992)), or those used in Dutch variety trials for silage yield (about 25 plants; Ebskamp (1981)). In variety trials or in selection trials such a comparison is relevant if the models used have been calibrated and validated for a single variety, and if the models incorporate the limiting factor of the trial (as the experiments are often executed under non-potential conditions).

Conclusions

A conclusion which stands out from the results is that the calibration executed in Chapter 6 not necessarily reduces prediction error. It does so for LINTUL, but not for SUCROS87. An important factor that determines the prediction error is the way in which parameter estimates are used to generate predictions. Using either averaged predictions, or averaged parameter vectors to generate predictions result in prediction errors that are smaller than the prediction error based on single calibration sets (both as an average over different calibration sets). Comparing selected model prediction errors to reference predictors (Table 7.9) allows to conclude that in all cases the models offer an improvement over a zero order forecast, a very simple prediction rule. The models are in most cases a satisfactory alternative to provide predictions for missing values in time series of forage maize yields (as compared to using the average yield). If yield has been measured at another location, model hindcasts are in most cases not an alternative. These conclusions are very sensitive to the procedure followed to generate the predictions.

The prediction error, if random, can be expressed in terms of a plot size that is equivalent in terms of its sampling error.

Table 7.9 The comparison of selected prediction errors for the two models before (nominal) and after calibration to errors of reference predictors (both in kg DM ha^{-1}). Numbers in the column heading refer to the following reference errors: 1 - Using one year to predict the next (Zero order forecast); 2 - Using the mean of available yield to calculate missing values (Average leave-one-out prediction error), 3 - Using the yields in Wageningen-Hoog to estimate those in Achterberg (Yield uncertainty). (+ : model is better; – : model is worse).

		Prediction error reference predictor		
		1	2	3
Selected prediction error		3553	2932	2145
SUCROS87 (nominal)	2476	+	+	–
SUCROS87 (calibrated)	2448	+	+	–
LINTUL (nominal)	3353	+	–	–
LINTUL (calibrated)	1688	+	+	+

Table 7.10 Prediction error expressed in terms of an equivalent experimental plot size (in number of plants), for independent plots (b = 1) and slightly correlated plots (b = 0.75). Number of plants cut off to nearest integer.

Selected prediction error		$b = 1$	$b = 0.75$
SUCROS87 (nominal)	2476	9	3
SUCROS87 (calibrated)	2448	9	3
LINTUL (nominal)	3353	5	1
LINTUL (calibrated)	1688	20	8

Chapter 8

General discussion and recommendations

8.1 Introduction

Simulation modelling is a tool in applied agricultural research. The results achieved using this tool should satisfy evaluation by concerned and involved parties. As all steps leading to the end result are potentially subject to discussion and evaluation, all steps should be executed in such a manner that they can be evaluated by the parties concerned.

The notion 'model auditing' was introduced to describe a systematic and accountable execution of specific steps towards model application (model development and model adaptation: cf. Figure 1.2 (Chapter 1)). Model auditing provides input to the evaluation.

The aim of research in this thesis was to develop methods to assess and quantify the predictive quality of simulation models, with the intent to contribute to construct evaluation i.e. discussion and evaluation of model studies by non-scientists. The case study focused on the prediction of forage maize yield under potential conditions in the Netherlands using two simulations models, SUCROS87 and LINTUL.

The methodology developed is evaluated and discussed in the first section of this chapter. In the second section results are discussed. In a third section the possibility of using the methodology in a different context, that of decision support, and the changes then required are considered. These considerations are used as the basis for recommendations.

8.2 Methodology

The methodology in this thesis is based on the premise that the simulation models have already passed their scientific certification. This is the argument that allows to exclude discussions regarding structural uncertainty - i.e. the mathematical structure of the models and the plant-scientific concepts used - and the software implementation from this thesis. It serves the practical purpose of further restricting the amount of work to be done.

In this case study the software code constitutes the non-modifiable basis on which to develop procedures for model auditing. Whereas the individual procedures become ad-hoc, as the models do not necessarily fit - and given the premise - can not be made to fit within well-defined mathematical templates, the methodology is developed along the lines sketched in Chapter 1, Figure 1.2.

In the approach the simulation model is first described in a systematic manner to simplify access to the software code (Chapter 2). It is subsequently analyzed to establish the information which allows to select the parameter estimation algorithm (Chapter 3).

After an analysis in which all numerical values in the model algorithms are re-assessed and classified as either constants or as (system)-parameters, the empirical basis of the simulation model is established in terms of parameter values and their uncertainty. To do so, parameter uncertainty is formally defined and the possible methods to establish parameter uncertainty are reviewed (Chapter 4). The parameter uncertainty for SUCROS87 and LINTUL are established on the basis of an extensive literature review.

The effect of parameter uncertainty on the model output is established using a Monte Carlo approach. The results of uncertainty analysis in terms of the output uncertainty are used as the

basis for a decision to calibrate a model. An analysis of output uncertainty in terms of the different parameter contributions is used to determine a parameter ranking (Chapter 5). To characterize the relative contribution of an uncertain parameter on model output, two measures - top and bottom marginal uncertainty contribution of a parameter - were used. The ranking based on this analysis reflects the combined effect of parameter uncertainty and that of the models sensitivity to the parameter. It is therefore used as a method to establish priorities in parameter estimation.

The analysis of model properties (Chapter 3) provides the arguments for a choice of the parameter estimation algorithm. The parameter estimation procedure is executed (Chapter 6) in such a way as to take problems of parameter identifiability, local optima, and parameter estimation bias into account, thus trying to avoid that the parameter is reduced to a mere regression constant. Parameter estimation is based on a random search algorithm, and is executed by repeatedly using a procedure in which an initially large number of parameters is reduced by dropping parameters from the estimation problem. Given that the random algorithm retains a set of parameter values that is gradually improved, the decision to drop parameters is based on the consistent shrinkage of the parameter variation with respect to the variation in the initial set. The consistency of shrinkage is determined over a number of independently parameterized repetitions of the algorithm. The initial parameter selection was based on the parameter contribution to output uncertainty (Chapter 5); to maximize parameter identifiability, parameters that could not be estimated very well were fixed at their nominal value.

The parameter estimates or combinations of parameter estimates are used to generate predictions. Prediction and measured data that were to be predicted are used to evaluate the predictive quality in terms that are relevant for the application. Predictive quality is based on the evaluation of the deviations between predictions and measurements given an anticipated relation between them, the so-called link hypothesis. Different options for link hypotheses are presented. To translate predictive quality in terms of usefulness of the simulation model prediction errors are compared to those of benchmark predictors (simple, often statistical, predictors; Chapter 7).

8.3 Methodological uncertainty

Of course, we would have liked to state not only that each of the procedures summarized in the previous section is accountable and leads to reproducible results, but also that the procedure has been formulated (and selected) on the basis of irrefutable criteria. However, this is not the case and a discussion of procedural choices made is unavoidable. It should be noted that conclusions in this case study are conditional on these choices, and may be sensitive to them. The degree of sensitivity is not clear, and has to be established through tests.

Analysis of parameter uncertainty and uncertainty contributions

In the review of parameter uncertainty (Chapter 4) the decision was made to evaluate parameter uncertainty for production instances 'all over the world'. The decision was based on the anticipated result of the parameter review if limited to Dutch production instances only. In that case

the number of instances recovered for each parameter was expected to become too small to allow a satisfying assessment of parameter uncertainty.

Even when based on a review of instances 'all over the world' the number of reported values for each parameter was too small to fit a probability distribution. The distributions, either a gamma or a beta distribution, were therefore chosen on the basis of theoretical parameter bounds. Apart from means and variances, the number of covariances (correlations) that could be retrieved from literature was very small. The possibility that (small) correlations exist between the parameters can not be excluded, nor can we exclude the possibility that if more data become available another distribution type would be selected.

Establishing parameter uncertainty for a different population will outweigh the effect of the distribution type chosen to describe parameter uncertainty. The effect of small correlations could be more important than the effect of distribution type as combinations of extreme parameter values would become less likely.

Initial parameter selection for calibration

The selection of the initial set of parameters is based on the contribution of the parameters to the uncertainty in final yield. Those parameters are selected for calibration that allow the largest expected reduction in the uncertainty of final yield, if they would become perfectly known. In Chapter 5 selection is based on the top and bottom marginal contribution to the uncertainty in final total aboveground dry matter weight averaged over a number of years. This criterion, however, is certainly not the only one possible. Selection could also be based on the marginal contributions to uncertainty evaluated over a number of output dates. By combining or aggregating these in different ways, parameters could be ranked on the basis of a vast number of criteria, each ranking possibly leading to a different initial selection of parameters for calibration. This points out that the use to which the model is put should serve as a guideline to select the criterion for parameter ranking.

Another aspect of parameter selection is the problem what to do with parameters for which no values could be retrieved from literature. In the approach chosen the parameters were fixed at their default value. However, other approaches, such as always including these parameters in the initial selection are possible.

Calibration

The procedural uncertainty in calibration results from the choices made in all previous procedures, combined with uncertainty regarding the estimation procedure. In the procedure presented in this thesis it was tried to take problems of discontinuous relationships between parameters and state variables, of parameter bias, of parameter identifiability and of local minima into account. As a theoretical basis for analysis is not available, the only way to test the effects of different assumptions is through extensive testing of the procedure using problems with known solutions (using artificial calibration sets), or using simpler problems (that can be analyzed with other techniques).

Apart from the test of the calibration result using the Simplex algorithm as an alternative to

the (default) Price algorithm, a limited test of the calibration procedure was executed for SUCROS87 using a slightly different sampling method (six independently drawn samples from the parameter distribution, as in the calibration procedure for LINTUL). Illustrative for the problems in calibrating the models is the fact that for one dataset the parameters retained in the estimation procedure after one dropcycle differed from those in Chapter 6, whereas the mean of the calibration criteria was within 1 s.d. of the best criterium presented in Chapter 6, suggesting that the procedure has 'found' the same minimal criterion value. If we disregard the difference between the sampling methods, this suggests that the parameter selection procedure based on dropcycles is sensitive to the number of repetitions. In that case parameters retained in one calibration could have been dropped in a repetition of the calibration for another set of initial samples. Obviously, given these non-robust results, parameter estimates derived using the calibration procedure are not comparable to those derived from designed experiments.

Validation

In the chapter on validation, a test on procedural uncertainty is executed. Different procedures to generate predictions on the basis of calibration results are compared and are shown to have an important effect on prediction error. This comparison leads to prefer averaging parameter values over averaging predictions.

8.4 Case study: discussion

The discussion in this section suggested by results of the individual chapters is intended to focus attention on areas of interest with respect to both models.

Model description and model analysis

The model description showed a large difference in the complexity of both models. In terms of the model structure SUCROS87 is more complex than LINTUL: in number of parameters, number of relations, and in number of feedback loops. Analysis of relations between state variables and individual parameters in SUCROS87 were shown to be discontinuous for some parameters. This effect was caused by the use of switches, functions with a discontinuous output, in the model source code implemented as IF-THEN-ELSE constructs. The magnitude of the discontinuity introduced is sometimes increased by a combination of these switches with threshold functions and adverse weather conditions. Threshold functions are functions with a discontinuous derivative, e.g. T_DVRV and T_DVRR in SUCROS87 (cf. Figure 2.2).

A common element in the models that requires some thought are the tabulated functions (partitioning tables among others). These were retained in their original form, but transformations of the table argument and the response allow to modify them (cf. Appendix 7). The parameters associated with these transformations are considered for calibration.

A justification for the use of switches and tables is that switches reflect biological concepts and that tables are of educational value, in that they stress the empirical nature of the relations

used. However, switches and to a certain extent their combination with threshold functions restrict the methodological options in parameter estimation; tables are a nuisance. Replacing both switches and tables by functions would simplify parameter estimation, both from literature data and in calibration.

Empirical basis

The empirical basis of the parameters used in the models was assessed in Chapter 4. The overview of parameter uncertainty for production instances 'all over the world' shows that uncertainty varies strongly between parameters. The coefficient of variation for individual parameters varied between a maximum of 143% to a minimum of 3%, with an average of 38% (for SUCROS87) and 32% (for LINTUL). It was not possible to select probability distributions on the basis of their fit. These were therefore selected on the basis of considerations regarding parameter ranges.

Results showed that in SUCROS87 the parameters associated with senescence, respiration (both growth - and maintenance respiration) and partitioning could not, or only partly, be retrieved from literature. In the case study there were two options: to use the default values for the parameters or to include the parameters in calibration. Other options - reformulate the model in terms suggested by the literature review, or execute the necessary experiments to determine the missing parameter values - were excluded. In the description of SUCROS87 the authors suggest to calibrate senescence parameters on the basis of field data. In this thesis the simpler option, to use the default values, was chosen.

Assessment of the empirical basis sometimes allows to consider options for model reformulation. To illustrate this, consider the development parameters. At present the parametrization of development in SUCROS87 is based on three temperature sums: temperature sum from sowing to 50% emergence (TSEMER), from 50% emergence to 50% silking (TSDVRV), and from 50% silking to maturity (TSDVRR). If, as for maize, the temperature sum to emergence and the temperature sum to silking are based on the same base temperature, a development stage defined as zero at sowing (as by de Groot et al. 1986), instead of zero at emergence in the present parameterization, uses more of the available data. These two possible parameterizations and the associated number of parameter values actually retrieved from literature are presented in Table 8.1.

Output uncertainty and parameter contributions to uncertainty

Based on the concepts of top and bottom marginal uncertainty contributions, the relative contribution of an uncertain parameter to model output uncertainty was determined for both models. In Tables 8.2 and 8.3 we have combined this relative contribution of parameters to output uncertainty with the number of retrieved parameter values. Parameters with a relatively large contribution that are not readily retrieved from literature suggest that either more effort should be put in being exhaustive in the literature review, or that there is a need for experimental research regarding these parameters. The tables show that in both models there are parameters (bold in both tables) that belong to this category.

Table 8.1 Two possible parameterizations of crop development in maize and the number of production instances in literature from which parameter values could be retrieved.

Development as used in SUCROS87 for maize at present	Number of parameter values	Alternative parametrization	Number of parameter values
sowing to emergence (TSEMER)	114	sowing to emergence	114
emergence to silking (TSDVRV)	16	sowing to silking	224
silking to maturity (TSDVRR)	81	silking to maturity	81

Table 8.2 A classification of parameters for SUCROS87, based on relative availability of reported parameter values (average = 20 instances), and top marginal uncertainty contribution of a parameter (average = 3%). Bold print indicates parameters to which attention should be paid, either experimentally, or in terms of literature retrieval. For an explanation of the abbrevations cf. Appendix 1.

SUCROS87		Top marginal uncertainty contribution	
		Below average	Above average
Relative availability	Below average	FRDFCA, FRDFCB DLYATR LAINI MAINLV AAMTMP ATRL, FRDFBB, FRDFBA, FRPAR ASRSO, ASRRT DVSSEN AAMDVS KDIF, SCV RGRL	**EFF** **MAINRT** **Q10** **AMX** **BSLA** **TSDVRV**
	Above average	TSEMER BLV	TSDVRR BSH

Table 8.3 A classification of parameters for LINTUL, based on relative availability of reported parameter values (average = 42 instances), and top marginal uncertainty contribution of a parameter (average = 8 %). Bold print indicates parameters to which attention should be paid, either experimentally, or in terms of literature retrieval. For an explanation of the abbrevations cf. Appendix 2.

LINTUL		Top marginal uncertainty contribution	
		Below average	Above average
Relative availability	Below average	DVLGTH RI, DVHALF FRPAR	**TSDVR** **FINTI**
	Above average	TSEMER LUE	

Calibration

The output uncertainty showed that the models had to be calibrated, given that population of production instances for which the default parameter values available for SUCROS87 were valid was unknown, whereas default parameter values for LINTUL were not available.

Whereas the uncertainty analysis strongly suggests that the model should be adapted to the local conditions, the default parameter values in SUCROS87 yield prediction results for locations near Wageningen that are difficult to improve upon. This suggests - not surprisingly - that the default parameters used in SUCROS87 are representative for conditions near Wageningen. The usefulness of uncertainty analysis as an indicator for the need to calibrate depends on the assumption that the parameter values for the application of interest are unknown. In some situations in which default parameter values are available, even if the population of production instances for which they are representative is unknown, it is more efficient to establish predictive quality first, before executing a time-consuming calibration procedure.

Validation

In Chapter 1 model evaluation was used to refer to two different steps: assessing output uncertainty on the basis of parameter uncertainty, and assessing the predictive quality of a model. The prediction error in final yield for both models was between 10 and 30% (Table 7.8), which is of the order found in analyses of the prediction error for potential yield for various other crops using SUCROS87 and LINTUL (Table 7.2, roughly between 5 and 25%).

A comparison of the model prediction errors to errors obtained from alternative simpler predictors (Chapter 7, benchmarks) allows to qualify the usefulness of the model relative to the simpler predictors. This comparison also has a bearing on the problem of model building and model choice. Model builders focus on the formulation of models that are in some sense optimal

for the intended application (cf. Chapter 1). The relation between complexity (increasing in the order benchmark-LINTUL-SUCROS87) and prediction error is presented in Figure 8.1. This figure shows that the variation in the prediction error for each single model is large compared to the variation of the mean prediction error over complexity. As this is due to the calibration set(s) used and the procedure followed to generate the prediction, the question of complexity is less important than the question just how to generate predictions for either LINTUL or SUCROS87 and what data are available for calibration, at least in this case study.

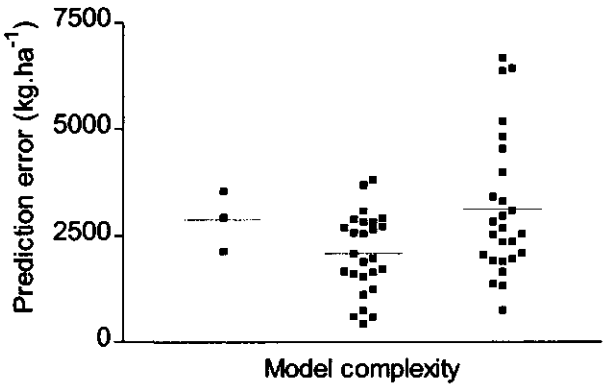

Figure 8.1 Prediction errors and their mean as a function of complexity (increasing in the order benchmark predictors, LINTUL, and SUCROS87). Data are jittered with respect to x-axis to present individual errors.

8.5 A broader perspective: applying the methodology in a different research context

The methodology presented in this thesis was developed within a intermediate-term theoretical research project at a university. A likely context for application is a short-term project for decision-support.

The question to be discussed in this section is which consequences this change in research context has for the proposed procedures. We will assume that the basic premises remain the same: the model used has already been formulated and is available in terms of a software code. We assume that the case study changes; we are dealing e.g. with a different model, a different crop and a different location. Finally we assume that the model is developed as a stand-alone source code, in contrast to models developed in modelling environments, such as MATLAB, or statistical environments, such as GENSTAT, in which some - but not all - of the recommendations presented in this discussion are straightforward to realize.

The model description (Chapter 2), the analysis of model structure, and the relations between state variables and parameters (Chapter 3) are independent of the application. Given that

the model already exists, there is no reason why these steps should be part of a project. However, in that case model documentation and analysis of model structure require that the model owner ensures their execution independent of any project, both in terms of funding and in terms of time available. We will assume that the number of parameters in the model is high and requires parameter selection prior to calibration. The analysis of parameter uncertainty (Chapter 4) required to provide the basis for parameter selection depends on the application. If the case study changes, the uncertainty of at least part of the parameters has to be re-assessed for the changed conditions. Given that a literature review of a single parameter takes minimally one week, this is a time-consuming activity, best executed external to a project. Again, this requires that the model owner ensures its execution independent of any project, both in terms of funding and in terms of time available.

Given the parameter uncertainty, the execution of uncertainty analysis and the estimation of parameter contributions to relevant model output is a relatively fast procedure. There are a number of requirements that - if met - allow an efficient (or less complex and time-consuming) execution:

1) Software is available that allows to generate the desired distributions.
2) The number of input files required is minimal: separate input files for different parameter values should not be required.

   In the proposed procedure parameter selection for calibration depends on the analysis of output uncertainty given the parameter uncertainty. If parameter uncertainty is application dependent, so is calibration. We assume that, unless the output uncertainty is sufficiently small, or unless default parameter values for the specific population of production instances are available, the model has to be calibrated. Given specific model properties, the calibration procedure may be time-consuming. The procedure executed in this thesis certainly is. There are a number of requirements to be met if an existing model is to be calibrated efficiently:
4) The model owner should provide or have access to a method for parameter estimation that takes model properties into account.
5) If the model owner is not able to execute an uncertainty analysis to serve as a basis for parameter selection, an alternative procedure for parameter selection should be available (unless of course the number of parameters is sufficiently small to be estimated from the calibration data).
6) The model owner should have access to calibration data and the associated input (e.g. meteorological data series), and these should have been checked for errors (verified). These three requirements are again either application independent (calibration procedure and parameter selection procedure) or time-consuming (determining or searching for calibration data). Obviously, this also requires a fundamental commitment.
7) Validation (Chapter 7) is the single most important procedure prior to model application. One may skip calibration, if e.g. parameter values established for the population of instances of interest are available, but comparison of validation results with benchmark predictors is the test of the model's usefulness. To be able to validate the model, the model owner should provide or have access to validation data (and their corresponding inputs). If this requirement is not met or can not be met, the conclusions should be based on the weakest link hypothesis, i.e. that ranks of predictions and (anticipated) measurements correspond.

In the context of this discussion, the work in this thesis may serve as an illustration for some of the points made, as it has benefitted from prior investment in the methodological and in the empirical basis. Specific procedures such as the parameter estimation algorithm (FSEOPT, Stol et al., 1992), but also subroutine libraries (TTUTIL, Rappoldt and van Kraalingen, 1990; WEATHER, van Kraalingen et al., 1991) had already become available in a form (the Fortran Simulation Environment (FSE), van Kraalingen, 1991) that allowed to embed the models considered. This thesis could not have been written without the longterm meteorological data from the Wageningen meteorological station, (Dept. Physics and Meteorology, 1954-1992) and without data from experiments and their description filed away and made available by individual scientists (cf. Chapter 6 and 7).

From the above it may be concluded that facilitating evaluation in the context of decision support comes at a price, as it requires investment, both in the methodological basis, and in the empirical basis associated with the model. Providing the empirical basis (parameter uncertainty and the data for calibration, validation and application) requires the most substantial effort, even more so as this empirical basis is both model- and application dependent. Providing the methodological basis is application independent - though model-dependent - and once available, often less of a worry. From this point of view the number of propositions dealing with or expressing worry about the accesibility of the empirical basis (Chapter 1) and a relative lack of similar propositions regarding the methodological basis can be understood.

The above discussion can be summarized in a number of recommendations for model-based applied research. The starting point for the work presented in this thesis is that an existing simulation model should only be made available for applied research (e.g. for decision support) if the procedures leading to the results can be made available to those concerned or interested and can be evaluated by them. Evaluation of a model and of the results of its application are facilitated, if apart from the models documentation and a description of its input requirements, the procedures required to present the uncertainties regarding the results can be efficiently executed and are easily documented. This is the case, if:

- The required input is available and has been described and verified.
- Model output can be efficiently generated for a large number of different parameter values in a single simulation session.
- A method for parameter estimation can be used that takes model properties into account.
- A procedure for parameter selection in parameter estimation has been formulated and the information required to execute this procedure (in this case study parameter uncertainty in terms of a multivariate probability distribution) is available.
- Validation data are available. If they are not available, conclusions should be based on the weakest link hypothesis, i.e. that ranks of predictions and (expected) measurements correspond.

Summary

Simulation modelling is a tool in applied agricultural research. The results achieved using this tool should satisfy evaluation by scientists and by concerned and involved parties. As all steps leading to the end result are potentially subject to discussion and evaluation, all steps should be executed in such a manner that they can be evaluated by the parties concerned.

Chapter 1 describes the different steps in model-based applied research and reviews the concerns among researchers in this field. Researchers try to formulate an ideal model, where 'ideal' is defined in terms reflecting the tension between the requirements of the application and scientific standards (e.g. simplicity vs. complexity). A coarse classification of the different steps towards application of a model indicates that there is a need for qualitative methodology and a need for quantitative methodology. In general terms, research in this thesis focuses on quantitative methodology. More specific the aim of research in this thesis was to develop methods to assess and quantify the predictive quality of simulation models, with the intent to contribute to construct evaluation, i.e. discussion and evaluation of model studies by non-scientists.

A case study is presented in which two models of different complexity, LINTUL and SUCROS87, are used to predict yield of forage maize under Dutch meteorological conditions. The models predict yield under potential conditions, i.e. temperature- and radiation limited yield, assuming other production factors to be optimal. The case study is based on two models to allow to indicate effects of model complexity.

In a first step the simulation models are described in a systematic manner (Chapter 2). Next (Chapter 3), the models are analysed in terms of properties that are relevant for parameter estimation, a step possibly required to adapt the model to the conditions of this case study. The analysis shows that the models contain switches, describing abrupt changes occurring in the crop (e.g. change of temperature driven leaf area growth to photosynthesis driven leaf growth; onset of leaf senescence). Some switches introduce discontinuities in the relation between state variables and parameters. It is concluded that these model properties prevent the use of methods based on linearisation of the model in its parameters.

In Chapter 4 an inventory of the empirical basis of the models in terms of parameter values and their uncertainty is made. To do so, parameter uncertainty is formally defined and the possible methods to establish it are discussed. Parameter uncertainty is defined as the uncertainty about the parameter value in a new production instance. It is modelled as the variation of a parameter value between production instances, corrected for the limited accuracy of the parameter estimates for the individual instances that have been studied. It is described in terms of probability distributions. Production instances are individual experimental instances in which a parameter has been estimated under conditions of potential growth. A statistical model is formulated to assess parameter uncertainty on the basis of a literature inventory.

The second part of Chapter 4 consists of the detailed presentation of the parameter values reported in literature. The review shows that for the population of maize production instances 'all over the world' the parameters have an average coefficient of variation of about 30%. Uncertainty itself is very variable between parameters (a coefficient of variation of 3% for the most secure, and of 143% for the most uncertain parameter).

This inventory of parameter uncertainty is used in Chapter 5 to assess and analyze the model

output uncertainty under the conditions of the case study (defined by Dutch meteorological conditions and management data). The analysis provides estimates of the contributions of individual parameters to the output uncertainty. The uncertainty contributions are defined as the expected reduction in yield prediction uncertainty should the parameter become perfectly known. Output uncertainty established on the basis of parameter uncertainty 'all over the world' is large. It is not to be expected that e.g. the average parameter vector yields the best prediction for Dutch conditions. The models must therefore be adapted to Dutch conditions by estimating the appropriate parameter values - they must be calibrated.

In Chapter 6 three parameter estimation procedures are presented and discussed. For a situation in which only very limited information is available, a calibration procedure similar to set-calibration is discussed. The results indicate that even very limited information allows to reduce initial uncertainty regarding parameter values. For a situation in which no parameter estimation algorithm is available, the possibility of manual point calibration is discussed. The discussion shows that in order to allow evaluation by parties other than the scientist a large number of decisions have to be documented. The third procedure is automatic calibration: an iterative computer algorithm is used to optimize the correspondence (fit) between measured and simulated data. In contrast to hand calibration it allows simultaneous adaptation of multiple parameters.

In parameter estimation one often has to cope with problems of poor identifiability, estimation bias and locally optimal fits. As the models contain a large number of parameters relative to the size and the nature of the datasets available for parameter estimation, it may well happen that optimal or non-optimal fits can be realized with widely diferent parameter vectors. Depending on the seriousness of the problem it is known as poor parameter identifiability or as parameter non-identifiability. The simplest solution would seem to be to estimate as few parameters as possible. However, in that case, the other parameters have to be fixed at their, possibly not entirely correct, default values. This may introduce a systematic error (estimation bias or fixing error) in the estimated parameter estimates. Locally optimal fits are fits that deteriorate if the parameter vector is slightly changed, whereas there exists a different parameter vector, that yields a better fit. Because of the discontinuities described in Chapter 3, local minima should be expected. The estimation procedure executed is an attempt to cope with these different problems.

The Price algorithm, a controlled random search algorithm, is the basis for the point estimation procedure executed. This algorithm was selected to cope with local optima. The algorithm evaluates the correspondence between measured and simulated data (fit) for a set of parameter vectors. In each iteration step, the parameter vector associated with the worst fit is discarded if a randomly generated parameter vector results in a better fit. After a number of iterations the ranges of all parameters have often shrunk compared with the ranges in the initial set.

The estimation procedure executed consists of a number of partly iterative steps. First, parameters are selected on the basis of a ranking derived from their contributions to output uncertainty. To keep the fixing error small, the parameter estimation algorithm is initially run with a relatively large number of parameters. In the estimation procedure the Price algorithm is applied iteratively in order to select an identifiable subset of the selected parameters. The most important selection criterion is shrinkage of the parameter range compared to its initial range.

In Chapter 7, the parameter estimates are used to generate predictions. A comparison

between predictions and measured data is used to evaluate the predictive quality of the models in terms that are relevant for the application. To do so, the concept of a link hypothesis is introduced. It defines the anticipated relation between prediction and measurement. Deviations from the anticipated relation are used to quantify predictive quality. Predictive quality was shown to depend strongly on the way in which predictions are generated. Predictions were generated using parameter values estimated from a single calibration set, were generated using the average of parameter values estimated from different calibration sets, and were generated as yield averaged over predictions based on different calibration sets. Predictions using several calibration sets yielded better predictions than predictions based on a single calibration set.

To express predictive quality in terms of relative usefulness of the simulation model, prediction errors are compared with those of simple, often purely statistical, benchmark predictors. The simplest benchmark predictor is a one year ahead predictor, the second is the predictive error made when using the mean over a number of yield data as a predictor to generate a single missing value (leave-one-out predictor), and a third is the prediction error made when using a series of yield data as the predictor for yield at a nearby location. Results showed that both models performed better than the simplest benchmark predictor and were the better alternative to generate missing values. LINTUL and SUCROS87 differed in their performance in relation to the third benchmark predictor. In this situation LINTUL performed better than the benchmark predictor, whereas SUCROS87 did not.

The methodology and the results are discussed in Chapter 8. Based on the results in Chapter 7, a more general conclusion in terms of the complexity of the models is that, at least in this case study, the variation in prediction error over the different ways to generate predictions was at least as important as the effect of model complexity on the prediction error. The discussion regarding the ideal model can not be isolated from the data and the ways used to generate a prediction.

The point of view in this thesis is that an existing simulation model should only be made available for applied research (for decision support) if the procedures leading to the results can be evaluated. A discussion of the procedures presented in this thesis in this research context, (Chapter 8) leads to the conclusion that this is only possible if the procedures can be efficiently executed and are easily documented. This is the case, if:

- The required input is available and has been described and verified.
- Model output can be efficiently generated for a large number of different parameter values in a single simulation session.
- A method for parameter estimation that takes model properties into account is available for use.
- A procedure for parameter selection in parameter estimation has been formulated and the information required (in this case study parameter uncertainty in terms of a multivariate probability distribution) to execute this procedure is available.
- Validation data are available, otherwise the conclusions should be based on the weakest link hypothesis, i.e. that ranks of predictions and system realizations correspond.

These recommendations are based on the assumption that models are provided as stand-alone software-codes, with little or no associated tools. Given this assumption these recommendations require actions that are not easily executed within the context of project-based, often time-limited, applied research. Execution requires investment in the methodological basis and in the empirical basis of the models.

Samenvatting

Simulatiemodellen zijn een stuk gereedschap binnen toegepast landbouwkundig onderzoek. De resultaten die met dit gereedschap bereikt worden, moeten evaluatie door wetenschappers en door betrokkenen doorstaan. Omdat alle stappen die tot het resultaat leiden in principe onderwerp van discussie kunnen worden, moeten deze stappen zodanig uitgevoerd worden dat evaluatie door betrokkenen mogelijk is.

Hoofdstuk 1 beschrijft de verschillende stappen binnen het op modellen gebaseerde toegepaste onderzoek en geeft een overzicht over wat in dit veld als een probleem gezien wordt. Onderzoekers proberen een ideaal model te formuleren. Dit ideaal wordt gekarakteriseerd in termen die spanning uitdrukken tussen de eisen vanuit de toepassing en eisen van wetenschappelijkheid (b.v. eenvoudig vs. complex). Een grove classificatie van de verschillende stappen noodzakelijk voor toepassing geeft aan dat er behoefte is aan zowel kwalitatieve alsook aan kwantitatieve methodologie.

Heel algemeen richt het werk in dit proefschrift zich op de kwantitatieve methodologie, meer in het bijzonder was het doel de ontwikkeling van methoden om de voorspellende kwaliteit van het model vast te stellen en te evalueren. De intentie was het daarbij bij te dragen aan 'construct evaluatie', een term die geïntroduceerd werd om discussie en evaluatie van model-studies door derden aan te geven.

In dit proefschrift wordt een case studie gepresenteerd, waarin twee modellen van verschillende complexiteit (LINTUL en SUCROS87) gebruikt worden om de opbrengst van snijmaïs onder Nederlandse weersomstandigheden te voorspellen. De modellen berekenen de opbrengst bij optimale omstandigheden, i.e. de temperatuur- en stralingsbeperkte opbrengst, aangenomen dat andere produktiefactoren optimaal zijn. De case studie is gebaseerd op twee modellen om effecten van modelcomplexiteit aan te kunnen geven.

In een eerste stap worden de simulatiemodellen systematisch beschreven (Hoofdstuk 2). Vervolgens (Hoofdstuk 3) worden de modellen geanalyseerd m.b.t. eigenschappen die relevant zijn voor parameterschatting, een mogelijk noodzakelijke stap waarin het model aan de nederlandse omstandigheden wordt aangepast. De analyse laat zien dat de modellen schakelaars bevatten, die abrupte overgangen in het gewas beschrijven (zoals de verandering van temperatuurgedreven bladontwikkeling naar photosynthesegedreven bladgroei, begin van bladafsterving). Een aantal schakelaars introduceren discontinuïteiten in de relatie tussen toestandsvariabelen en parameters. Deze modeleigenschappen verhinderen het gebruik van methoden die gebaseerd zijn op linearisering van het model in de parameters.

In Hoofdstuk 4 wordt de empirische basis van de modellen geïnventariseerd in termen van parameterwaarden en hun onzekerheid. Om dat te kunnen doen wordt de parameter-onzekerheid gedefinieerd als de onzekerheid van de parameterwaarde in een nieuwe produktie-instantie. Parameteronzekerheden worden beschreven in termen van kansverdelingen. Produktie-instanties zijn individuele experimentele instanties, waarin een parameter geschat wordt. Een statistisch model wordt gebruikt om de parameteronzekerheid op basis van de literatuur samen te vatten.

Het tweede gedeelte van Hoofdstuk 4 bestaat uit het gedetailleerde overzicht van de informatie over parameterwaarden in de literatuur. Het overzicht laat zien dat voor de populatie van maïsproduktie-instanties over de hele wereld de parameters een gemiddelde

variatiecoefficient van ongeveer 30% hebben. De onzekerheid zelf wisselt sterk tussen parameters (een variatiecoefficient van 3% voor de zekerste, en van 143% voor de meest onzekere parameter.).

Deze inventarisatie van de parameteronzekerheid wordt in Hoofdstuk 5 gebruikt om de onzekerheid van de modeluitvoer onder de omstandigheden van de case studie (nederlandse meteorologische omstandigheden en teeltparameters) vast te stellen en te analyseren. De analyse levert schattingen van de bijdragen van individuele parameters aan de onzekerheid van de modeluitvoer. De onzekerheidsbijdragen van parameters zijn gedefinieerd als de verwachte reductie in de onzekerheid van de voorspelling, mocht de parameter precies bekend worden. Onzekerheid in de oogstvoorspelling gebaseerd op de parameteronzekerheid 'van over de hele wereld' is groot. Het is niet te verwachten dat bijvoorbeeld de gemiddelde parametervector de beste voorspelling onder nederlandse omstandigheden geeft. De modellen moeten daarom worden aangepast aan nederlandse omstandigheden (gecalibreerd) door de parameterwaarden voor die produktieinstantie te schatten.

In Hoofdstuk 6 worden drie parameterschattingsmethoden gepresenteerd en besproken. In een situatie waarin alleen beperkte informatie beschikbaar is, wordt een calibratieprocedure besproken die vergelijkbaar is met set-calibratie. De resultaten tonen aan dat zelfs zeer beperkte informatie de initiële onzekerheid van de parameterwaarden kan verkleinen. In een situatie waarin geen parameterschattingsalgoritme beschikbaar is wordt de mogelijkheid van een punt-calibratie met de hand gepresenteerd. Een voorbeeld laat zien dat om derden de procedure te laten evalueren een fors aantal beslissingen moet worden vastgelegd. De derde procedure is punt-calibratie met een optimalisatie algoritme: een iteratief computeralgoritme wordt gebruikt om de overeenkomst (fit) tussen gemeten en gesimuleerde data te optimaliseren. In tegenstelling tot calibratie met de hand laat deze procedure gelijktijdige aanpassing van meerdere parameters toe.

In parameterschatting moet men oplossingen vinden voor slechte identificeerbaarheid, systematische schatfouten en lokle optima. Omdat de modellen een groot aantal parameters bevatten, zeker gelet op de grootte en aard van de datasets op basis waarvan ze geschat moeten worden, kan het gebeuren dat optimale fit kan worden gerealiseerd met sterk verschillende parametervectoren. Afhankelijk van de mate waarin dit probleem optreedt, heeft men te maken met slechte parameter identificeerbaarheid of met parameter niet- identificeerbaarheid. De eenvoudigste oplossing in dit geval is het om zo min mogelijk parameters te schatten. In dat geval moeten de niet geschatte parameters op een vaste waarde gesteld worden. Een minder juiste of mogelijk onjuiste waarde kan tot een systematische fout (bias of fixing error) leiden bij de schatting van de andere parameters. Lokaal optimale fits zijn fits die slechter worden bij een verandering van de parameterwaarden, terwijl er een parameterwaarde bestaat die een betere citeriumwaarde levert. Vanwege de discontinuïteiten beschreven in Hoofdstuk 3 moeten lokle minima verwacht worden. De calibratieprocedure die uitgevoerd wordt is een poging deze verschillende problemen te ondervangen. Het Price-algoritme, een zoekalgoritme gebaseerd op gestuurd toeval, is de basis voor de gepresenteerde parameterschattingsprocedure. Het algoritme was gekozen om met het probleem van lokle optima te ondervangen. Het algoritme vergelijkt de overeenkomst tussen metingen en simulatie (de fit) voor een verzameling parametervectoren. In elke iteratiestap wordt de parametervector geassocieerd met de op dat moment slechtste fit

weggegooid, wanneer de gegenereerde vector tot een betere aanpassing leidt. Na een aantal iteraties is het interval van de parameters vaak gekrompen t.o.v. het begininterval.

De uitgevoerde schattingsprocedure bestaat uit een aantal, gedeeltelijk iteratieve, stappen. De eerste stap is de keuze van de parameters op basis van een rangschikking afgeleid uit hun bijdrage aan de voorspelonzekerheid. Om de systematische fout klein te houden, wordt het parameterschattings-algoritme in eerste instantie met een relatief groot aantal parameters uitgevoerd. In een iteratieve toepassing van het Price algoritme wordt het aantal geschatte parameters gereduceerd om zo een identificeerbare subset van parameters te kiezen. Het belangrijkste selectiecriterium is de krimp van het parameterinterval relatief t.o.v het begininterval.

In Hoofdstuk 7 worden de parameterschattingen gebruikt voor voorspellingen. Een vergelijking tussen voorspellingen en metingen wordt gebruikt om de voorspellende kwaliteit van modellen te beschrijven in termen die relevant zijn voor toepassingen. Om dat te kunnen doen wordt het begrip 'link hypothese' geïntroduceerd, als de geanticipeerde relatie tussen voorspelling en meting. Afwijkingen van de geanticipeerde relatie worden gebruikt om de voorspelkwaliteit te kwantificeren. De voorspelkwaliteit blijkt heel erg af te hangen van de manier waarop de voorspellingen gegenereerd worden. In dit proefschrift werden deze op meerdere manieren gegenereerd: gebruikmakend van een enkel calibratieresultaat, gebruikmakend van parameterwaarden gemiddeld over meerdere calibratieresultaten, en gemiddelde voorspellingen gebruikmakend van meerdere calibratieresultaten. Voorspellingen gebruikmakend van meerdere calibratie-resultaten waren beter dan die gebaseerd op een enkel calibratieresultaat.

Om de voorspelkwaliteit te vertalen naar de bruikbaarheid van het simulatiemodel wordt de voorspelfout van het model vergeleken met de voorspelfout van eenvoudige, puur statistische, testvoorspellers. De eenvoudigste testvoorspeller is de '1 jaar vooruit voorspeller', een tweede is het gemiddelde over de opbrengsten over een aantal jaren (waarin de te voorspellen opbrengst ontbreekt), dat gebruikt word om het ontbrekende getal te voorspellen. Een derde testvoorspeller is een serie opbrengsten van een andere, dichtbijgelegen, lokatie. De resultaten lieten zien dat beide modellen beter waren dan de eenvoudigste testvoorspeller, en dat zij ook ontbrekende waarden in een reeks beter voorspellen. LINTUL en SUCROS87 verschilden t.o.v. de derde testvoorspeller. In die vergelijking bleek LINTUL beter te zijn dan de testvoorspeller, dit in tegenstelling tot SUCROS87.

De methodologie en de resultaten staan in Hoofdstuk 8 ter discussie. Op basis van de resultaten in Hoofdstuk 7 is een conclusie ten aanzien van de complexiteit van de modellen die dat - in deze studie - de variatie van de voorspelfout als gevolg van keuzes in de procedure om tot een voorspelling te komen op z'n minst even belangrijk is als de variatie in de voorspelfout als gevolg van verschillen in modelcomplexitieit. De discussie over het ideale model kan niet geïsoleerd worden gezien van de data en de manieren waarop voorspellingen gegenereerd worden.

Uitgaand van het standpunt dat een bestaand simulatiemodel alleen beschikbaar mag worden gesteld voor toegepast (beleidsondersteunend) onderzoek, wanneer de procedures die tot het gepresenteerde resultaat leiden ter discussie kunnen worden gesteld, leidt de discussie in Hoofdstuk 8 tot de aanbeveling dat de gebruikte procedures efficiënt moeten kunnen worden

uitgevoerd en gemakkelijk moeten kunnen worden gedocumenteerd. Dit is het geval, wanneer:
- De vereiste invoer beschreven en beschikbaar is, en op fouten gecontroleerd.
- Modeluitvoer efficiënt kan worden gegenereerd voor een groot aantal parameterwaarden, en in een enkele simulatiesessie.
- Een parameterschattingsmethode die rekening houdt met modeleigenschappen beschikbaar is
- Een procedure voor keuze van parameters in parameterschattingen geformuleerd is en de informatie die voor uitvoering nodig is beschikbaar is (in deze studie parameteronzekerheid als multivariate kansverdeling).
- Validatiegegevens beschikbaar zijn; wanneer dit niet het geval is, moeten de conclusies van een studie gebaseerd worden op de zwakste linkhypothese: de rangschikking van voorspelling en toekomstige meting komt overeen.

Deze aanbevelingen zijn gebaseerd op de aanname dat modellen worden geleverd als op zichzelf staande software, met weinig of geen ondersteunende software. Op basis van deze aanname vereisen de aanbevelingen acties die niet eenvoudig kunnen worden uitgevoerd in de context van beleidsondersteunend onderzoek, waarvoor vaak maar beperkt tijd beschikbaar is. Om te voldoen aan deze voorwaarden moet worden geïnvesteerd in de methodologische en de empirische basis van de modellen.

References

Aggarwal, P.K., 1995. Uncertainties in plant, soil and weather inputs used in crop growth models: implications for simulated outputs and their application. Agricultural Systems 48(3): 361-384.

Aitken, A.P., 1973. Assessing systematic errors in rainfall-runoff models. Journal of hydrology 20: 121–136.

Aldenberg, T., J.H. Janse, and P.R.G. Kramer, 1995. Fitting the dynamic model PClake to a multi-lake survey through Bayesian Statistics. Ecological Modelling 78: 83-99.

Allen, L.H., C.S. Yocum, and E.R. Lemon, 1964. Photosynthesis under field conditions. VII. Radiant energy exchanges within a corn crop canopy and implications in water use efficiency. Agronomy Journal 57(6): 575-580.

Bair, E.S., 1994. Model (in)validation - a view from the courtroom. Ground water 32(4): 530-531.

Baker, C.H. and R.B. Curry, 1976. Structure of agricultural simulators: a philosophical view. Agricultural Systems 1: 201-218.

Banks., J., D. Gerstein, and S.P. Searles, 1988. Modelling processes, validations, and verification of complex simulation models: a survey. Pages 13-18 in: Balci, O. (Ed.). Methodology and validation. San Diego: Soc. for Comp. Sim., Simulation Series 19(1).

Bard, Y., 1974. Non-linear parameter estimation. New York: Academic press, 341 p.

Bastiaans, L., 1993. Understanding yield reduction in rice due to leaf blast. PhD-thesis, Wageningen Agricultural University, 127 p.

Bastiaanssen, W.G.M., 1995. Regionalization of surface flux densities and moisture indicators in composite terrain. A remote sensing approach under clear skies in Mediterranean climates. PhD-thesis, Wageningen Agricultural University, 273 p.

Beck, M.B., 1981. Hard or soft environmental systems? Ecological Modelling 11: 233-251.

Beck, M.B., 1987. Water quality modeling: a review of the analysis of uncertainty. Water Resources Research 23(8): 1393-1442.

Berge, H.F.M. ten, 1986. Heat and water transfer at the bare soil surface. Aspects affecting thermal imagery. PhD-thesis, Wageningen Agricultural University, 214 p.

Berge, H.F.M. ten, 1991. Building research capacity for systems research at national agricultural research centres: SARP's experience. Pages 515-538 in: Penning de Vries, F.W.T, P.S. Teng, and K. Metselaar (Eds). Systems approaches for agricultural development. Dordrecht: Kluwer.

Berge, H.F.M. ten, K. Metselaar, M.J.W. Jansen, E.M. de San Augustin, and T. Woodhead, 1995. The SAWAH riceland hydrology model. Water Resources Research 31(11): 2721-2732.

Beumer, L., A. van Gameren, B. van der Hee, and J. Paelinck, 1978. A study of the formal structure of J.W. Forrester's urban dynamics model. Urban Studies 15: 167-177.

Bignon, J., 1990. Agrométéorologie et physiologie du maïs grain dans la Communauté Européenne. Bruxelles-Luxembourg: Centre Commune de Recherche, Commission des Communautés Européennes, 195 p.

Birch, C.J., 1996. Modelling the ontogeny, leaf area and dry matter accumulation and distribution in diverse cultivars of maize. PhD-thesis, University of Queensland, 258 p.

Blokker, K.J., 1984. Computergestuurde voorlichting: een decisiegericht voorlichtingskundig onderzoek naar Epipre en andere geautomatiseerde informatiesystemen in de landbouwvoorlichting. PhD-thesis, Wageningen Agricultural University, 389 p.

Blower, S.M. and H. Dowlatabadi, 1994. Sensitivity and uncertainty analysis of complex models of disease transmission: an HIV model, as an example. International Statistical Review 62(2): 229-243.

Blower, S.M., D. Hartel, H. Dowlatabadi, R.M. Anderson, and R.M. May, 1991. Drugs, sex and HIV; a mathematical model for New York City. Phil. Trans. R. Soc. Lond. B 321: 171-187.

Boesten, J.J.T.I., 1986. Behaviour of herbicides in soil: simulation and experimental assessment. PhD-thesis, Wageningen Agricultural University, 263 p.

Boons-Prins, E.R., G.H.J. de Koning, C.A. van Diepen, and F.W.T. Penning de Vries, 1993. Crop specific simulation parameters for yield forecasting across the European Community. Simulation reports CABO-TT 32. Wageningen: AB-DLO, [123] p.

Bot, G.P.A., 1994. Vanuit behoud naar beweging. Inaugurele rede, Landbouwuniversiteit Wageningen, 2 juni 1994, 34 p.

Botkin, D.B., 1977. Bits, bytes and IBP (editorial). Bioscience 27(6): 385.

Botterweg, P., 1995. The user's influence on model calibration results: an example of the model SOIL, independently calibrated by two users. Ecological Modelling 81: 71-81.

Bouman, B.A.M, 1991. Linking X-band radar backscattering and optical reflectance with crop growth models. PhD-thesis, Wageningen Agricultural University, 169 p.

Bouman, B.A.M., H. van Keulen, H.H. van Laar, and R. Rabbinge, 1996. The 'School of de Wit' crop growth simulation models: a pedigree and historical overview. Agricultural Systems 52(2/3): 171-198.

Bredehoeft, J.D. and L.F. Konikow, 1993. Ground water models: validate or invalidate. Ground water 31(2): 178-179.

Breiman, L., J.H. Friedman, R.A. Olshen, and C.J. Stone, 1984. Classification and regression trees. Belmont (CA.): Wadsworth, 358 p.

Bresler, E. and G. Dagan, 1988a. Variability of yield of an irrigated crop and its causes. 1. Statement of the problem and methodology. Water Resources Research 24(3): 381-387.

Bresler, E. and G. Dagan, 1988b. Variability of yield of an irrigated crop and its causes. 2. Input data and illustration of results. Water Resources Research 24(3): 389-394.

Britton, C.M. and J.D. Dodd, 1976. Relationships of photosynthetically active radiation and shortwave irradiance. Agricultural Meteorology 17: 1-7.

Bruggeman, J.P., 1996. Formalizing organizational ecology. Logical and mathematical investigations in organization theory. PhD-thesis, University of Amsterdam, 114p.

Bunnik, N.J., 1978. The multispectral reflectance of shortwave radiation by agricultural crops in relation with their morphological properties. Mededelingen Landbouwhogeschool Wageningen 78-1, 176 p.

Bunting, E.S., 1976. Accumulated temperature and maize development in England. Journal Agricultural Science Cambridge 87: 577-583.

Carroll, J.J., 1985. Global transmissivity and diffuse fraction of solar radiation for clear and

cloudy skies as measured and as predicted by bulk transmission models. Solar Energy 35(2): 105-118.

Chaves, L.J. and J.B. de Miranda Filho, 1992. Plot size for progeny selection in maize. Theoretical and Applied Genetics 84: 963-970.

Chen De-Xing, M.B. Coughenour, A.K. Knapp, and C.E. Owensby, 1994. Mathematical simulation of C_4 grass photosynthesis in ambient and elevated CO_2. Ecological Modelling 73: 63-80.

Cleveland, W.S. and R. McGill, 1984. Graphical perception: Theory, experimentation, and application to the development of graphical methods. Journal of the American Statistical Association 79(387): 531-554.

Collatz, G.J., M. Ribas-Carbo, and J.A. Berry, 1992. Coupled photosynthesis-stomatal conductance model for leaves of C_4 plants. Australian Journal Plant Physiology 19: 519-538.

Colson, J., D. Wallach, A. Bouniols, J.-B. Denis, and J.W. Jones, 1995. Mean square error of yield prediction by SOYGRO. Agronomy Journal 87: 397-402.

Cook, T.D. and L.C. Leviton, 1980. Reviewing the literature: a comparison of traditional methods with meta-analysis. Journal of Personality 48(4): 449-472.

Cooke, R.M., 1991. Experts in uncertainty. Opinion and subjective probability in science. New York: Oxford University Press, 321 p.

Cooke, R.M., 1992. Expert judgement study on atmospheric dispersion and deposition: main report and annexes. (Report 91-81 of the faculty of technical mathematics and informatics) Delft: University of Technology, 260 p.

Cooper, H.M., 1984. The integrative research review: A systematic approach. (2^{nd} edition) Newbury Park: Sage, 157p.

Cross, H.Z. and M.S. Zuber, 1972. Prediction of flowering dates in maize based on different methods of estimating thermal time. Agronomy Journal 64: 351-355.

Dagan, G. and E. Bresler, 1988. Variability of yield of an irrigated crop and its causes. 3. Numerical simulation and field results. Water Resources Research 24(3): 395-401.

Dale, V.H., H.I. Jager, R.H. Gardner, and A.E. Rosen, 1988. Using sensitivity and uncertainty analysis to improve predictions of broad scale forest development. Ecological Modelling 42: 165-178.

Daynard, T.B., 1972. Relationships among black layer formation, grain moisture percentage, and heat unit accumulation in corn. Agronomy Journal 64: 716-719.

Demetriades-Shah, T.H., M. Fuchs, E.T. Kanemasu, and I. Flitcroft, 1992. A note of caution concerning the relationship between cumulated intercepted solar radiation and crop growth. Agricultural and Forest Meteorology 58: 193-207; Comment/Reply: Agricultural and Forest Meteorology 68: 213-242.

Dept. Physics and Meteorology, 1954-1992. Maandoverzicht weergegevens Station Wageningen 'De Haarweg'. Vakgroep Natuur- en Weerkunde (Dept. Physics and Meteorology), Wageningen Agricultural University.

Derieux, M. and R. Bonhomme, 1982a. Heat unit requirements for maize hybrids in Europe. Results of the European FAO sub-network. I. sowing - silking period. Maydica 27: 59-77.

Derieux, M. and R. Bonhomme, 1982b. Heat unit requirements for maize hybrids in Europe. Results of the European FAO sub-network. II. Period from silking to maturity. Maydica 27:

79-96.

Derieux, M. and R. Bonhomme, 1990. Heat unit requirements of maize inbred lines for pollen shedding and silking: results of the European FAO network. Maydica 35: 41-46.

Dorp, F. van, 1977. Simulatie van veranderingen in concentraties van voedingselementen rond plantewortels. PhD-thesis, Wageningen Agricultural University, 162 p.

Doucet, P. and P.B. Sloep, 1992. Mathematical modeling in the life sciences. New York: Horwood, 490 p.

Downing, D.J., R.H. Gardner, and F.O. Hoffman, 1985. An examination of response surface methodologies for uncertainty analysis in assessment models. Technometrics 27: 151-163; C/R: Technometrics (1986) 28: 91-93.

Draper, D., 1987. A review of "Summing up: the science of reviewing research." by R.J. Light and D.B. Pillemer. Journal of the American Statistical Association: 349-350.

Duan, Q., S. Sorooshian, and V. Gupta, 1992. Effective and efficient global optimization for conceptual rainfal-runoff models. Water Resources Research 28(4): 1015-1031.

Dubois, D. and H. Prade, 1989. Fuzzy sets, probability and measurement. European Journal of Operational Research 40: 135-154.

Ebskamp, A.G., 1981. Het cultuurwaarde-onderzoek van snijmaisrassen in Nederland (I) Wijze van beproeving. Bedrijfsontwikkeling 12(2): 175-181.

Fairfield Smith, H., 1938. An empirical law describing heterogeneity in the yields of agricultural crops. Journal Agricultural Science 28: 1-23.

Feddes, R.A., 1990. Waar blijft het water... Inaugurele rede, Landbouwuniversiteit Wageningen, 13 september 1990, 39 p.

Flénet, F., J.R. Kiniry, J.E. Board, M.E. Westgate, and D.C. Reicosky, 1996. Row spacing effects on light extinction coefficients of corn, sorghum, soybean, and sunflower. Agronomy Journal 88:185-190.

Flowers, J.H., and T.A. Hauer, 1993. "Sound" alternatives to visual graphics for exploratory data analysis. Behavior Research Methods, Instruments & Computers 25(2): 242-249.

Fresco, L.O., 1992. Zo niet nu, wanneer dan? Inaugurele rede, Landbouwuniversiteit Wageningen, 21 mei 1992, 39 p.

Friedman, J.H., 1991. Multivariate adaptive regression splines. Annals of Statistics 19: 1-67; C/R: 68-141.

Gallo, K.P. and C.S.T. Daughtry, 1986. Techniques for measuring intercepted and absorbed photosynthetically active radiation in corn canopies. Agronomy Journal 78: 752-756.

Gass, S.I., 1983. Decision aiding models: validation, assessment, and related issues for policy analysis. Operations Research 31(40): 603-631.

Gill, P.E., W. Murray, and M.H. Wright, 1981. Practical optimization. London: Academic press, 401 p.

Gilmore jr., E.C. and J.S. Rogers, 1958. Heat units as a method of measuring maturity in corn. Agronomy Journal 50: 611-615.

Glantz, M.H., 1982. Consequences and responsibilities in drought forecasting: The case of Yakima, 1977. Water Resources Research 18(1): 3-13.

Goicoechea, A., 1988. Expert system models for inference with imperfect knowledge: a comparative study. Journal of Statistical Planning and Inference 20: 245-277.

Goossens, L.H.J., R.M. Cooke, F. Woudenberg, and P. van den Torn, 1992. Probit functions and expert judgement: protocol and results. (report) Delft: University of Technology, 102 p.

Gosse, G., C. Varlet-Grancher, R. Bonhomme, M. Chartier, J-M. Allirand, and G. Lemaire, 1986. Production maximale de matiere seche et rayonnement solaire intercepte par un couvert vegetal. Agronomie 6(1): 47-56.

Goudriaan, J., 1977. Crop micrometeorology: a simulation study. Simulation Monographs, Wageningen: Pudoc, 249 p.

Goudriaan, J., 1986. A simple and fast numerical method for the computation of daily totals of crop photosynthesis. (short communication). Agricultural and Forest Meteorology 38: 249-254.

Goudriaan, J., 1988. The bare bones of leaf-angle distribution in radiation models for canopy photosynthesis and energy exchange. Agricultural and Forest Meteorology 43: 155-169.

Goudriaan, J., 1993. Model, schaal en aggregatieniveau. Inaugurele rede, Landbouwuniversiteit Wageningen, 30 september 1993, 29 p.

Goudriaan, J., 1996a. Predicting crop yields under global change. Chapter 14 in: Walker, B. and W. Steffen, (Eds). Global change and terrestrial ecosystems. International geosphere-biosphere programme book series. Cambridge: Cambridge university press.

Goudriaan, J., 1996b. The end of modelling? Pages 115-120 in: Stein, A., F.W.T. Penning de Vries, and P.J. Schotman (Eds). Models in action. Proceedings of a seminar series 1995-1996. Quantitative approaches in systems analysis 6. Wageningen: DLO Research Institute for Agrobiology and Soil fertility, the C.T. de Wit Graduate School for Production Ecology.

Goudriaan, J. and H.H. van Laar, 1994. Modelling potential crop growth processes. -Textbook with exercises. Current issues in production ecology, volume 2. Dordrecht: Kluwer, 238 p.

Green, I.R.A. and D. Stephenson, 1986. Criteria for comparison of single event models. Hydrological Sciences 31(3): 395-411.

Greenwood, D.J., J.J. Neeteson, and A. Draycott, 1985. Response of potatoes to N fertilizer: dynamic model. Plant and Soil 85: 185-203.

Grinsven, H.J.M. van (Ed.), 1993. Modelling water, carbon and nutrient cycles in forests: application of 16 simulation models to a spruce stand at Solling, Germany: Proceedings of a workshop held in Leusden, the Netherlands. Ecological Modelling 83(1-2).

Grinsven, H.J.M. van, C.T. Driscoll, and A. Tiktak, 1995. Workshop on comparison of forest-soil-atmosphere models: preface. Ecological Modelling 83: 1-6.

Groot, J.J.R., M.J. Kropff, F.J.H. Vossen, C.J.T. Spitters, and R. Rabbinge, 1986. A decimal code for the developmental stages of maize and its relation to accumulated heat units. Netherlands Journal of Agricultural Science 34: 67-73.

Groot, J.J.R, P. de Willigen, and E.L.J. Verberne (Eds), 1991. Nitrogen turnover in the soil-crop system: modelling of biological transformations, transport of nitrogen and nitrogen use efficiency: Proceedings of a workshop held at the institute of Soil Research, Haren, The Netherlands. Fertilizer Research 27(2-3) [volume].

Habekotté, B. 1996. Evaluation of seed yield determining factors in winter oilseed rape under potential growth conditions (Brassica Napus L.) by means of crop growth modelling (submitted to Field Crops Research). Chapter 4 in: Winter oilseed rape - analysis of yield formation and crop type design for higher yield potential. PhD-thesis, Wageningen

Agricultural University, 156 p.
Håkanson, L. and R.H. Peters, 1995. Predictive limnology - Methods for predictive modelling. Amsterdam: SPB, 464 p.
Hallauer, A.R., 1964. Estimation of soil variability and convenient plot size from corn trials. Agronomy Journal 56: 493-499.
Hamby, D.M., 1994. A review of techniques for parameter sensitivity analysis of environmental models. Environmental Monitoring and Assessment 32: 135-154.
Hamilton, M.A., 1991. Model validation: an annotated bibliography. Communications in Statistical Theory and Methodology 20(7): 2207-2266.
Hanway, J.J., 1963. Growth stages of corn. Agronomy Journal 55: 487-492.
Harrison, S.R., 1990. Regression of a model on real system output: an invalid test of model validity. Agricultural Systems 34: 183-190.
Hastie, T.J. and R.J. Tibshirani, 1990. Generalized additive models (Monographs on statistics and applied probability; 43). London: Chapman and Hall, 335 p.
Hedges, L.V. and I. Olkin, 1985. Statistical methods for meta-analysis. Orlando FL: Academic press, 370p.
Hedges, L.V., 1987. How hard is hard science, how soft is soft science? The empirical cumulativeness of research. American Psychologist 42(2): 443-455.
Heinen, M., 1997. Dynamics of water and nutrients in closed, recirculating cropping systems in glasshouse horticulture - with special attention to lettuce grown in irrigated sand beds. PhD thesis, Wageningen Agricultural University, 270 p.
Hendrickson, J.D., S. Sorooshian, and L.E. Brazil, 1988. Comparison of Newton-type and direct search algorithms for calibration of conceptual rainfall-runoff models. Water Resources Research 24(5): 691-700.
Henize, J., 1984. Critical issues in evaluating socio-economic models. Page 557-590 in: T.I. Ören et al. (Eds). Simulation and model-based methodologies: An integrative view. NATO ASI series Vol. F10, Springer.
Henten, E.J. van., 1994. Greenhouse climate management: an optimal control approach. PhD-thesis, Wageningen Agricultural University, 329 p.
Heuvelink, E., 1996. Dry matter partitioning in tomato: validation of a dynamic simulation model. Annals of Botany 77: 71-80.
Hillel, D., 1987. Modelling in soil physics: a critical review. Page 35-42 in: Future developments in soil science research. Madison: Soil Science Society of America.
Hocking, R.R., 1976. The analysis and selection of variables in linear regression. Biometrics 32: 1-49.
Hogarth, R.M., 1975. Cognitive processes and the assessment of subjective probability distributions. Journal of the American Statistical Association, 70(350): 271-289, Comment/Reply: 290-294.
Holmes, K.J., W.S. Chu, and D.R Erickson, 1989. Automated calibration of a contaminant transport model for a shallow sand aquifer. Groundwater 27(4): 501-508.
Hopmans, J.W. and E. Guttierez-Rave, 1988. Calibration of a root water uptake model in spatially variable soils. Journal of Hydrology 103(1-2): 53-65.
Hora, S.C, and R.L. Iman, 1986. A comparison of maximum/bounding and Bayesian/Monte

Carlo for fault tree uncertainty analysis. Sandia report SAND 85-2839.

Hordijk, L., 1993. In: Slobbe, L. Milieuverkenning beperkt zich tot evaluatie - ex-RIVM'er Hordijk mist zicht op aanvullend beleid. LT-journaal 12 (8 juli 1993), 11-12.

Hornberger, G.M. and R.C. Spear, 1980. Eutrophication in the Peel inlet - I. The problem defining behavior and a mathematical model for the phosphorus scenario. Water Research 14: 29-42.

Hough, M.N., 1990. Agrometeorological aspects of the crops in the United Kingdom and Ireland. A review for Sugar Beet, Oilseed Rape, Peas, Wheat, Barley, Oats, Potatoes, Apples and Pears. Luxembourg: Joint Research Centre, Commission of the European Community, 303 p.

Ibbitt, R.P. and T. O'Donnell, 1971. Fitting conceptual catchment models. Journal of the hydraulics division, Proceedings of the American society of civil engineers HY 9: 1331-1342.

Iman, R.L. and W.J. Conover, 1982. A distribution-free approach to inducing rank correlations among input variables. Communication in Statistics B11: 311-334.

Iman, R.L. and J.C. Helton, 1988. An investigation of uncertainty and sensitivity analysis techniques for computer models. Risk Analysis 8(1): 71-90.

Iman, R.L. and S.C. Hora, 1990. A robust measure of uncertainty importance for use in fault tree system analysis. Risk Analysis 10: 401-406.

Jackson, G.B., 1980. Methods for integrative reviews. Review of Educational Research 50(3): 438-460.

Jansen, A.A.M., J. Oude Voshaar, B. Keen, and J.A.P. Heesterbeek, 1995. Modelbouw en statistiek voor landbouw en milieu. Kwantitatieve Methoden 16(49): 61-68.

Jansen, M.J.W., W.A.H. Rossing, and R.A. Daamen, 1994. Monte Carlo estimation of uncertainty contribution from several independent multivariate sources. Pages 334-343 in: Grasman, J. and G. van Straten (Eds). Predictability and Nonlinear modelling in natural sciences and economics. Dordrecht: Kluwer.

Janssen, P.H.M., 1994. Assessing sensitivities and uncertainties in models; a critical evaluation. Pages 344-361 in: Grasman, J. and G. van Straten (Eds). Predictability and non-linear modelling in natural sciences and economics. Dordrecht: Kluwer.

Janssen, P.H.M. and P.S.C. Heuberger, 1995. Calibration of process-oriented models. Ecological Modelling 83: 55-66.

Johnston, P.R. and D.H. Pilgrim, 1976. Parameter optimization for watershed models. Water Resources Research 12(3): 477-486.

Jones, C.A., J.R. Kiniry, P.T. Dijke, D.B. Farmer, and D.C. Godwin, 1986. CERES-Maize: a simulation model of maize growth and development. College station: Texas A&M University press, 194 p.

Jong, J.B.R.M. de, 1980. Een karakterisering van de zonnestraling in Nederland. (Deel I, Deel II: bijlagen). Scriptie T.U. Eindhoven, 97+67 p.

Kabat, P., B. Marshall, and B.J. van den Broek, 1995. Comparison of simulation results and evaluation of parameterization schemes. Page. 439-501 in: Kabat, P., B. Marshall, B.J. van den Broek, J. Vos, and H. van Keulen (Eds). Modelling and parameterization of the soil-plant-atmosphere system; a comparison of potato growth models. Wageningen: Wageningen Press.

Kadane, J.B., J.M. Dickey, R.K. Winkler, W.S. Smith, and S.C. Peters, 1980. Interactive elicitation of opinion for a normal linear model. Journal of the American Statistical

Association 75(372): 845-854.

Karnieli, A.M., M.H. Diskin, and L.J. Lane, 1994. CELMOD5 - a semi distributed cell model for conversion of rainfall into runoff in semi-arid watersheds. Journal of Hydrology 157(1-4): 61-85.

Kaše, M. and J. Čatský, 1984. Maintenance and growth components of dark respiration rate in leaves of C_3 and C_4 plants as affected by leaf temperature. Biologia plantarum (Prague) 26(6): 461-470.

Keesman, K., 1989. On the dominance of parameters in structural models of ill-defined systems. Applied Mathematics and Computation 30:133-147.

Keulen, H. van, 1995. Dierlijke produktiesystemen: een poging tot integratie. Inaugurele rede, Landbouwuniversiteit Wageningen, 21 december 1995, 20 p.

Kim, C.P., 1995. The water budget of heterogeneous areas - impact of soil and rainfall variability. PhD-thesis, Wageningen Agricultural University, 182 p.

Kiniry, J.R., C.A. Jones, J.C. O'Toole, R. Blanchet, M. Cabelguenne, and D.A. Spanel, 1989. Radiation use efficiency in biomass accumulation prior to grain-filling for five grain-crop species. Field Crops Research 20: 51-64.

Klaassen, H.L., G. Ooms, J. Paelinck, and W. Willemsen, 1980. A study of the formal structure of J.W. Forrester's world dynamics model. Ecological Modelling 8: 259-274.

Kleijnen, J.P.C., 1987. Too many factors: screening. Chapter 28 in: Statistical tools for simulation practitioners. Statistics: textbooks and monographs vol. 76. New York: Marcel Dekker.

Kleijnen, J.P.C, 1995. Verification and validation of simulation models. European Journal of Operational Research 82: 145-162.

Klemeš, V., 1986. Dilettantism in hydrology: Transition or destiny? Water Resources Research 22(9): 177S-188S.

Klepper, O., 1989. A model of carbon flow in relation to macrobenthic food supply in the Oosterschelde estuary (S.W. Netherlands). PhD-thesis, Wageningen Agricultural University, 270 p.

Klepper, O. and D.I. Rouse, 1991. A procedure to reduce parameter uncertainty for complex models by comparison with a real system output illustrated on a potato growth model. Agricultural Systems 36: 375-395.

Klepper, O. and E.M.T. Hendrix, 1994. A method for robust calibration of ecological models under different types of uncertainty. Ecological Modelling 74: 161-182.

Kocabas, Z., R.A. Mitchell, J. Craigon, and J.N. Perry, 1993. Sensitivity analysis of the ARCWHEAT1 crop model: the effect of changes in radiation and temperature. Journal of Agricultural Science (Cambridge) 120: 149-158.

Konikow, L.F. and J.D. Bredehoeft, 1992. Ground water models cannot be validated. Advances in Water Resources 15: 75-83.

Koning, G.H.J. de, M.J.W. Jansen, E.R. Boons-Prins, C.A. van Diepen, and F.W.T. Penning de Vries, 1993. Crop growth simulation for regional yield forecasting across the European community. (Simulation reports CABO-TT 31) Wageningen: AB-DLO, 105 pp.

Kooman, P.L. and C.J.T. Spitters, 1995. A coherent set of modules to simulate potato growth. Pages 253-274 in: Kabat P., B. Marshall, B.J. van den Broek, J. Vos, and H. van Keulen

(Eds). Modelling and parameterization of the soil-plant-atmosphere system. A comparison of potato growth models. Wageningen: Wageningen pers.

Kowalik, J. and M.R. Osborne, 1968. Methods for unconstrained optimization problems. Modern analytic and computational methods in science and mathematics; a group of monographs and advanced textbooks; no. 12. New York: [-], 148 p.

Kraalingen, D.W.G. van, 1991. The FSE System for crop simulation. Simulation report CABO-TT no. 23. Wageningen: CABO, 77 p.

Kraalingen, D.W.G. van, W. Stol, P.W.J. Uithol, M. Verbeek, 1991. User manual of CABO/TPE Weather system. CABO/TPE internal communication. Wageningen: CABO, 27 p.

Kroon, L.J.M., 1985. Profile derived fluxes above inhomogeneous terrain: a numerical approach. PhD-thesis, Wageningen Agricultural University, 159 p.

Kropff, M.J., 1989. Quantification of SO_2 effects on physiological processes, plant growth and crop production. PhD-thesis, Wageningen Agricultural University, 201 p.

Kropff, M.J., 1996. Strategisch balanceren. Onkruidkunde als toegepaste plantenecologie. Inaugurele rede. Landbouwuniversiteit Wageningen, 18 april 1996, 34 p.

Kropff, M.J., H.H van Laar, and R.B. Matthews, 1994. ORYZA1. An ecophysiological model for irrigated rice production. Wageningen: SARP Research Proceedings. DLO-research institute for Agrobiology and soil fertility, 110 p.

Kropff, M.J., S.E. Weavers, L.A.P. Lotz, J.L. Lindquist, W. Joenje, B.J. Schnieders, N.C. van Keulen, T.R. Migo, and F.F. Fajardo, 1993. Understanding crop-weed interactions in field situations. Chapter 8 in: Kropff, M.J. and H.H. van Laar (Eds). Modelling crop-weed interactions, Wallingford: CAB international in cooperation with IRRI.

Kuzelka, I., 1990. Nitrate risk management. Page 489-507 in: Bardossy, A., I. Bogardi, and R.D. Kuzelka (Eds). Nitrate contamination: exposure, consequence, and control. Proceedings of the NATO Advanced Research Workshop NATO ASI Series G: Ecological Sciences, Volume 30. Berlin: Springer Verlag.

Landry, M. and M. Oral, 1993. In search of a valid view of model validation for operations research. European Journal of Operational Research 66: 161-167.

Lanen, H.A.J. van, 1991. Qualitative and quantitative physical land evaluation: an operational approach. PhD-thesis, Wageningen Agricultural University, 195 p.

Latesteijn, H.C. van, and R. Rabbinge, 1992. Scenario's als hulpmiddel voor beleidmakers. Grond voor keuzen - een toelichting. Spil 5: 12-16; Discussion: Spil 5: 17-39.

Lauciani, E. and A. Ponticiello, 1993. Photosynthetically active radiation (PAR): empirical determination and literature survey. FAO Agrometeorology series, working paper Nr. 7, 85 p.

Laviolette, M., J.W. Seaman jr., J.D. Barrett, and W.H. Woodall, 1995. A probabilistic and statistical view of fuzzy methods. Technometrics 37(5): 249-261; C/R 37(5): 262-292.

Lawlor, D.W., 1987. Photosynthesis: metabolism, control and physiology. Harlow: Longman, 262 p.

Leffelaar, P.A., 1987. Dynamics of partial anaerobiosis, denitrification, and water in soil: experiments and simulation. PhD-thesis, Wageningen Agricultural University, 117 p.

Leijnse, A., 1996. Modellering van grondwaterkwaliteit: zin en onzin. Inaugurele rede, Landbouwuniversiteit Wageningen, 28 maart 1996, 36 p.

Lenthe, J. van, 1993. A blueprint of ELI: A new method for eliciting subjective probability

distributions. Behavior Research Methods, Instruments & Computers 25(4): 425-433.
Leutscher, K.J., 1995. Operational management in pot plant production. PhD-thesis, Wageningen Agricultural University, 285 p.
Light, R.J. and D.B. Pillemer, 1984. Summing up - the science of reviewing research. Cambridge: Harvard University Press, 191 p.
Loague, K. and R.E. Green, 1991. Statistical and graphical methods for evaluating solute transport models: overview and application. Journal of Contaminant Hydrology 7: 51-73.
Loehle, C., 1987. Errors of construction, evaluation, and inference: a classification of sources of error in ecological models. Ecological Modelling 36: 297-314.
Lourens, P., 1984. The formalization of knowledge by specification of subjective probability distributions: an experimental approach. PhD-thesis, Groningen, the Netherlands, 111 p.
Lumb, F.E., 1964. The influence of cloud on hourly amounts of total solar radiation at the sea surface. Quarterly Journal of the Royal Meteorological Society 90: 43-56, C/R: 493-495.
Maas, S.J. and J.R. Dunlap, 1989. Reflectance, transmittance and absorptance of light by normal, etiolated and albino corn leaves. Agronomy Journal 81: 105-110.
Mayer, D.G. and D.G. Butler, 1993. Statistical validation. Ecological Modelling, 68: 21-32.
Mayer, D.G., M.A. Stuart, and A.J. Swain, 1994. Regression of real world data on model output: an appropriate overall test of validity. Agricultural Systems 45: 93-104.
McKay, M.D., R.J. Beckman, and W.J. Conover, 1979. A comparison of three methods for selecting values of input variables in the analysis of output from a computer code. Technometrics 21(2): 239-245.
McVoy, C.W., K.C. Kersebaum, B. Diekkrüger, and D. Söndgerath, 1995. Modelling of Geo-Biosphere processes - Validation of agroecosystem models. Papers presented at the international workshop, Braunschweig, Germany, 12-14 May 1993. Ecological Modelling 81.
Mederski, H.J., M.E. Miller, and C.R. Weaver, 1973. Accumulated heat units for classifying corn hybrid maturity. Agronomy Journal 65: 743-747.
Meinke, H., 1996. Improving wheat simulation capabilities in Australia from a cropping systems perspective. PhD-thesis, Wageningen Agricultural University, 270 p.
Mesplé, F., M. Troussellier, C. Casellas, and P. Legendre, 1996. Evaluation of simple statistical criteria to qualify a simulation. Ecological Modelling 88: 9-18.
Miglietta, F., 1992. Simulation of wheat ontogenesis. PhD-thesis, Wageningen Agricultural University, 124 p.
Miller, A.J., 1990. Subset selection in regression. Monographs on statistics and applied probability vol. 40. London: Chapman and Hall, 229 p.
Miller, D.R., G. Butler, and L. Bramall, 1976. Validation of ecological system models. Journal of Environmental Management 4: 383-401.
Monteith, J.L., 1981. Epilogue: themes and variations. Plant and Soil 58: 305- 309.
Monteith, J.L. and M.H. Unsworth, 1990. Principles of environmental physics. London: Arnold, 291 p.
Montgomery, D.C. and E.A. Peck, 1992. Introduction to linear regression analysis. Second edition. New York: Wiley, 527 p.
Morgan, J.N. and J.A. Sonquist, 1963. Problems in the analysis of survey data, and a proposal. Journal of the American Statistical Association 58: 415-434.

Morris, M.D., 1991. Factorial sampling plans for preliminary computational experiments. Technometrics 33-2: 161-174.

Mosteller, F., A.F. Siegel, E. Trapido, and C. Youtz, 1981. Eye fitting straight lines. American Statistician 35(30): 150-152.

Mous, S.L.J., 1993. Identification of the movement of water in unsaturated soil: the problem of identifiability of the model. Journal of Hydrology 143(1/2): 153-167.

Mous, S.L.J., 1994. On identification of non-linear systems. PhD-thesis, Wageningen Agricultural University, 129 p.

Muchow, R.C., 1994. Effect of nitrogen on yield determination in irrigated maize in tropical and subtropical environments. Field Crops Research 38: 1-13.

Nagel, E., 1971. The structure of science - problems in the logic of scientific explanation. 3rd impression, London: Routledge and Kegan Paul, 618 p.

Nance, R.E. and O. Balci, 1987. Simulation model management: objectives and requirements. Page 4328-4333 in: M.G. Singh (Ed.). Systems and control encyclopaedia: Theory, technology, applications. Oxford: Pergamon Press, 8 vols.

Nelder, J.A. and R. Mead, 1965. A simplex method for function minimization. Computer Journal 7: 308-313.

Nevers, N. de, 1973. Enforcing the Clean Air Act of 1970. Scientific American 228(6): 14-21, 122.

Nonhebel, S., 1993. The importance of weather data in crop growth simulation models and assessment of climatic change effects. PhD-thesis, Wageningen Agricultural University, 144p.

Nonhebel, S., 1994. Inaccuracies in weather data and their effects on crop growth simulation results. I. Potential production. Climate Research 4: 47-60.

O'Neill, R.V. and R.H. Gardner, 1979. Sources of uncertainty in ecological models. Page 447-463 in: Zeigler, B.P., M.S. Elzas, G.J. Klir, and T.I. Ören (Eds). Methodology in systems modelling and simulation. Amsterdam: North Holland Publishing.

Oenema, O., 1996. De kunst van het goed bemesten. Inaugurele rede, Landbouwuniversiteit Wageningen, 13 juni 1996, 26 p.

Olseth, J.A., A. Skartveith, 1993. Characteristics of hourly global irradiance modelled from cloud data. Solar Energy 51(3): 197-204.

Otegui, M.E., M.G. Nicolini, R.A. Ruiz, and P.A. Dodds, 1995. Sowing date effects on grain yield components for different maize genotypes. Agronomy Journal 87: 29-33.

Ouwerkerk, I., and H. Drenth, 1986. Verslag beregeningsonderzoek naar het gewas snijmais 1985. Interne mededeling nr. 439. Lelystad: PAGV.

Owen, A.B., 1992. A central limit theorem for Latin hypercube sampling. J. Royal Statistical Soc.-B 54(2): 541-551.

Palanisamy, S., F.W.T. Penning de Vries, S. Mohandass, T.M. Thiyagarajan, and A.A. Kareem, 1993. Simulation in pre-testing of rice genotypes in Tamil Nadu. Page 63-75 in: Penning de Vries, F.W.T., P. Teng, and K. Metselaar (Eds). Systems approaches for agricultural development. Dordrecht: Kluwer in cooperation with IRRI.

Patoine, M. and J.P. Fortin, 1992. Modification of information model for irrigation management. Canadian Agricultural Engineering 34(4): 305-317.

Paul, W. and E. Witte, 1995. Robust approaches to N-leaching under uncertainties. Pages V.B.1-

5 in: Mathematical modelling and simulation in agriculture and bio-industries, Volume III, IMACS/IFAC symposium, May 9-12, Brussels, Belgium.

Payne, R.W. and P.W. Lane (Eds), 1993. Genstat 5 - Release 3 Reference manual. Oxford: Clarendon Press, 796 p.

Penning de Vries, F.W.T. and C.J.T. Spitters, 1991. The potential for improvement in crop yield simulation. Page 123-140 in: Muchow R.C. and J.A. Bellamy (Eds). Climatic risk in crop production models and management for the semiarid tropics and subtropics. Wallingford: CAB-International.

Penning de Vries, F.W.T., R. Rabbinge, D.M. Jansen, and A. Bakema, 1988. Transfer of systems analysis and simulation in agriculture to developing countries. Agricultural Administration and Extension 29: 85-96.

Penning de Vries, F.W.T., D.M. Jansen, H.F.M. ten Berge, and A. Bakema, 1989. Simulation of ecophysiological processes of growth of several annual crops. Simulation Monographs 29. Wageningen: PUDOC, 271 p.

Philip, J.R., 1991. Soils, natural science, and models. Soil Science 151(1): 91-98.

Plucknett, D.L. and D.L. Winkelmann, 1995. Technology for sustainable agriculture. Scientific American 273(3): 148-152.

Power, M., 1993. The predictive validation of ecological and environmental models. Ecological Modelling 68: 33-50.

Press, W.H., S.A. Teukolsky, W.T. Vetterling, and B.P. Flannery, 1992. Numerical recipes in FORTRAN -The art of scientific computing. Second edition. Cambridge: Cambridge University Press, 963 p.

Price, W.L., 1979. A controlled random search procedure for global optimisation. The Computer Journal 20(4): 367-370.

Querner, E.P., 1993. Aquatic weed control within an integrated water management framework. PhD-thesis, Wageningen Agricultural University, 204 p.

Rabbinge, R. and H.C. van Latesteijn, 1992. Long term options for land use in the European Community. Agricultural Systems 40: 195-210.

Rabbinge, R., 1985. De brugfunctie van de gewasecologie. Inaugurele rede, Landbouwhogeschool Wageningen, 12 december 1985, 18 p.

Rabbinge, R., S.A. Ward, and H.H. van Laar (Eds), 1989. Simulation and systems management in crop protection. Simulation Monographs 32. Wageningen: PUDOC, 420 p.

Rao, C.R., 1973. Linear statistical inference and its applications. New York: Wiley, 625 p.

Rao, P.S.C., J.W. Jones, and G. Kidder, 1989. Development, validation and applications of simulation models for agro-ecosystems: problems and perspectives. Page 253-259 in: M. Clarholm and L. Bergström (Eds). Ecology of arable land. Dordrecht: Kluwer.

Rappoldt, C., 1992. Diffusion in aggregated soil. PhD-thesis, Wageningen Agricultural University, 204 p.

Rappoldt, C. and D.W.G. van Kraalingen, 1990. Reference manual of the FORTRAN utility library TTUTIL with applications. Simulation report CABO-TT nr. 20. Wageningen: CABO, 122 p.

Reckhow, K.H., 1988. A comparison of robust Bayes and classical estimators for regional lake models of fish response to acidification. Water Resources Research 24(7): 1061-1068.

Reed, K.L., K.A. Rose, and R.C. Whitmore, 1984. Latin hypercube analysis of parameter sensitivity in a large model of outdoor recreation demand. Ecological Modelling 24: 159-169.

Richter, O. and D. Söndgerath, 1990. Chapter 5 in: Parameter estimation in ecology - The link between data and models. Weinheim: VCH.

Ridder, N. de and H. van Keulen, 1997. Estimating biomass through transfer functions based on simulation model results: a case study for the Sahel. Agricultural Water Management 28: 57-71.

Rogers, P., 1978. On the choice of the 'appropriate model' for water resources planning and management. Water Resources Research 14(6): 1003-1010.

Rossing, W., 1991. Simulation of damage in winter wheat caused by the grain aphid Sitiobion avenae. 2. Construction and evaluation of a simulation model. Netherlands Journal of Plant Pathology 97: 25-54.

Rossing, W.A.H., R.A. Daamen, and M.J.W. Jansen, 1994a. Uncertainty analysis applied to supervised control of aphids and brown rust in winter wheat. Part 1. Quantification of uncertainty in cost-benefit calculations. Agricultural Systems 44: 419-448.

Rossing, W.A.H., R.A. Daamen, and M.J.W. Jansen, 1994b. Uncertainty of predictions in supervised pest control in winter wheat, its price and its causes. Pages 334-343 in: Grasman, J. and G. van Straten (Eds). Predictability and non-linear modelling in natural sciences and economics. Dordrecht: Kluwer.

Rötter, R., 1993. Simulation of the biophysical limitations to maize production under rainfed conditions in Kenya - Evaluation and application of the model WOFOST. Materialen zur Ostafrika-Forschung, Heft 12. PhD-thesis University Trier. 261 p.

Rykiel Jr., E.J., 1996. Testing ecological models: the meaning of validation. Ecological Modelling 90: 229-244.

Sage, A.P., 1987. Validation. Page 4996 in: Singh, M.G. (Ed.). Systems analysis and control encyclopedia: theory, technology, applications. Oxford: Pergamon.

Saltelli, A., T.H. Andres, and T. Homma, 1993. Sensitivity analysis of model output. An investigation of new techniques. Computational Statistics & Data Analysis 15: 211-238.

Saltelli, A., T.H. Andres, and T. Homma, 1995. Sensitivity analysis of model output. Performance of the iterated fractional factorial design method. Computational Statistics & Data Analysis 20: 387-407.

Sargent, R.G., 1984. Simulation model validation. Page 536-555 in: Ören, T.I. et al. Simulation and model-based methodologies: an integrative view.

Savage, L.J., 1971. Elicitation of personal probabilities and expectations. Journal of the American Statistical Association 66: 783-801.

Schans, D.A. van der, H.M.G. van der Werf, and W. van den Berg, 1993. Produktie- en kwaliteitsverloop bij snijmaïs - Verslag van onderzoek naar het optimale oogstijdstip bij snijmaïs (Production and quality of silage-maize during ripening - Report of research into the optimum time of harvest) (PAGV-verslag 155). Lelystad: PAGV, 75 p.

Scholte, K., 1987. Relationship between cropping frequency, root rot and yield of silage maize on sandy soils. Netherlands Journal Agricultural Science 35: 473-486.

Scholten, H. and M.W.M. van der Tol, 1994. SMOES: a simulation model for the Oosterschelde ecosystem. Part II: Calibration and validation. Hydrobiologia 282/283: 453-474.

Schouwenaars, J.M., 1990. Problem-oriented studies on plant-soil-water relations: Sowing strategies for maize in rainfed agriculture in Southern Mozambique. Water management in bog relicts in the Netherlands. PhD-thesis, Wageningen Agricultural University, 175 p.

Seber, G.A.F., and C.J. Wild, 1989. Nonlinear regression. New York: Wiley, 768p.

Sellers, P.J., 1989. Vegetation-canopy spectral reflectance and biophysical processes. Chapter 8 in: Asrar, G. (Ed.). Theory and applications of optical remote sensing. New York: Wiley, 734 p.

Shaeffer, D.L., 1980. A model evaluation methodology applicable to environmental assessment models. Ecological Modelling 8: 275-295.

Sheng, G.M., 1994. Software verification, model validation, and hydrogeologic modelling aspects in nuclear waste disposal system simulations - a paradigm shift. PhD- thesis, Wageningen Agricultural University, [295 p.]

Sheng, G.M., M.S. Elzas, T.I. Ören, and B.T. Cronhjort., 1993. Model validation: a systemic and systematic approach. Reliability Engineering and System Safety 42: 247-259.

Sibma, L., 1987. Ontwikkeling en groei van mais (*Zea mays L.*) onder Nederlandse omstandigheden. Wageningen: PUDOC, 57 p.

Sinclair, T.R. and C.T. de Wit, 1976. Analysis of the carbon and nitrogen limitations to soybean yield. Agronomy Journal 68: 319-324.

Smaling, E.M.A., 1993. An agro-ecological framework for integrated nutrient management - with special reference to Kenya. PhD-thesis, Wageningen Agricultural University, 250 p.

Smit, A.B., 1996. PIEteR: a field specific bio-economic production model for decision support in sugar beet growing. PhD-thesis, Wageningen Agricultural University, 201 p.

Smith, E.P. and K.A. Rose, 1995. Model goodness-of-fit analysis using regression and related techniques. Ecological Modelling 77: 49-64.

Sorooshian, S., 1991. Parameter estimation, model identification, and model validation: conceptual-type models. Chapter 20 in: Bowles, D.S. and P.E. O'Connell (Eds). Recent advances in the modeling of hydrological systems. Dordrecht: Kluwer.

Sorooshian, S. and V.K. Gupta, 1985. The analysis of structural identifiability: theory and application to conceptual rainfall-runoff models. Water Resources Research 21(4): 487-495.

Sorooshian, S., Q. Duan, and V.K. Gupta, 1993. Calibration of rainfall-runoff models: application of global optimization to the Sacramento Soil Moisture Accounting Model. Water Resources Research 29(4): 1185-1194.

Spear, R.C. and G.M. Hornberger, 1980. Eutrophication in Peel Inlet -II. Identification of critical uncertainties via generalized sensitivity analysis. Water Research 14: 43-49.

Spear, R.C., T.M. Grieb, and N. Shang, 1994. Parameter uncertainty and interaction in complex environmental models. Water Resources Research: 30(11): 3159-3169.

Spetzler, C.S. and C.A. Staël von Holstein, 1975. Probability encoding in decision analysis. Management Science 22(3): 340-358.

Spitters, C.J.T., 1986. Separating the diffuse and direct component of global radiation and its implications for modeling canopy photosynthesis. Part II. Calculation of canopy photosynthesis. Agricultural and Forest Meteorology 38: 231-242.

Spitters, C.J.T., 1990. Crop growth models: Their usefulness and limitations. Acta Horticulturae 267: 349-368.

Spitters, C.J.T. and A.H.C.M. Schapendonk, 1990. Evaluation of breeding strategies for drought tolerance in potato by means of crop growth simulation. Plant and Soil 123: 193-203.

Spitters, C.J.T., H.A.J.M. Toussaint, and J. Goudriaan, 1986. Separating the diffuse and direct component of global radiation and its implications for modeling canopy photosynthesis. Part I. Components of incoming radiation. Agricultural and Forestry Meteorology 38: 217- 229.

Spitters, C.J.T., H. van Keulen, and D.W.G. van Kraalingen, 1989. A simple and universal crop growth simulator: SUCROS87. Pages 147-181 in: Rabbinge, R., S.A. Ward, and H.H. van Laar (Eds). Simulation and systems management in crop protection. Simulation Monographs 32. Wageningen: PUDOC.

Steen, J.F.J. van, 1992. Elicitation of expert judgement on parameters in atmospheric dispersion models. Final report. (TNO-report ref nr. 91-325) Apeldoorn: TNO Institute of Environmental and Energy Technology, [140 p.]

Stein, M., 1987. Large sample properties of simulations using latin hypercube sampling. Technometrics 29(2): 143-151. C/R Technometrics 32: 367.

Stigter, C.J. and V.M.M. Musabilha, 1982. The conservative ratio of photosynthetically active to total radiation in the tropics. Journal of Applied Ecology 19: 853-858.

Stijn, Th. L. van, E.D. de Goede, and J. Dijkzeul, 1994. Kwaliteit als basis voor modelontwikkeling. H_2O 27(20): 599-601.

Stol, W., D.I. Rouse, D.W.G. van Kraalingen, and O. Klepper, 1992. FSEOPT, a FORTRAN program for calibration and uncertainty analysis of simulation models. Simulation reports CABO-TT 24. Wageningen: CABO-DLO, 24 p.

Straten, G. van, 1991. Heer en meester? Systeemtechnologie in landbouw en milieu. Inaugurele rede, Landbouwuniversiteit Wageningen, 7 februari 1991, 40 p.

Stroosnijder, L., 1976. Infiltratie en herverdeling van water in grond. PhD-thesis, Wageningen Agricultural University, 213 p.

Stroosnijder, L., 1991. Verdrinken en verdrogen. Inaugurele rede, Landbouwuniversiteit Wageningen, 2 mei 1991, 40 p.

Stutterheim, N.C., 1995. Towards higher nitrogen efficiency in European rice cultivation: a case study for the Camargue. PhD-thesis, Wageningen Agricultural University, 94 p.

Stuyt, L.C.P.M., 1992. The water acceptance of wrapped subsurface drains. PhD-thesis, Wageningen Agricultural University, 305 p.

Swaney, D.P., J.W. Jones, and J.W. Mishoe, 1986. A combined simulation-optimization approach for predicting crop yields. Agricultural Systems 20: 133-157.

Tadmor, H.N., A. Brieghet, I. Noy-Meir, R.W. Benjamin, and E. Eyal, 1975. An evaluation of the calibrated weight-estimate method for measuring production in annual vegetation. Journal of Range Management 28(1): 65-69.

Tarantola, A. 1987. Inverse problem theory: methods for data fitting and model parameter estimation. Amsterdam: Elsevier, 613 p.

Terlouw, P., 1989. Subjective probability distributions: a psychometric approach. PhD thesis, Groningen, the Netherlands, 139 p.

Theil, H., 1961. Economic forecasts and policy. North Holland: Amsterdam, 567 p.

Thornley, J.H.M. and I.R. Johnson, 1990. Plant and crop modelling. A mathematical approach to plant and crop physiology. Oxford: Clarendon Press, 669 p.

Uehara, G. and Tsuji, G.Y., 1991. The IBSNAT project. Page 505-513 in: Penning de Vries, F.W.T, P.S. Teng, and K. Metselaar (Eds). Systems approaches for agricultural development. Dordrecht: Kluwer.

Varlet-Grancher, C. and R. Bonhomme, 1982. Efficience de la conversion de l'énergie solaire par un couvert végétal. Acta Œcologica/Œcologia Plantarum 3(17)1: 3-26.

Varlet-Grancher, C., G. Gosse, M. Chartier, H. Sinoquet, R. Bonhomme, and J.-M. Allirand, 1989. Mise au point: rayonnement solaire absorbé ou intercepté par un couvert végétal. Agronomie 9: 419-439.

Velde, H.A. te, 1984. Constraints on maize production in northern latitudes. Chapter 36 in: Gallagher, E.J. (Ed.). Cereal production. Proceedings of the second international summer school in agriculture held by the Royal Dublin Society in co-operation with W. K. Kellogg Foundation. Butterworths in ass. with Royal Dublin Society.

Veldkamp, A., 1991. Quaternary river terrace formation in the Allier basin of France - a reconstruction based on sand bulk geochemistry and 3-D modelling. PhD-thesis, Wageningen Agricultural University, 172 p.

Verboom, J., 1996. Modelling fragmented populations: between theory and application in landscape planning. PhD-thesis, Wageningen Agricultural University, 118 p.

Vereecken, H., M. Vanclooster, M. Swerts, and J. Diels, 1991. Simulating water and nitrogen behaviour in soils cropped with winter wheat. Fertilizer Research 27: 233-243.

Verheul, M.J., 1992. Seedling growth of maize genotypes under chilling conditions. PhD-thesis, Eidgenössische Technische Hochschule Zürich, 101 p.

Verkenningscommissie toepassing modelbouw en simulatietechniek DLO, 1986. Aspecten en knelpunten bij modelbouw en simulatie in het onderzoek. Verslag van de verkenningscommissie toepassing modelbouw en simulatietechniek. Den Haag: DLO. Landbouw-economisch instituut, 26 p.

Versteeg, M.N. and H. van Keulen, 1986. Potential crop production prediction by some simple calculation method as compared with computer simulations. Agricultural Systems 19: 249-272.

Voet, H. van der and G.M.J. Mohren, 1994. An uncertainty analysis of the process-based growth model FORGRO. Forest Ecology and Management 69: 157-166.

VROM, 1989. Nationaal Milieubeleidsplan. Omgaan met risico's. De risicobenadering in het milieubeleid. Centrale directie voorlichting en externe betrekkingen, Van Alkemadelaan 85, 2597 AC Den Haag, 28 p.

Wainer, H., 1984. How to display data badly. American Statistician 38(2): 137-147.

Waldman, S.E. and R.W. Rickman, 1990. Programming techniques to expedite communication of scientific models. Agrononomy Journal 82: 356-359.

Wallach, D. and B. Goffinet, 1989. Mean squared error of prediction as a criterion for evaluating and comparing system models. Ecological Modelling 44: 299-306.

Walter, E., 1982. Identifiability of state space models; with aplications to transformation systems. Lecture notes in biomathematics. Berlin: Springer Verlag, 202 p.

Welch, W.J., R.J. Buck, J. Sacks, H.P. Wynn, T.J. Mitchell, and M.D. Morris, 1992. Screening, predicting, and computer experiments. Technometrics 34(1): 15-25.

Wetenschappelijke raad voor het regeringsbeleid (WRR), 1992. Grond voor keuzen - vier

perspectieven voor de landelijke gebieden in de Europese Gemeenschap. (Rapporten aan de regering). 's-Gravenhage: SDU, 210 p.

Whisler, F.D., B. Acock, D.N. Baker, R.F. Fye, H.F. Hodges, J.R. Lambert, E. Lemmon, J.M. McKinion, and V.R. Reddy, 1986. Crop simulation models in agronomic systems. Advances in Agronomy 40: 141-208.

Wijngaard, P.J.M., A.K. Bregt, and E.G.J. Bogaart, 1995. Verbetering software-kwaliteit kan leiden tot beter imago landbouwkundig onderzoek. Page 53-61 in: Engelbart, F., H. Hogeveen, T.P.H. Troquay, J.J.F. Wien (Eds). Voordrachten VIAS-symposium 1995: Informatica toepassingen in de agribusiness. (Agro-informaticareeks 9.)

Williams, W.A., R.S. Loomis, and C.R. Lepley, 1965. Vegetative growth of corn as affected by population density. I. Productivity in relation to interception of solar radiation. Crop Science 5: 211-215.

Wit, C.T. de, 1968. Theorie en model. Inaugurele rede, Landbouwhogeschool Wageningen, 10 oktober 1968, 13 p.

Wit, C.T. de, 1970. Dynamic concepts in biology. Page 17-23 in: I. Šetlik (Ed.). Prediction and measurement of photosynthetic productivity: Proceedings of the IBP/PP Technical Meeting Třeboň, 14-21 September 1969. Wageningen: PUDOC.

Wit, C.T. de, 1993. Philosophy and terminology. Pages 3-9 in: Leffelaar, P.A. (Ed.). On systems analysis and simulation of ecological processes - with examples in CSMP and FORTRAN. (Current issues in production ecology volume 1.) Dordrecht: Kluwer.

Wit, C.T. de, R. Brouwer, and F.W.T. Penning de Vries, 1970. The simulation of photosynthetic systems. Pages 47-70 in: Šetlik, I. (Ed.). Prediction and measurement of photosynthetic productivity. (Proceedings IBP -Plant production technical meeting Třeboň.) Wageningen: PUDOC.

Wit, C.T. de and F.W.T. Penning de Vries, 1985. Predictive models in agricultural production. Phil. Trans. R. Soc. Lond. B 310: 309-315.

Wit, C.T. de, J. Goudriaan, H.H. van Laar, F.W.T. Penning de Vries, R. Rabbinge, H. van Keulen, L. Sibma, and C. de Jonge, 1978. Simulation of assimilation, respiration and transpiration of crops. Simulation Monographs. Wageningen: PUDOC, 141 p.

Woldt, W., C. Dou, and I. Bogardi, 1995. Using fuzzy methods to consider imprecision in groundwater flow models. Page 203-212 in: Wagner, B.J., T.H. Illangasekare, and K.H. Jensen (Eds). Models for assessing and monitoring groundwater quality. Wallingford: IAHS Press.

Woolley, J.T., 1971. Reflectance and transmittance of light by leaves. Plant physiology 47: 656-662.

Wopereis, M.C.S., 1993. Quantifying the impact of soil and climate variability on rainfed production. PhD-thesis, Wageningen Agricultural University, 188 p.

Wösten, J.H.M., 1988. Using texture and other soil properties to predict unsaturated soil hydraulic functions. Soil Science Society of America Journal 52: 1762-1770.

Wösten, J.H.M., 1990. Use of soil survey data to improve simulation of water movement in soils. PhD-thesis, Wageningen Agricultural University, 103 p.

Yocum, C.S., L.H. Allen, and E.R. Lemon, 1964. Photosynthesis under field conditions VI. Solar radiation balance and photosynthetic efficiency. Agronomy Journal 56(3): 249-253.

Appendix 1 List of abbreviations used in SUCROS87

Different model elements (Type) are designated in the following way:

S	: State variables	I	: Input
R	: Rate variables	P	: Parameter
F, G	: Auxiliary calculations	L	: Logical switch
T	: Tabulated function	O	: Output variable

Acronym	Type	Explanation	Unit
AADVS	P	Parameter used to modify the argument of table AMDVS	[-]
AAMDVS	P	Parameter used to modify the response of table AMDVS	[-]
AAMTMP	P	Parameter used to modify the response of table AMTMP	[-]
ABDVS	P	Parameter used to modify the argument of partitioning tables	[-]
ACB	P	Parameter used to modify the response of the partitioning to cob	[-]
ACDVS	P	Parameter used to modify the argument of table SLA	[-]
ADDTMP	P	Parameter used to modify the argument of table AMTMP	[-]
ADVRR	P	Parameter introduced to allow for a non-linear effect of temperature on crop development rate DVR after anthesis	[-]
ADVRV	P	Parameter introduced to allow for a non-linear effect of temperature on crop development rate DVR before anthesis	[-]
ALV	P	Parameter used to modify the response of the partitioning to leaves	[-]
AMAX	F	Actual CO_2 assimilation rate at light saturation for individual leaves	kg (CO_2) ha^{-1} (leaf) h^{-1}
AMDVS	T	Factor accounting for effect of development stage (DVS) on AMX	[-]
AMTMP	T	Factor accounting for effect of daytime temperature (DDTMP) on AMX	[-]
AMX	P	Potential CO_2 assimilation rate at light assimilation for individual leaves	kg(CO_2) ha^{-1} (leaf) h^{-1}
ART	P	Parameter used to modify the response of the partitioning to root	[-]
ASH	P	Parameter used to modify the response of the partitioning to shoot	[-]
ASLA	P	Parameter used to modify the response of table SLA	[-]
ASO	P	Parameter used to modify the response of the partitioning to storage organs (grain)	[-]
ASRCB	P	Assimilate requirement for dry matter production of cob	kg(CH_2O) kg^{-1} (d.m.)
ASRLV	P	Assimilate requirement for dry matter production of leaves	kg(CH_2O) kg^{-1} (d.m.)
ASRQ	F	Assimilate (CH_2O) requirement for dry matter production	kg(CH_2O) kg^{-1} (d.m.)
ASRRT	P	Assimilate requirement for dry matter production of roots	kg(CH_2O) kg^{-1} (d.m.)
ASRSO	P	Assimilate requirement for dry matter production of storage organs	kg(CH_2O) kg^{-1} (d.m.)
ASRST	P	Assimilate requirement for dry matter production of stems	kg(CH_2O) kg^{-1} (d.m.)
ASSIM	F	Function calculating gross canopy assimilation rate	kg CO_2 ha^{-1} (ground) h^{-1}
AST	P	Parameter used to modify the response of the partitioning to stem	[-]
ASTRO	F	Function calculating maximum daily global radiation	[-]
ATRL	P	Limit of the ratio actual-extraterrestrial radiation below which the fraction diffuse radiation equals 1	[-]

Acronym	Type	Explanation	Unit
ATRM	P	Limit of the ratio actual-extraterrestrial radiation above which the fraction diffuse radiation less than 1	[-]
BABVS	P	Parameter used to modify the argument of table AMDVS	[-]
BAMDVS	P	Parameter used to modify the response of table AMDVS	[-]
BAMTMP	P	Parameter used to modify the response of table AMTMP	[-]
BCB	P	Parameter used to modify the response of the partitioning to cob	[-]
BCDVS	P	Parameter used to modify the argument of table SLA	[-]
BDDTMP	P	Parameter used to modify the argument of table AMTMP	[-]
BLV	P	Parameter used to modify the response of the partitioning to leaves	[-]
BSH	P	Parameter used to modify the response of the partitioning to shoot	[-]
BSLA	P	Parameter used to modify the response of table SLA	[-]
BST	P	Parameter used to modify the response of the partitioning to stem	[-]
DAVTMP	I	Daily average temperature	°C
DAY	S	Day number (1 January = 1)	d
DAYSOW	P	Sowing day, day number since 1 January	d
DDTMP	I	Daily average daytime temperature	°C
DELT	I	Time step of summation (= time step of model)	d
DEMERG	R	Daily increase in temperature sum for emergence	°C d^{-1}
DLYATR	P	Correction daily course atmospheric transmission (Spitters, 1986)	[-]
DLV	R	Death rate of leaves	kg(leaf) ha^{-1} d^{-1}
DS0	I	Daily maximum total solar radiation	J m^{-2} d^{-1}
DTEFF	R	Daily effective temperature for leaf area growth in juvenile phase	°C
DTGA	F/G	Daily total gross assimilation of the crop	kg(CO_2) ha^{-1} (ground) d^{-1}
DVR	R	Development rate	d^{-1}
DVRR	T	Function (DAVTMP) for DVR in pre-anthesis phase	d^{-1}
DVRV	T	Function (DAVTMP) for DVR in post anthesis phase	d^{-1}
DVS	S	Development stage of the crop	[-]
DVSJUV	P	Development stage of the crop after which juvenile stage ends; (cf. LAIJUV)	[-]
DVSKIL	P	Development stage after which leaf death due to low temperatures is possible	[-]
DVSRDR	P	Development stage after which leaf death accelerates	[-]
DVSSEN	P	Development stage of the crop where leaf death starts	[-]
EFF	P	Initial light use efficiency for individual leaves	(kg(CO_2)ha^{-1} (leaf) h^{-1})(J m^{-2}(leaf) s^{-1})
EMERG	S	Temperature sum for the crop to emerge	°C d
FCB	T	Fraction of shoot d.m. increase allocated to cob (DVS)	[-]
FLV	T	Fraction of shoot d.m. increase allocated to leaves (DVS)	[-]
FRPAR	P	Fraction of global radiation as photosynthetically active radiation	[-]
FSH	T	Fraction of total d.m. increase allocated to shoots (DVS)	[-]
FST	T	Fraction of shoot d.m. increase allocated to stems (DVS)	[-]
FRDFAA	P	Regression coefficient for direct - diffuse radiation partitioning	[-]

Acronym	Type	Explanation	Unit
FRDFBA	P	Regression coefficient for direct - diffuse radiation partitioning	[-]
FRDFBB	P	Regression coefficient for direct - diffuse radiation partitioning	[-]
FRDFCA	P	Regression coefficient for direct - diffuse radiation partitioning	[-]
FRDFCB	P	Regression coefficient for direct - diffuse radiation partitioning	[-]
GCB	R	Dry matter (d.m.) growth rate of cob	kg(cob) ha^{-1} (ground) d^{-1}
GLAI	R	Net growth rate of leaf area index	ha(leaf) ha^{-1} (ground) d^{-1}
GLV	R	D.m. growth rate of leaves	kg(leaves) ha^{-1} (ground) d^{-1}
GPHOT	F	Daily total gross assimilation	kg(CH_2O) ha^{-1} (ground) d^{-1}
GRT	R	D.m. growth rate of roots	kg(roots) ha^{-1} (ground) d^{-1}
GSH	F	D.m. growth rate of shoots	kg(shoot) ha^{-1} (ground) d^{-1}
GSO	R	D.m. growth rate of storage organs	kg(st. organ) ha^{-1} (ground) d^{-1}
GST	R	D.m. growth rate of storage organs	kg(stem) ha^{-1} (ground) d^{-1}
GTW	F/G	Total d.m. growth rate of crop	kg(total d.m.) ha^{-1} (ground) d^{-1}
IEMERG	L	Variable that indicates if emergence has taken place; in this description L_2	[-]
KDFTHE	P	Theoretical extinction coefficient (Spitters, 1986)	ha(ground) ha^{-1} (leaf)
KDIF	P	Extinction coefficient for diffuse PAR flux	ha(ground) ha^{-1} (leaf)
L_1	L	Switch checking whether or not the crop has been sown.	[-]
L_2	L	Switch checking emergence of the crop evaluated at the end of the day	[-]
L_3	L	Switch which flags whether the crop is mature or not. If the crop is mature the simulation is stopped.	[-]
L_4	L	Switch activating leaf weight senescence	[-]
L_5	L	Switch increasing senescence due to high temperatures	[-]
L_6	L	Switch activating senescence due to low temperatures	[-]
L_7	L	Switch activating leaf area growth proportional to leaf weight growth rate	[-]
L_8	L	Switch activating the crop development rate in the reproductive phase	[-]
LAINI	P	Extrapolated leaf area at field emergence	cm^2 $plant^{-1}$
LAI	S	Leaf area index	ha(leaf) ha^{-1} (ground)
LAII	F	Initial leaf area index	ha(leaf) ha^{-1} (ground)
LAJUV	F/G	Growth rate leaf area index of the crop in juvenile stage	ha(leaf) ha^{-1} (ground) d^{-1}
LAIJUV	P	Leaf area index of the crop after which juvenile stage ends; (cf. DVSJUV)	ha(leaf) ha^{-1} (ground)
LAMAT	F/G	Growth rate leaf area index of the crop after juvenile stage	ha(leaf) ha^{-1} (ground) d^{-1}
LAT	P	Latitude of the site	[degrees] [decimal minute]
MAINCB	P	Maintenance respiration coefficient of cob	kg(CH_2O) kg^{-1} (d.m.) d^{-1}
MAINLV	P	Maintenance respiration coefficient of leaves	kg(CH_2O) kg^{-1} (d.m.) d^{-1}
MAINRT	P	Maintenance respiration coefficient of roots	kg(CH_2O) kg^{-1} (d.m.) d^{-1}
MAINSO	P	Maintenance respiration coefficient of storage organs	kg(CH_2O) kg^{-1} (d.m.) d^{-1}
MAINST	P	Maintenance respiration coefficient of stems	kg(CH_2O) kg^{-1} (d.m.) d^{-1}
MAINT	F	Maintenance respiration (CH_2O) of the crop at actual temperature and development stage	kg(CH_2O) ha^{-1} (ground) d^{-1}
MASS	F	Reference level of maintenance respiration	kg(CH_2O) ha^{-1} (ground) d^{-1}

Acronym	Type	Explanation	Unit
MTEFF	F	Factor accounting for the effect of temperature on maintenance respiration	[-]
MNDVS	F	Factor accounting for effect of development stage on maintenance respiration	[-]
NPL	P	Plant density	plants m^{-2}
OGEM	P	Average projection of leaves on a horizontal surface for a spherical angle distribution (Goudriaan, 1988)	[-]
Q10	P	Factor reflecting the increase in maintenance respiration with a 10°C rise in temperature	[-]
RDD	I	Daily global radiation	J m^{-2} d^{-1}
RDR	F	Relative death rate of leaves	d^{-1}
RDRDV	F	Component of relative death rate of leaves	d^{-1}
RDRH	F	Component of relative death rate of leaves	d^{-1}
RDRL	F	Component of relative death rate of leaves	d^{-1}
RDRLT	F/G	Component of relative death rate of leaves	d^{-1}
RGRL	P	Relative growth rate during exponential leaf area growth	$(°C\ d)^{-1}$
SCV	P	Scattering coefficient of leaves for PAR	[-]
SHRDRL	P	Regression coefficient for early leaf death	$°C^{-1}$
SHRDRH	P	Regression coefficient for late leaf death	$°C^{-1}$
SLA	T	Specific leaf area of new leaves (DVS)	ha(leaf) kg^{-1} (ground)
TADRW	F	Total above-ground dry weight	kg(d.m.)ha^{-1} (ground)
TBSVRR	P	Base temperature in crop development rate after anthesis	°C
TBSVRV	P	Base temperature in crop development rate before anthesis	°C
TBSEM	P	Base temperature for crop emergence	°C
TBSJUV	P	Base temperature for juvenile leaf area growth	°C
TBSKIL	P	Temperature below which leaves are killed	°C
TBSRDR	P	Base temperature for senescence - leaf death	°C
TMMN	I	Daily minimum temperature	°C
TMMX	I	Daily maximum temperature	°C
TOTASS	F	Function calculating daily total gross assimilation	kg CO_2 ha^{-1} (ground) d^{-1}
TSDVRR	P	Temperature sum after anthesis, given TBSVRR	°C d
TSDVRV	P	Temperature sum to anthesis, given TBSVRV	°C d
TSEMER	P	Temperature sum until emergence	°C d
TSUMEM	S	Temperature sum after emergence determining juvenile leaf area growth	°C d
WCB	S	Dry matter weight of cob	kg(cob) ha^{-1}
WLVD	S	Dry matter weight of dead leaves	kg(dead leaves) ha^{-1}
WLVG	S	Dry matter weight of green leaves	kg(green leaves) ha^{-1}
WRT	S	Dry matter weight of roots	kg(roots) ha^{-1}
WSO	S	Dry matter weight of storage organs	kg(storage organs) ha^{-1}
WST	S	Dry matter weight of stems	kg(stems) ha^{-1}

Appendix 2 List of abbreviations used in LINTUL

Different model elements (Type) are designated in the following way:

S	: State variables	I	: Input
R	: Rate variables	P	: Parameter
F, G	: Auxiliary calculations	L	: Logical switch
T	: Tabulated function	O	: Output variable

Acronym	Type	Explanation	Unit
DAVTMP	I	Daily average temperature	°C
DAY	S	Day of year (1 January = 1)	d
DAYSOW	P	Sowing day, day number since 1 January	d
DEMERG	R	Daily increase in temperature sum for emergence	°C
DVLGTH	P	Duration of phase of decreasing interception in dimensionless development stage	[-]
DVR	R	Development rate of the crop	d^{-1}
DVRR	T	Function for DVR in post-anthesis phase	d^{-1}
DVRV	T	Function for DVR in pre-anthesis phase	d^{-1}
DVHALF	P	Development stage until 50% reduction of the fraction intercepted light	[-]
DVS	S	Development stage of the crop	[-]
EMERG	S	Temperature sum for crop to emerge	°C
FINT	F	Function reducing light use efficiency as a function of development stage	[-]
FINTI	P	Initial interception capacity per plant for definition cf. appendix 2	m^2 $plant^{-1}$
FINTL	F/G	Function reducing light use efficiency as a function of development stage	[-]
FINTS	F/G	Function reducing light use efficiency as a function of development stage	[-]
FRPAR	P	Fraction of photosynthetically active radiation in global radiation	[-]
FSH	T	Fraction of d.m. increase allocated to shoot (DVS)	[-]
FSO	T	Fraction of shoot d. m. increase allocated to grain (DVS)	[-]
GRT	R	Dry matter growth rate of roots	kg (roots) ha^{-1} (ground) d^{-1}
GSO	R	Dry matter growth rate of storage organs (grain)	kg (grain) ha^{-1} (ground) d^{-1}
GW	R	D.m. growth rate of aboveground organs (Total aboveground minus grain)	kg (d.m.) ha^{-1} (ground) d^{-1}
GTAW	R	Total aboveground growth rate of crop	kg (aboveground dry matter) ha^{-1} d^{-1}
L_1	L	Switch checking whether or not the crop has been sown	[-]
L_2	L	Switch checking emergence of the crop is evaluated at the end of the day	[-]
L_3	L	Switch which checks whether the crop is mature or not. If the crop is mature the simulation is stopped.	[-]
L_4	L	Switch which controls the evaluation of the decrease in the light use efficiency	[-]
LUE	P	Light use efficiency, conversion of intercepted PAR into dry matter.	kg (d.m.) MJ^{-1}
NPL	P	Plant density	plants m^{-2}
PAR	F	Conversion of global radiation in photosynthetically active radiation	[-]

Acronym	Type	Explanation	Unit
RDD	I	Daily total solar radiation	$J\ m^{-2}\ d^{-1}$
RI	P	Initial relative growth rate of interception capacity	[-]
TBASE	P	Base temperature used in crop development rate	°C
TBSEM	P	Base temperature for crop emergence	°C
TSDVR	P	Temperature sum to anthesis, given TBASE, in LINTUL assumed equal to temperature sum from anthesis to maturity.	°C d
TSEMER	P	Temperature sum until emergence, given TBSEM	°C d
TMMN	I	Daily minimum temperature	°C
TMMX	I	Daily maximum temperature	°C
W	S	Dry matter weight aboveground organs minus grain dry matter weight	kg (aboveground organs) ha^{-1}
WSO	S	Dry matter weight grain (storage organ)	kg (grain) ha^{-1}

Appendix 3 Listing of arguments of the rate calculations in SUCROS87

This appendix hierarchically describes the relations between all model elements used in SUCROS87. The structure in this appendix corresponds to the levels of detail as given in the description of the model (Section 2.2). The backslash (\) indicates that the following model elements are used at the next level of detail.
As an example Table A3.1 gives the arguments of the function used to calculate the state variable root weight S_WRT as:

S_WRT = S_WRT + L_2( ) * L_3( ) * R_GRT( )

To calculate the state variable S_WRT, L_2, L_3 and R_GRT are used. Table A3.1 shows which arguments are used in the calculation of L_2, L_3 and R_GRT. To calculate R_GRT, T_FSH and F_GTW are used; the arguments used in the calculation of these two elements is defined at level 2 of the description, and so on.

Basic equations model:

S_DAY = S_DAY + 1
S_EMERG = S_EMERG + L_1( ) * [1 - L_2( )] * R_DEMERG( )
S_DVS = S_DVS + L_2( ) * L_3( ) * R_DVR( )
S_TSUMEM = S_TSUMEM + L_2( ) * L_3( ) * R_DTEFF( )
S_WLVG = S_WLVG + L_2( ) * L_3( ) * [R_GLV( )-R_DLV( )]
S_WLVD = S_WLVD + L_2( ) * L_3( ) * R_DLV( )
S_WST = S_WST + L_2( ) * L_3( ) * R_GST( )
S_WSO = S_WSO + L_2( ) * L_3( ) * R_GSO( )
S_WRT = S_WRT + L_2( ) * L_3( ) * R_GRT( )
S_WCB = S_WCB + L_2( ) * L_3( ) * R_GCB( )
S_LAI = S_LAI + L_2( ) * L_3( ) * R_GLAI( )

Table A3.1 Description of the arguments of the function S_WRT.

First level	2nd level	3rd level	...	...	...	...
L_2	\S_EMERG					
	\P_TSEMER					
L_3	\S_DVS					
R_GRT	\T_FSH	\S_DVS				
	\F_GTW	\G_GTW	\F_GPHOT	\F_DTGA	\...	\S_DAY
	\	\	\	\	\...	\S_LAI
	\	\	\	\	\...	\I_RDD
	\	\	\	\	\...	\P_FRPAR
	\	\	\	\	\...	\P_DLYATR
	\	\	\	\	\...	\P_ATRL
	\	\	\	\	\...	\P_ATRM
	\	\	\	\	\...	\P_FRDFBA
	\	\	\	\	\...	\P_FRDFBB
	\	\	\	\	\...	\P_FRDFBCA
	\	\	\	\	\...	\P_FRDFCB
	\	\	\	\	\...	\P_OGEM
	\	\	\	\	\...	\P_KDFTHE
	\	\	\	\	\...	\P_LAT
	\	\	\	\	\...	\P_SCV
	\	\	\	\	\...	\P_AMX
	\	\	\	\	\...	\P_EFF
	\	\	\	\	\...	\P_KDIF
	\	\	\	\	\...	\T_AMDVS \S_DVS
	\	\	\	\	\...	\T_AMTMP \I_DDTMP
	\	\	\F_MAINT	\F_MASS	\S_WLVG	
	\	\	\	\	\S_WLVD	
	\	\	\	\	\S_WST	
	\	\	\	\	\S_WRT	
	\	\	\	\	\P_MAINLV	
	\	\	\	\	\P_MAINST	
	\	\	\	\	\P_MAINRT	
	\	\	\	\	\P_MAINCB	
	\	\	\	\	\P_MAINSO	
	\	\	\	\	\S_WSO	
	\	\	\	\	\S_WCB	
	\	\	\	\F_MTEFF	\I_DAVTMP	
	\	\	\	\	\P_Q10	
	\	\	\	\F_MNDVS	\S_WLVG	
	\	\	\	\	\S_WLVD	
	\	\	\F_ASRQ	\P_ASRLV		
	\	\	\	\P_ASRST		
	\	\	\	\P_ASRSO		
	\	\	\	\P_ASRCB		
	\	\	\	\P_ASRRT		
	\	\	\	\T_FSH	\S_DVS	
	\	\	\	\T_FLV	\S_DVS	
	\	\	\	\T_FST	\S_DVS	
	\	\	\	\T_FCB	\S_DVS	

Rate equations model:

```
R_DEMERG \I_DAVTMP
         \P_TBSEM

R_DVR    \L_8        \S_DVS
         \T_DVRV     \I_DAVTMP
         \T_DVRR     \I_DAVTMP

R_DTEFF  \I_DAVTMP
         \P_TBSJUV
         \L_7        \S_DVS
         \           \S_LAI
         \           \P_DVSJUV
         \           \P_LAIJUV

R_DLV    \S_WLVG
         \L_4        \S_DVS
         \           \P_DVSSEN
         \F_RDR      \F_RDRV     \L_5        \S_DVS
         \           \           \           \P_DVSRDR
         \           \           \F_RDRL     \I_DAVTMP
         \           \           \           \P_SHRDRL
         \           \           \           \P_TBSRDR
         \           \           \F_RDRH     \I_DAVTMP
         \           \           \           \P_SHRDRH
         \           \           \           \P_TBSRDR
         \           \F_RDRLT    \L_6        \S_DVS
         \           \           \           \P_DVSKIL
         \           \           \G_RDRLT    \I_DAVTMP
         \           \           \           \P_TBSKIL

R_GLAI   \F_GLA      \L_7        \S_DVS
         \           \           \S_LAI
         \           \           \P_DVSJUV
         \           \           \P_LAIJUV
         \           \F_LAJUV    \R_DTEFF    \.....
         \           \           \G_LAJUV    \S_TSUMEM
         \           \           \           \P_RGRL
         \           \           \           \F_LAII     \P_NPL
         \           \           \           \           \P_LAINI
         \           \F_LAMAT    \R_GLV      \.....
         \           \           \R_DLV      \.....
         \           \           \T_SLA      \S_DVS
```

Remark: Calculations replaced by the dots (....) refer to other rates and a logical defined below.

Rate equations model continued:

```
R_GLV   \T_FLV    \S_DVS
        \T_FSH    \S_DVS
        \F_GTW    \G_GTW    \F_GPHOT   \F_DTGA   \...    \S_DAY
        \         \         \          \         \...    \S_LAI
        \         \         \          \         \...    \I_RDD
        \         \         \          \         \...    \P_FRPAR
        \         \         \          \         \...    \P_DLYATR
        \         \         \          \         \...    \P_ATRL
        \         \         \          \         \...    \P_ATRM
        \         \         \          \         \...    \P_FRDFBA
        \         \         \          \         \...    \P_FRDFBB
        \         \         \          \         \...    \P_FRDFBCA
        \         \         \          \         \...    \P_FRDFCB
        \         \         \          \         \...    \P_OGEM
        \         \         \          \         \...    \P_KDFTHE
        \         \         \          \         \...    \P_LAT
        \         \         \          \         \...    \P_SCV
        \         \         \          \         \...    \P_AMX
        \         \         \          \         \...    \P_EFF
        \         \         \          \         \...    \P_KDIF
        \         \         \          \         \...    \T_AMDVS \S_DVS
        \         \         \          \         \...    \T_AMTMP \I_DDTMP
        \         \         \F_MAINT   \F_MASS   \S_WLVG
        \         \         \          \         \S_WLVD
        \         \         \          \         \S_WST
        \         \         \          \         \S_WRT
        \         \         \          \         \P_MAINLV
        \         \         \          \         \P_MAINST
        \         \         \          \         \P_MAINRT
        \         \         \          \         \P_MAINCB
        \         \         \          \         \P_MAINSO
        \         \         \          \         \S_WSO
        \         \         \          \         \S_WCB
        \         \         \          \F_MTEFF  \I_DAVTMP
        \         \         \          \         \P_Q10
        \         \         \          \F_MNDVS  \S_WLVG
        \         \         \          \         \S_WLVD
        \         \         \F_ASRQ    \P_ASRLV
        \         \         \          \P_ASRST
        \         \         \          \P_ASRSO
        \         \         \          \P_ASRCB
        \         \         \          \P_ASRRT
        \         \         \          \T_FSH    \S_DVS
        \         \         \          \T_FLV    \S_DVS
        \         \         \          \T_FST    \S_DVS
        \         \         \          \T_FCB    \S_DVS
```

Remark: F_GTW is also used in the calculation of R_GRT, R_GST, R_GSO and R_GCB. It did not seem necessary to repeat this structure, so in the cases where this structure occurs again, this is indicated by F_GTW ..... , where the above structure should be substituted.

Rate equations model continued:

```
R_GST   \T_FSH    \S_DVS
        \T_FST    \S_DVS
        \F_GTW    \.....

R_GSO   \T_FSH    \S_DVS
        \T_FLV    \S_DVS
        \T_FST    \S_DVS
        \T_FCB    \S_DVS
        \F_GTW    \.....

R_GRT   \T_FSH    \S_DVS
        \F_GTW    \.....

R_GCB   \T_FSH    \S_DVS
        \T_FCB    \S_DVS
        \F_GTW    \.....
```

Switches:

```
L_1     \S_DAY
        \P_DAYSOW
L_2     \S_EMERG
        \P_TSEMER
L_3     \S_DVS
        \2
```

Appendix 4 Listing of arguments of the rate calculations in LINTUL

This appendix hierarchically describes the relations between all model elements used in LINTUL. The structure in this appendix corresponds to the levels of detail as given in the description of the model (Section 2.3). The backslash (\) indicates that the following model elements are used at the next level of detail.

Basic equations model:

```
S_DAY    = S_DAY +1
S_EMERG = S_EMERG + L_1( ) * [1 – L_2( )] * R_DEMERG( )
S_DVS    = S_DVS + L_2( ) * L_3( ) * R_DVR( )
S_W      = S_W + L_2( ) * L_3( ) * R_GW( )
S_WSO    = S_WSO + L_2( ) * L_3( ) * R_GSO( )
```

Rate equations model:

```
R_DEMERG \I_DAVTMP
         \P_TBSEM

R_DVR    \I_DAVTMP
         \P_TBASE
         \P_TSDVR

R_GW     \T_FSO     \S_DVS
         \F_GTAW    \P_LUE
         \          \F_PAR      \P_FRPAR
         \          \           \I_RDD
         \          \F_FINT     \F_FINTL\P_FINTI
         \          \           \          \P_NPL
         \          \           \          \G_FINTL   \P_FINTI
         \          \           \          \          \P_RI
         \          \           \          \          \S_DVS
         \          \           \F_FINTS   \L_5       \P_DVHALF
         \          \           \          \          \S_DVS
         \          \           \          \G_FINTS   \P_DVHALF
         \          \           \          \          \P_DVLGTH
         \          \           \          \          \S_DVS
```

R_GSO is calculated as complement to R_GW; it depends on the same model elements.

Switches:

```
L_1      \I_DAY
         \P_DAYSOW

L_2      \S_EMERG
         \P_TSEMER

L_3      \S_DVS

L_4      \S_DVS
         \P_DVHALF
```

Appendix 5 Default values for parameters in SUCROS87

AMX	= 70	KDIF	= 0.6
ATRL	= 0.22	KDFTHE	= 0.8
ATRM	= 0.35	LAINI	= 6.69
ASRSO	= 1.49	LAIJUV	= 0.75
ASRLV	= 1.463	LAT	= 52.
ASRST	= 1.513	MAINSO	= 0.01
ASRRT	= 1.444	MAINLV	= 0.03
ASRCB	= 1.47	MAINST	= 0.015
DAYSOW	= 119.	MAINRT	= 0.015
DLYATR	= 0.4	MAINCB	= 0.02
DVSJUV	= 0.3	NPL	= 11.13
DVSKIL	= 1.25	OGEM	= 0.5
DVSRDR	= 1.35	Q10	= 2.
DVSSEN	= 1.	RGRL	= 0.0294
EFF	= 0.45	SCV	= 0.2
FRPAR	= 0.5	SHRDRH	= 0.0030
FRDFBA	= 1.47	SHRDRL	= 0.0005
FRDFBB	= 1.66	TBSEM	= 3.
FRDFCA	= 0.15	TBSJUV	= 10.
FRDFCB	= 0.1	TBSKIL	= 6.
TSEMER	= 120.	TBSRDR	= 8.

AMDVST	= 0., 1.0; 1.3, 1.0; 1.6, 0.5; 2.0, 0.25; 2.5, 0.25
AMTMPT	= –30., 0.01; 9., 0.05; 16., 0.80; 18., 0.94; 20., 1.0; 30., 1.0; 40., 0.75
DVRVT	= –30., 0.0; 10., 0.0; 30., 0.0471
DVRRT	= –30., 0.0; 10., 0.0; 30., 0.0471
SLAT	= 0., 0.004; 0.7, 0.001; 2.5, 0.001
FSHTB	= 0.0, 0.600; 0.1, 0.630; 0.2, 0.660; 0.3, 0.690; 0.4, 0.730; 0.5, 0.770; 0.6, 0.810; 0.7, 0.850; 0.8, 0.900; 0.9, 0.940; 1.0, 1.000; 2.5, 1.000
FLVTB	= 0.0, 0.7; 0.25 0.7; 0.80, 0.15; 0.95, 0.0; 2.5, 0.0
FSTTB	= 0.0, 0.3; 0.25, 0.3; 0.80, 0.85; 0.95, 0.45; 1.1, 0.0; 2.5, 0.0
FCBTB	= 0.0, 0.0; 0.80, 0.0; 0.95, 0.55; 1.1, 1.0; 1.2, 0.0; 2.5, 0.0

Appendix 6 Default values for parameters in LINTUL

The parameters in LINTUL are partly based on SUCROS87' parameters, and partly based on the literature review in Chapter 4.

FRPAR	= 0.5	TSEMER	= 80.
RI	= 10.86	LUE	= 1.868
DVLGTH	= 2.257	TBASE	= 10.
DVHALF	= 1.894	TSDVR	= 425.
TBSEM	= 10.	FINTI	= 0.00008024

DVRVT	= –30., 0.0; 10., 0.0; 30., 0.0471
DVRRT	= –30., 0.0; 10., 0.0; 30., 0.0471
FSOTB	= 0.0, 0.0; 1.1, 0.0; 1.2, 1.0; 2.5, 1.0
FSHTB	= 0.0, 0.600; 0.1, 0.630; 0.2, 0.660; 0.3, 0.690; 0.4, 0.730; 0.5, 0.770; 0.6, 0.810; 0.7, 0.850; 0.8, 0.900; 0.9, 0.940; 1.0, 1.000; 2.5, 1.000

Appendix 7 Parameterizing tabulated functions in SUCROS87 and LINTUL

The crop growth simulation model SUCROS87 contains tabulated functions, which cause problems in parameter calibration procedures. Some possibilities to modify these tables are proposed. Transformation of the function argument and the function response is proposed as a simple and robust method to modify tabulated functions using a small number of parameters. The problem of overparameterization is investigated for the partitioning functions used in SUCROS87.

Introduction

In SUCROS87 functions used in the calculation of rates are defined in various ways. Some functions are defined by tables, e.g. the dry matter partitioning functions and the dynamic specific leaf area. The use of these tables in the model requires a function (subroutine) for interpolation. In SUCROS87 a linear interpolation function is used. A complete overview of the tabulated functions used in SUCROS87 is given in Table A7.1.

Table A7.1 Overview of all tabulated functions used in SUCROS87

Tabulated function	Short characterization
F_X(DVS)	Partitioning functions
SLA(DVS)	Specific leaf area
AMDVS(DVS)	Photosynthesis reduction function due to crop age
DVR(DAVTMP)	Development rate
AMTMP(DDTMP)	Photosynthesis reduction function due to temperature

Tabulated functions are not very practical in calibration. Modification of complete tables by hand is laborious due to the large number of parameters; and although single values are easily calibrated both by hand and in other parameter estimation procedures, calibration of the complete functions by parameter estimation procedures is hardly possible without a parameterized description.

Some possible approaches and the approach chosen

There are three approaches to this problem. The first approach is to replace the table by a

function based on theory; the second approach is to replace the table by a descriptive function. In a third approach the table is retained as input but modified within the model by transformation of the in- and output of the interpolating function. We are pushing and pulling the table into a different shape within the model calculations, while this approach requires less parameters than the modification of each table entry individually.

In this appendix we will elaborate the transformation approach, which can be implemented in such a way that the model itself is not modified in any way. The approach has the additional advantage that it can be applied to all tabulated functions in a uniform way. A disadvantage of this approach is that discontinuities in partial derivatives of state variables with respect to parameters are not removed. This disadvantage does not occur when fitting a smooth function to the tables.

Below we will introduce three often-used transformations. The basic idea is to transform the original scale to the scale of real numbers and apply a linear transformation. The simplest transformation is a linear transformation; more complex transformations are necessary to ensure that the variable remains larger than zero and that the transformation is monotonous or that the requirements for a monotonous transformation of a fraction (between zero and one) are met. The transformation of partitioning tables will be treated as an additional special case.

Transformation of the scale of real numbers

The definition of the transformation is given as:

$$z^* = az + b$$

subject to $a > 0$. This condition ensures that transformation of any two values z_1, z_2 with $z_1 > z_2$ yields $z_1^* > z_2^*$. This transformation of function argument and function values does not require rescaling if there are no restrictions on x and/or y axis values.

Transformations of positive scales

This transformation is defined by:

$$z^* = bz^a$$

subject to $a,b > 0$. This condition ensures that transformation of any two values z_1, z_2 with $z_1 > z_2$ yields $z_1^* > z_2^*$. Furthermore the transformed value of zero is just zero. When a and b equal 1, the transformation is the identity transformation $z = z^*$. This transformation does not require rescaling if the x and y-values are restricted to values greater than zero. Note that the transformation is linear on logarithmic scale:

$$\ln z^* = a \ln z + \ln b$$

Transformation of fractions

This transformation has the property that when the original argument z is a fraction, the resulting transformed value z^* is also a fraction. The transformation is defined by:

$$z^* = \frac{b\left(\frac{z}{1-z}\right)^a}{1+b\left(\frac{z}{1-z}\right)^a}$$

subject to $a,b > 0$. This condition ensures that transformation of any two values z_1, z_2 with $z_1 > z_2$ yields $z_1^* > z_2^*$. Furthermore the transformed value of zero is just zero. When a and b equal 1, the transformation is the identity transformation $z = z^*$. This transformation does not require rescaling if the arguments and function values are fractions. Note that the transformation is linear on a logit scale:

$$\text{logit}\, z^* = \log\frac{z^*}{1-z^*} = \log\left(b\left(\frac{z}{1-z}\right)^a\right) = a\log\frac{z}{1-z} + \log b = a\,\text{logit}\, z + \log b$$

Transformation of a partitioning function

The partitioning functions in SUCROS87 are defined on a specific scale. The responses are positive numbers F_i, which sum to one:

$$\sum_{i=1}^{n} F_i = 1$$

For instance, the partitioning function for the aboveground material (shoot) and function for the fraction belowground (root) must sum to one. The partitioning functions within the aboveground material (leaves, stem, cob and storage organs) should also sum to one. Additionally we would like the transformation of the response to have the property that zero remains zero and one remains one. The transformation chosen is a modification of the transformation for a positive scale:

$$F_i^* = \frac{b_i \cdot F_i^{a_i}}{\sum_{j=1}^{n} b_j \cdot F_j^{a_j}}$$

which ensures that

$$\sum_{i=1}^{n} F_i = 1$$

Parameters a_i and b_i transform the y-axis and should be positive. Under these conditions transformation of any two values F_{i1} and F_{i2} say, where $F_{i1} > F_{i2}$ yields $F_{i1}^* > F_{i2}^*$. When all a_i and b_i are equal to one the original partitioning is reproduced. In Table A7.2 an overview of proposed transformations of the tabulated functions used in SUCROS87 is given.

Table A7.2 Tabulated functions in SUCROS87 and the proposed transformation of argument and response in view of scale properties.

Tabulated function	Properties of scale transformation	
	Argument	Response
F_X(DVS)	positive	partitioning
SLA(DVS)	positive	positive
AMDVS(DVS)	positive	fraction
DVR(DAVTMP)	real	positive
AMTMP(DDTMP)	real	fraction

Some additional considerations

In some cases specific conditions have to be met which require modifications or additional transformations. For some crops the formation of certain organs is not allowed after a certain development stage: grain filling in maize, for instance, only occurs after female flowering. Development of a crop in SUCROS87 after emergence is determined by the development stage; a dimensionless state variable, which is zero at emergence, one at flowering and two at physiological maturity. DVS is expressed as a variable between 0 and 2. To ensure that function values at DVS = 0, 1 and 2 are not modified the following transformation is useful:

$$z = \frac{\mathrm{DVS}}{2}$$

$$z^* = \frac{(\frac{z}{1-z})^a}{1+(\frac{z}{1-z})^a}$$

This is a special case of the transformation of a fraction where we use the property that $z^*(1/2) = 1/2$, under the condition that $b = 1$.

The function DVR(T), where T is the daily average temperature can be used as a simple example to illustrate some possible effects in parameterization. The tabulated function is a

simple piecewise linear function which can be written as:

$$\mathrm{DVR} = \begin{cases} 0 & T < T_B \\ A(T - T_B) & T \geq T_B \end{cases}$$

A two parameter linear transformation on temperature would yield the following result:

$$\mathrm{DVR} = \begin{cases} 0 & T < T_B \\ A(b_x T + c_x - T_B) & T \geq T_B \end{cases}$$

This equation can be simplified to yield:

$$\mathrm{DVR} = \begin{cases} 0 & T < T_B \\ A^*(T - T_B^*) & T \geq T_B \end{cases}$$

A two parameter positive scale transformation of the response yields:

$$\mathrm{DVR}^* = \begin{cases} 0 & T < T_B^* \\ b_y A^*(T - T_B^*)^{a_y} & T \geq T_B^* \end{cases}$$

This equation can also be described by three parameters:

$$\mathrm{DVR}^{**} = \begin{cases} 0 & T < T_B^* \\ A^{**}(T - T_B^*)^{a_y} & T \geq T_B^* \end{cases}$$

In this example, if we had not tried to write the table as a function, four parameters would have been introduced to calibrate the table. The above analysis shows that three are sufficient. Obviously transforming tables is not always efficient in terms of the number of parameters used. This effect - using four parameters, where three would suffice - is called overparameterization. In this case it can be prevented because the table is easily written as a function.

Modification of partitioning tables

As shown in the simple example above the effects of the parameters introduced in the calibration of tabulated functions may be very similar. In that case it is undesirable to use all parameters. In the following we will consider the partitioning functions and analyse whether overparameterization occurs and how the number of parameters in the partitioning functions can be reduced.

In Table A7.3 the structure of the partitioning in SUCROS87 is presented. In the case of SUCROS87 fourteen parameters (six partitioning functions with two parameters each and an

additional two to transform the argument development stage (DVS)) have to be introduced for the full transformation of all partitioning functions. The equation describing the full transformation is:

$$F_i^* = \frac{b_i (F_i[\beta\, x^{\alpha}])^{a_i}}{\sum_{j=1}^{n} b_j (F_j[\beta x^{\alpha}])^{a_j}}$$

where also

$$\sum_{i=1}^{n} F_i^* = 1$$

A first reduction in the number of parameters is possible as two groups of partitioning functions (Table A7.3: partitioning -1 and -2) sum to one. This condition allows to choose two parameters b_i modifying the function value as equal to one, as in each of these groups one partitioning can be calculated from the other partitioning functions. For maize partitioning to the leaves is not allowed after male flowering; for the modification of the argument DVS the transformation of a fraction under the condition $\beta = 1$ is necessary, further reducing the number of parameters.

After these considerations eleven parameters would then parameterize the partitioning tables. A further simplification seems possible as for small changes the effects of the transformation using 'a' are similar to the linear transformation using 'b'-parameters. That the effects are similar can be shown if the derivatives of the partitioning functions at a_i, $b_i = 1$ are calculated:

$$\frac{\partial F_i^*}{da_j} = \begin{cases} \left(F_i - F_i^2\right) \ln F_i & i = j \\ -F_i F_j \ln F_j & i \neq j \end{cases}$$

$$\frac{\partial F_i^*}{db_j} = \begin{cases} F_i - F_i^2 & i = j \\ -F_i F_j & i \neq j \end{cases}$$

These equations show that the sensitivity of the parameters 'a_j' and 'b_j' are very similar.

Table A7.3 The structure of the partitioning functions in SUCROS87.

Partitioning-1	Partitioning-2
FSH(DVS)	FLV(DVS)
	FST(DVS)
	FCB(DVS)
	FSO(DVS)
FRT(DVS)	

To check the possibilities to reduce the number of parameters necessary to calibrate the partitioning tables a sensitivity analysis for eleven parameters (1% change, meterological year 1985, default parameters) was executed.

To approximate the parameter effect in a real calibration the partitioning functions were evaluated every fourteen days during the growing season that year. The difference quotients with respect to the parameters were calculated; these can be regarded as the effect of a parameter on the partitioning tables. The effects of two parameters were regarded as similar if the difference quotients at given times for a given partitioning were similar. The similarity measure used was the cosine of the angle between the vectors defined by the effects of the parameters. If the angle is zero or 180 degrees, the cosine equals one and the similarity is then maximal; the similarity is minimal when the angle equals ninety degrees and the cosine equals zero.

The pairwise comparison of the effects for the eleven parameters is given in Table A7.4 as cosine values. The effect of the parameter modifying the partitioning to the storage organs was zero and is left out. This was due to the sampling schedule, as the parameter is only active over a short period of time. The results show that the effects of the pairs of parameters bsh-art, ast-blv and ast-bst are notably similar. This would suggest that from these pairs only one parameter should be chosen. However, there are two other options to further reduce the number of parameters used:

1) only use the 'b_i' parameters,

2) only use the 'a_i' parameters.

These two strategies would exclude parameter pairs with the most similar effects on the partitioning function. The results of the first option are presented in Table A7.5; the results of the second are presented in Table A7.6. It is seen that the pair bst-blv (Table A5.5) and the pair ast-alv (Table A5.6) still cause similar changes. A third possibility is to use the 'b_i' parameters but additionally use a single a_i value for each group of organs whose partitioning functions sum to one.

This last option was investigated using a sensitivity analysis, with all '*b*'-parameters and two '*a*'-parameters. One '*a*'-parameter ('ashrt') modifies the partitioning functions FSH and FRT; the second parameter ('atag') modifies the partitioning functions FCB, FLV, FST and FSO. Large values of parameter '*a*' will increase differences between partitioning functions, while they are decreased for small values of the parameter '*a*'. This statement is better understood if a partitioning between two organs in case of very large '*a*' and in case of very small '*a*' is considered. The results of the analysis are presented in Table A7.7. Again specific pairs of parameters cause similar effects.

The analyses allow to conclude that a clear-cut advice which parameters to use in calibration is not possible; perfect similarity which would justify such a conclusion does not occur. Other criteria, such as interpretability, the number of parameters necessary and a pragmatic argument derived from the quality of the calibration result will be necessary to make a final choice. The interpretability of the effects of the '*b*'-parameters and the small number of parameters necessary would be an argument in favour of this parameterization.

Table A7.4 Similarity of effects of parameters on partitioning functions; pairwise comparison.

	abdvs	art	bsh	ash	blv	alv	bst	ast	bcb	acb
abdvs	1									
art	0.14	1								
bsh	0.10	0.72	1							
ash	0.04	0.32	0.33	1						
blv	0.18	0	0	0	1					
alv	0.13	0	0	0	0.42	1				
bst	0.17	0	0	0	0.68	0.6	1			
ast	0.16	0	0	0	0.83	0.64	0.85	1		
bcb	0.01	0	0	0	0	0	0.11	0.07	1	
acb	0.00	0	0	0	0	0	0.05	0.03	0.14	1

Table A7.5 Similarity of changes in partitioning functions caused by '*b*'-parameters.

	bsh	blv	bst	bcb
bsh	1			
blv	0	1		
bst	0	0.68	1	
bcb	0	0	0.11	1

Table A7.6 Similarity of changes in partitioning functions caused by '*a*'-parameters.

	abdvs	art	ash	alv	ast	acb
abdvs	1					
art	0.14	1				
ash	0.04	0.32	1			
alv	0.13	0	0	1		
ast	0.16	0	0	0.64	1	
acb	0.00	0	0	0	0.03	1

Table A7.7 Similarity of parameter effects on partitioning functions using six parameters.

	atag	ashrt	bsh	blv	bst	bcb
atag	1					
ashrt	0	1				
bsh	0	0.68	1			
blv	0.44	0	0	1		
bst	0.47	0	0	0.68	1	
bcb	0.03	0	0	0	0.11	1

The transformations used in this thesis

Apart from the above considerations, transformations chosen were also determined by the available data, and by the problems occurring if transformed arguments are outside the range of the interpolation routine. These problems were solved by treating the argument of the function as a fraction and rescaling to the original scale.

Based on the above considerations the parameters additionally introduced are presented in Table A7.8. The table used to calculate the development rate was replaced by a linear function, given the discussion regarding overparametrization. The development rate DVR is calculated as follows:

DVR = [1 – L_8( )] * F_DVRV( ) + L_8( ) * F_DVRR( )

where L_8 switches the calculation from one function to another, and F_DVRV and F_DVRR are auxiliary functions calculated as:

F_DVRV = max(0, 1/TSDVRV*(DAVTMP–TBSVRV)**ADVRV)

F_DVRR = max(0, 1/TSDVRR*(DAVTMP–TBSVRR)**ADVRR)

where TBSVRV is the base temperature before silking (female flowering) and is equal to the base temperature after silking (TBSVRR). 1/TSDVRV is the increase in development stage per degree Celsius above the base temperature between emergence and flowering, while 1/TSDVRR is the increase in development stage per degree Celsius above the base temperature between silking and physiological maturity. ADVRV and ADVRR are parameters which would allow to incorporate non-linear responses to temperature. These parameters were set to 1, and assumed to be constant. The same approach was implemented in LINTUL, but was simplified to:

DVR = max(0, 1/TSDVR*(DAVTMP–TBASE)**ADVR)

Table A7.8 Additional parameters introduced in the model SUCROS87 to allow to modify the different tables. The asterisk indicates the parameters that could be estimated from literature data (cf. Chapter 4); other parameters were set to 1 or to their default value.

Tabulated function	Properties of scale transformation			
	Argument	Parameters introduced	Response	Parameters introduced
F_X(DVS)	rescaled fraction	ABDVS	partitioning	ASH, BSH* ALV, BLV* AST, BST ACB, BCB ASO ART
SLA(DVS)	rescaled fraction	ACDVS, BCDVS	positive	ASLA, BSLA*
AMDVS(DVS)	rescaled fraction	AADVS, BADVS	fraction	AAMDVS*, BAMDVS
DVR(DAVTMP) replaced by functions	real	TBSVRV, TSDVRR* TBSVRR, TSDVRR*	positive	ADVRV ADVRR
AMTMP(DDTMP)	rescaled fraction	ADDTMP, BDDTMP	fraction	AAMTMP*, BAMTMP

In case of the dependence of leaf photosynthesis on development stage, and temperature a single scaling factor was estimated which modifies the response value. The following transformation of the response was chosen:

$$AMDVS^{*} = \frac{AMDVS^{AAMDVS}}{AMDVS^{AAMDVS} + (1 - AMDVS)^{AAMDVS}}$$

The same type of transformation was chosen for AMTMP.

The statistical description of parameter uncertainty requires that the partitioning functions are estimated from a single instance. No instance was retrieved from literature, which allowed to fit all shoot partitioning functions to the data simultaneously. Restricting the argument values to a specific DVS-range allowed to select instances which could be used to fit leaf and

stem partitioning simultaneously. A single parameter was estimated, transforming the response value of the partitioning function to leaves. The transformation chosen is given as:

$$FLV^{*} = \frac{BLV \cdot FLV}{BLV \cdot FLV + FST}$$

The transformed stem partioning function is calculated as 1 – FLV over the period considered. In the shoot-root partitioning function the same transformation was used and the associated parameter (BSH) was estimated.

The function SLA, a tabulated function of DVS, was parameterized through a transformation of the SLA-values. SLA has a positive value; the response was transformed using a multiplicative transformation (SLA = BSLA*SLA). The calculation was identical to the calculation of the partitioning coefficients.

Whereas the same approach can be used for the partitioning tables in LINTUL, relevant partitioning coefficients (notably partitioning to grain) could not be determined from the data.

Conclusions

The above transformations allow to modify tabulated functions used in SUCROS87 and LINTUL without modifying the basic input. The modifications necessary are slight and depend on the scales on which the tabulated function is measured. When reparameterizing tabulated functions as proposed in this appendix the problem of overparameterization should be borne in mind. Two examples are discussed, which show that in the case of the development rate the conclusion regarding parameterization is clear-cut, whereas in the case of the partitioning functions conclusions are less clear-cut and other arguments should be considered, before adopting a parameterization. The arguments to be considered are number of parameters used and interpretability of the parameterization.

Appendix 8 Additional references used in the determination of parameter uncertainty

Development stage and emergence

Aitken, Y., 1980. The early maturing character in maize in relation to temperature and photoperiod. Zeitschrift für Acker- und Pflanzenbau -Journal of Agronomy and Crop Science 149: 89-106.

Alessi, J. and J.F. Power, 1971. Corn emergence in relation to soil temperature and seeding depth. Agronomy Journal 63: 717-719.

Allison, J.C.S. and T.B. Daynard, 1979. Effect of change in time of flowering, induced by altering photoperiod or temperature, on attributes related to yield in maize. Crop Science 19(1): 1-4.

Andrew, R.H., F.P. Ferwerda, and A.M. Strommen, 1956. Maturation and yield of corn as influenced by climate and production technique. Agronomy Journal 48: 231-236.

Beauchamp, E.G. and D.J. Lathwell, 1967. Root-zone temperature effects on the early development of maize. Plant and Soil 26(2): 224-234.

Benvenuti, A. and P. Belloni, 1990. Plant growth and dry matter yield in maize in relation to cultivar and density. Agricoltura Mediterranea. 120: 422-428.

Bloc, D., J.-P. Gouet, J. Lorgeou, H. Haudecoeur, and D. Taillet, 1984. Besoins en temperatures des variétes de maïs (inscrites au catalogue en 1982). Pau: Association generale des producteurs de maïs, 58 p.

Bonaparte, E.E.N.A., 1975. The effects of temperature, daylength, soil fertility and soil moisture on leaf number and duration to tassel emergence in Zea mays. Annals of Botany 39: 853-861.

Breuer, C.M., R.B. Hunter, and L.W. Kannenberg, 1976. Effects of 10- and 20- hour photoperiod treatments at 20- and 30°C on rate of development of a single-cross maize hybrid. Canadian Journal of Plant Science 56: 795-798.

Bunting, E.S., 1976. Accumulated temperature and maize development in England. Journal Agricultural Science (Cambridge) 87: 577-583.

Coelho, D.T. and R.F. Dale, 1980. An energy-crop growth variable and temperature function for predicting corn growth and development: planting to silking. Agronomy Journal 72: 503-510.

Coligado, M.C. and D.M. Brown, 1975. A bio-photo-thermal model to predict tassel-initiation time in corn. Agricultural Meteorology 15: 11-31.

Cross, H.Z. and M.S. Zuber, 1972. Prediction of flowering dates in maize based on different methods of estimating thermal time. Agronomy Journal 64: 351-355.

Darrah, L.L., 1976. Altitude and environmental responses of entries in the 1972-1973 Eastern African maize variety trial. East African Agricultural and Forestry Journal 41(4): 273-288.

Daynard, T.B., 1972. Relationships among black layer formation, grain moisture percentage, and heat unit accumulation in corn. Agronomy Journal 64: 716-719.

Derieux, M. and R. Bonhomme, 1982a. Heat unit requirements for maize hybrids in Europe. Results of the European FAO sub-network. I. sowing - silking period. Maydica 27: 59-77.

Derieux, M. and R. Bonhomme, 1982b. Heat unit requirements for maize hybrids in Europe. Results of the European FAO sub-network. II. Period from silking to maturity. Maydica 27: 79-96.

Derieux, M. and R. Bonhomme, 1990. Heat unit requirements of maize inbred lines for pollen shedding and silking: results of the European FAO network. Maydica 35: 41-46.

Dobben, W.H. van, 1962. Influence of temperature and light conditions on dry-matter distribution, development rate and yield in arable crops. Netherlands Journal of Agricultural Science 10(5): 377-389.

Dwyer, L.M., H.N. Hayhoe, and J.L.B. Culley, 1990. Prediction of soil temperature from air temperature for estimating corn emergence. Canadian Journal of Plant Science 70: 619-628

Gilmore jr., E.C. and J.S. Rogers, 1958. Heat units as a method of measuring maturity in corn. Agronomy Journal 50: 611-615.

Grant, R.F., 1989. Simulation of maize phenology. Agronomy Journal 81: 451-457.

Groot, J.J.R., M.J. Kropff, F.J.H. Vossen, C.J.T. Spitters, and R. Rabbinge, 1986. A decimal code for the developmental stages of maize and its relation to accumulated heat units. Netherlands Journal of Agricultural Science 34: 67-73.

Gubbels, G.H., 1974. Growth of corn seedlings under low temperatures as affected by genotype, seed size, total oil and fatty acid content of the seed. Canadian Journal of Plant Science 54: 425-426, note.

Hanke, P. and U. Koss, 1961. Beziehungen zwischen einigen Witterungsfaktoren und der phänologischen Entwicklung von Mais in Bad Lauchstadt. Albrecht Thaer-Archiv 5(9): 705-722.

Hanway, J.J. and W.A. Russell, 1969. Dry-matter accumulation in corn plants: comparisons among single cross hybrids. Agronomy Journal 61: 947-951.

Hanway, J.J., 1963. Growth stages of corn. Agronomy Journal 55: 487-492.

Hayhoe, H.N. and L.M. Dwyer, 1990. Relationship between percentage emergence and growing degree days for corn. Canadian Journal of Soil Science 70: 493-497.

Hilson, M.T. and L.H. Penny, 1965. Dry matter accumulation and moisture loss during maturation of corn grain. Agronomy Journal 57: 150-153.

Hodges, T. and D.W. Evans, 1992. Leaf emergence and leaf duration related to thermal time calculations in CERES-maize. Agronomy Journal 84: 724-730.

Karlen, D.L., E.J. Sadler, and C.R. Camp, 1987. Dry matter, phosphorus, and potassium accumulation rates by corn on Norfolk loamy sand. Agronomy Journal 79: 649-656.

Karlen, D.L., R.L. Flannery, and E.J. Sadler, 1988. Aerial accumulation and partitioning of nutrients by corn. Agronomy Journal 80: 232-242.

Leng, E.R., 1951. Time-relationships in tassel development of inbred and hybrid corn. Agronomy Journal 43: 445-449.

Lenga, F.K. and B.A. Keating, 1990. Use of thermal time to predict phenology of Kenyan maize germplasm. East African Agricultural and Forestry Journal 55(3): 103-123.

Linvill, D.E., R.F. Dale, and H.F. Hodges, 1978. Solar radiation weighting for weather and corn growth models. Agronomy Journal 70: 257-263.

Magoon, C.A., 1932. Response of sweet corn to varying temperatures from time of planting

to canning maturity. (Technical bulletin 312). Washington D.C.: USDA.
Mederski, H.J., M.E. Miller, and C.R. Weaver, 1973. Accumulated heat units for classifying corn hybrid maturity. Agronomy Journal 65: 743-747.
Miller, D.E., 1968. Emergence and development of sweet corn as influenced by various soil mulches. Agronomy Journal 60: 369-371.
Neild, R.E., 1982. Temperature and rainfall influences on the phenology and yield of grain sorghum and maize: a comparison. Agricultural Meteorology 27: 79-88.
Phipps, R.H., R.J. Fulford, and F.C. Crofts, 1975. Relationships between the production of forage maize and accumulated temperature, Ontario heat units and solar radiation. Agricultural Meteorology 14: 385-397.
Russell, W.K., C.W. Stuber, 1983. Effects of photoperiod and temperatures on the duration of vegetative growth in maize. Crop Science 23: 847-850.
Schans, D.A. van der, H.M.G. van der Werf, and W. van den Berg, 1993. Produktie- en kwaliteitsverloop bij snijmaïs - Verslag van onderzoek naar het optimale oogstijdstip bij snijmaïs (Production and quality of silage-maize during ripening - Report of research into the optimum time of harvest) (PAGV-verslag 155). Lelystad: PAGV, 75 p.
Schneider, E.C. and S.C. Gupta, 1985. Corn emergence as influenced by soil temperature, matric potential and aggregate size distribution. Soil Science Society of America Journal 49: 415-422.
Schröder, J., 1990. Optimale plantgetal van snijmaïs en van korrelmaïs. (PAGV-verslag 108) Lelystad, PAGV, 84 p.
Shaw, R.H. and H.C.S. Thom, 1951a. On the phenology of field corn, the vegetative period. Agronomy Journal 43: 9-15.
Shaw, R.H. and H.C.S. Thom, 1951b. On the phenology of field corn, silking to maturity. Agronomy Journal 43: 541-546.
Shaw, R.H., 1988. Chapter 10: Climatic requirement. Pages 591-623 in: Sprague, G.F. and J.W. Dudley (Eds). Corn improvement. Agronomy monograph No. 18. Madison, WI. USA: ASA-CSSA-SSSA.
Sibma, L., 1987. Ontwikkeling en groei van maïs onder nederlandse omstandigheden Gewassenreeks 1. Wageningen: Pudoc, 57p.
Singh, P.M., J.R. Gilley, and W.E. Splinter, 1976. Temperature thresholds for corn growth in a controlled environment. Transactions of the ASAE: 1152-1155.
Stamp, P., 1980. Chlorophyll content, PEP-carboxylase activity and plant growth of the genotypes of a maize double cross hybrid. Zeitschrift für Pflanzenzüchtung (Journal of plant Breeding) 85: 128-139.
Struik, P.C., M. Doorgeest, and J.G. Boonman, 1986. Environmental effects on flowering characteristics and kernel set of maize. Netherlands Journal of Agricultural Science 34: 469-484.
Tekrony, D.M., D.B. Egli, and D.A. Wickham, 1989. Corn seed vigor effect on no-tillage field performance - I. Field emergence. Crop Science 29: 1523-1528.
Warrington, I.J. and E.T. Kanemasu, 1983. Corn growth response to temperature and photoperiod I: seedling emergence, tassel initiation and anthesis. Agronomy Journal 75: 749-754.

Werf, H.M.G. van der, and H. Hoek, 1989. Teeltvervroeging bij maïs: Het effect van bodembedekkend plastic folie en uitplanten op de opbrengst van maïs (The effect of plastic mulch and planting on yield of maize). (PAGV-verslag 79). Lelystad: PAGV, 19 p.

Willis, W.O., W.E. Larson, and D. Kirkham, 1957. Corn growth as affected by soil temperature and mulch. Agronomy Journal 49: 323-327.

Leaf scattering coefficient (SCV)

Bunnik, N.J., 1978. The multispectral reflectance of shortwave radiation by agricultural crops in relation with their morphological properties. Mededelingen Landbouwhogeschool Wageningen 78-1, 176 p.

Goudriaan, J., 1977. Crop micrometeorology: a simulation study. Wageningen: PUDOC, 249 p.

Maas, S.J. and J.R. Dunlap, 1989. Reflectance, transmittance and absorptance of light by normal, etiolated and albino corn leaves. Agronomy Journal 81: 105-110.

Sellers, P.J., 1989. Vegetation-canopy spectral reflectance and biophysical processes. Chapter 8 in: Asrar, G. (Ed.). Theory and applications of optical remote sensing. New York: Wiley, 734 p.

Woolley, J.T., 1971. Reflectance and transmittance of light by leaves. Plant physiology 47: 656-662.

Yocum, C.S., L.H. Allen, and E.R. Lemon, 1964. Photosynthesis under field conditions VI. Solar radiation balance and photosynthetic efficiency. Agronomy Journal 56(3): 249-253.

Extinction coefficient (KDIF)

Allen, L.H., C.S. Yocum, and E.R. Lemon, 1964. Photosynthesis under field conditions. VII. Radiant energy exchanges within a corn crop canopy and implications in water use efficiency. Agronomy Journal 57(6): 575-580.

Birch, C.J., 1996. Modelling the ontogeny, leaf area and dry matter accumulation and distribution in diverse cultivars of maize. PhD-thesis, University of Queensland, 258 p.

Flénet, F., J.R. Kiniry, J.E. Board, M.E. Westgate, and D.C. Reicosky, 1996. Row spacing effects on light extinction coefficients of corn, sorghum, soybean, and sunflower. Agronomy Journal 88:185-190.

Gallo, K.P. and C.S.T. Daughtry, 1986. Techniques for measuring intercepted and absorbed photosynthetically active radiation in corn canopies. Agronomy Journal 78: 752-756.

Hatfield, J.L. and R.E. Carlson, 1979. Light quality distributions and spectral albedo of three maize canopies. Agricultural Meteorology 20: 215-226.

Loomis, R.S., W.A. Williams, W.G. Duncan, A. Dovrat, and F. Nunez-A., 1968. Quantitative description of foliage display and light absorption in field communities of corn plants. Crop Science 8: 352-356.

Monteith, J.L. and M.H. Unsworth, 1990. Principles of environmental physics. London: Arnold, 291 p.

Rabbinge, R., S.A. Ward, and H.H. van Laar (Eds), 1989. Simulation and systems

management in crop protection. Simulation Monographs 32. Wageningen: PUDOC, 420 p.
Richardson, A.J. and C.L. Wiegand, 1989. Canopy leaf display effects on absorbed, transmitted and reflected solar radiation. Remote Sensing of Environment 29: 15-24.
Sivakumar, M.V.K. and S.M. Virmani, 1984. Crop productivity in relation to interception of photosynthetically active radiation. Agricultural and Forest Meteorology 31: 131-141.
Tollenaar, M. and A. Aguilera, 1992. Radiation use efficiency of an old and a new maize hybrid. Agronomy Journal 84: 536-541.
Varlet-Grancher, C., G. Gosse, M. Chartier, H. Sinoquet, R. Bonhomme, and J.-M. Allirand, 1989. Mise au point: rayonnement solaire absorbé ou intercepté par un couvert végétal. Agronomie 9: 419-439.

Fraction photosynthetically active radiation (FRPAR)

Britton, C.M. and J.D. Dodd, 1976. Relationships of photosynthetically active radiation and shortwave irradiance. Agricultural Meteorology 17: 1-7.
Howell, T.A., D.W. Meek, and J.L. Hatfield, 1983. Relationship of photosynthetically active radiation to shortwave radiation in the San Joaquin valley. Agricultural Meteorology 28: 157-175.
Lauciani, E. and A. Ponticiello, 1993. Photosynthetically active radiation (PAR): empirical determination and literature survey. FAO Agrometeorology series, working paper Nr. 7, 85 p.
Lee, YS and Jung, YS, 1988. Utilization of photosynthetically active radiation in maize. Research reports of the rural development administration soils and fertilizer, Korean republic 1988: 30(1): 31-35; based on abstract in CAB.
Monteith, J.L. and M.H. Unsworth, 1990. Principles of environmental physics. London: Arnold, 291 p.
Sceicz, G., 1974. Solar radiation for plant growth. Journal of Applied Ecology 11: 617-636.
Stigter, C.J. and V.M.M. Musabilha, 1982. The conservative ratio of photosynthetically active tot total radiation in the tropics. Journal of Applied Ecology 19: 853-858.
Varlet-Grancher, C., G. Gosse, M. Chartier, H. Sinoquet, R. Bonhomme, and J.-M. Allirand, 1989. Mise au point: rayonnement solaire absorbé ou intercepté par un couvert végétal. Agronomie 9: 419-439.

Parameters determining the ratio diffuse over total global radiation

Chendo, M.A.C. and A.A.L. Maduekwe, 1994. Hourly global and diffuse radiation of Lagos, Nigeria - correlation with some atmospheric parameters. Solar Energy 52(3): 247-251.
Jong, J.B.R.M. de, 1980. (in Dutch). Een karakterisering van de zonnestraling in Nederland. (Deel I, Deel II: bijlagen). Scriptie T.U. Eindhoven, 97+67 p.
Reindl, D.T., W.A. Beckman, and J.A. Duffie, 1990. Diffuse fraction correlations. Solar Energy 45(1): 1-7.
Skartveit, A. and J.A. Olseth, 1987. A model for the diffuse fraction of hourly global radiation. Solar Energy 38(4): 271-274.

Dependence of atmospheric transmission on solar elevation (DLYATR)

Carroll, J.J., 1985. Global transmissivity and diffuse fraction of solar radiation for clear and cloudy skies as measured and as predicted by bulk transmission models. Solar Energy 35(2): 105-118.

Jong, J.B.R.M. de, 1980. (in Dutch). Een karakterisering van de zonnestraling in Nederland. (Deel I, Deel II: bijlagen). Scriptie T.U. Eindhoven, 97+67 p.

Lumb, F.E., 1964. The influence of cloud on hourly amounts of total solar radiation at the sea surface. Quarterly Journal of the Royal Meteorological Society 90: 43-56, C/R: 493-495.

Monteith, J.L. and M.H. Unsworth, 1990. Principles of environmental physics. London: Arnold, 291 p.

Olseth, J.A., A. Skartveith, 1993. Characteristics of hourly global irradiance modelled from cloud data. Solar Energy 51(3): 197-204.

Leaf photosynthesis parameters AMX and EFF

Bunce, J.A., 1990. Afternoon inhibition of photosynthesis in maize. I. Evidence, and relationship to stand density. Field Crops Research 24: 251-260.

Long, S.P., T.M. East, and N.R. Baker, 1983. Chilling damage to photosynthesis in young Zea mays. I. Effects of light and temperature variation on photosynthetic CO_2 assimilation. Journal of Experimental Botany 34(139): 177-188.

Tollenaar, M., 1989. Response of dry matter accumulation in maize to temperature: II. Leaf photosynthesis. Crop Science 29: 1275-1279.

Thiagarajah, M.R., L.A. Hunt, and J.D. Mahon, 1981. Effects of position and age on leaf photosynthesis in corn. Canadian Journal of Botany 59: 28-33.

Venanzi, G., S. Pasqualin, N. Pocceschi, M. Antonielli, and C.F. Soldatini, 1988. Leaf gas exchange and carboxylating enzymes in three hybrids of Zea mays of different earliness. Agricoltura Mediterranea 118: 366-370.

Ward, D.A., 1987. The temperature acclimation of photosynthetic responses to CO_2 in Zea mays and its relationship to the activities of photosynthetic enzymes and the CO_2 concentrating mechanism of C_4 photosynthesis. Plant, Cell and Environment 10: 407-411.

Dependence of maximum leaf photosynthesis on temperature (AMTMP)

Bird, I.F., M.J. Cornelius, and A.J. Keys, 1977. Effects of temperature on photosynthesis by maize and wheat. Journal Experimental Botany 28(104): 519-524.

Duncan, W.G. and J.D. Hesketh, 1968. Net photosynthetic rates, relative leaf growth rates, and leaf numbers of 22 races of maize grown at eight temperatures. Crop Science 8: 670-674.

Lee, S.S. and G.O. Estes, 1982. Corn physiology in short season and low temperatures environments. Agronomy Journal 74: 327-331.

Tollenaar, M., 1989. Response of dry matter accumulation in maize to temperature: II Leaf photosynthesis. Crop Science 29: 1275-1279.

Venanzi, G., S. Pasqualin, N. Pocceschi, M. Antonielli, and C.F. Soldatini, 1988. Leaf gas exchange and carboxylating enzymes in three hybrids of *Zea mays* of different earliness. Agricoltura Mediterranea 118: 366-370.

Winter, S.R. and J.W. Pendleton, 1970. Results of changing light and temperature regimes in a corn field and temperature effects on the apparent photosynthesis of individual leaves. Agronomy Journal 62(2): 181-184.

Dependence of maximum leaf photosynthesis on development stage (AMDVS)

Bethenod, O., C. Jacob, J.-C. Rode, and J.-F. Morot-Gaudry, 1982. Influence de l'âge sur les characteristiques photosynthetiques de la feuille de maïs. Agronomie 2(2): 159-166.

Thiagarajah, M.R., L.A. Hunt, and J.D. Mahon, 1981. Effects of position and age on leaf photosynthesis in corn. Canadian Journal of Botany 59: 28-33.

The temperature dependence of maintenance respiration (Q10)

Kaše, M. and J. Čatský, 1984. Maintenance and growth components of dark respiration rate in leaves of C_3 and C_4 plants as affected by leaf temperature. Biologia plantarum (Prague) 26(6): 461-470.

Penning de Vries, F.W.T., 1972. Respiration and growth. Pages 327-346 in: Rees A.R., K.E. Coeckshull, and D.W. Hand (Eds). Crop processes in controlled environments: proceedings of an international symposium held at Littlehampton, Sussex, July 1971. London: Academic Press.

Penning de Vries, F.W.T., 1974. The cost of maintenance processes in plant cells. Annals of Botany 39: 77-92.

Assimilate requirement: the parameter ASRSO

Penning de Vries, F.W.T., H.H. van Laar, and M.C.M. Chardon, 1983. Bioenergetics of growth of seeds, fruits and storage organs. Pages 37-59 in: Smith, W.H. and S.J. Banta (Eds), Potential productivity of field crops under different environments. Los Baños, Philippines: IRRI.

Assimilate requirement: the parameter ASRRT

Frossard, J.-S., 1986. Essai de modélisation de l'utilisation de l'énergie dans les racines. Application a l'effet de la témperature racinaire chez des plantules de deux hybrides de maïs. Agronomie 6(4): 401-408.

Machado, E.C. and A. Pereira, 1990. Respiração de crescimento e de manutenção da planta inteira das raízes e da parte aérea em milho e arroz. Pesquisa Agropecuaria Brasileira 25(6): 925-933.

Maintenance respiration: the parameter MAINLV

Penning de Vries, F.W.T., 1972. Respiration and growth. Pages 327-346 in: Rees A.R., K.E. Coeckshull, and D.W. Hand (Eds). Crop processes in controlled environments: proceedings of an international symposium held at Littlehampton, Sussex, July 1971. London: Academic Press.

Penning de Vries, F.W.T., 1974. The cost of maintenance processes in plant cells. Annals of Botany 39: 77-92.

Maintenance respiration: the parameter MAINRT

Frossard, J.-S.,1986. Essai de modélisation de l'utilisation de l'énergie dans les racines. Application a l'effet de la témperature racinaire chez des plantules de deux hybrides de maïs. Agronomie, 6(4): 401-408.

Penning de Vries, F.W.T., 1972. Respiration and growth. Pages 327-346 in: Rees A.R., K.E. Coeckshull, and D.W. Hand (Eds). Crop processes in controlled environments: proceedings of an international symposium held at Littlehampton, Sussex, July 1971. London: Academic Press.

Partitioning to the leaves (BLV)

Brimmer, C. and J. König, 1877. Trockengewichts-Bestimmungen beim Mais in siebentägigen Vegetations-Perioden. Landwirthschaftliche Jahrbücher 6: 839-848.

Cal, J.P. and R.L. Obendorf, 1972. Differential growth of corn (Zea mays L.). Hybrids seeded at cold root zone temperatures. Crop Science 12(5): 572-575.

Cooper P.J.M. and R. Law, 1977. Soil temperature and its association with maize yield variations in the highlands of Kenya. Journal Agricultural Science Cambridge 89: 355-363.

Demotes-Mainard, S. and S. Pellerin, 1992. Effect of mutual shading on the emergence of nodal roots and the root/shoot ratio of maize. Plant and Soil 147: 87-93.

Duley, F.L. and M.F. Miller, 1921. The effect of a varying supply of nutrients upon character and composition of the maize plant at different periods of growth. University of Missouri, College of Agriculture, Agricultural Experiment Station, Research Bulletin 42.

Gardner, F.P., R. Valle, and D.E. McCloud, 1990. Yield characteristics of ancient races of maize compared to a modern hybrid. Agronomy Journal 82: 864-868.

Gisiger, L., 1965. Ertragsbildung und Verlauf der Nährstoffaufnahme von Körnermaïs. Die Grüne (Schweizerische landwirtschaftliche Zeitschrift), April 1965, 17: 593-603.

Haggar, R.J., and D.C. Couper, 1972. Effects of plant population and fertilizer nitrogen on growth and components of yield of maize grown for silage in Nigeria. Experimental. Agriculture 8: 251-263.

Hanway, J.J. and W.A. Russell, 1969. Dry-matter accumulation in corn plants: comparisons among single cross hybrids. Agronomy Journal 61: 947-951.

Hanway, J.J., 1962. Corn growth and composition in relation to soil fertility: I. growth of

different plant parts and relation between leaf weight and yield. Agronomy Journal 54(2): 145-148.

Hunter, R.B., M. Tollenaar, and C.M. Breuer, 1977. Effects of photoperiod and temperature on vegetative and reproductive growth of a maize hybrid. Canadian Journal of Plant Science 57: 1127-1133.

Karlen, D.L., R.L. Flannery, and E.J. Sadler, 1988. Aerial accumulation and partitioning of nutrients by corn. Agronomy Journal 80: 232-242.

Kiesselbach, T.A., 1963. The use of water by crop plants. Agricultural Experiment Station special report 11. Lincoln: University of Nebraska, 126 p.

Kreusler, U., A. Prehn, and G. Becker, 1877. Beobachtungen über das Wachsthum der Maispflanze. Landwirthschaftliche Jahrbücher 6: 759-785.

Kreusler, U., A. Prehn, and G. Becker, 1877. Beobachtungen über das Wachsthum der Maispflanze. Landwirthschaftliche Jahrbücher 6: 787-800.

Kreusler, U., A. Prehn, and R. Hornberger, 1878. Beobachtungen über das Wachsthum der Maispflanze. Landwirthschaftliche Jahrbücher 7: 536-563.

Kreusler, U., A. Prehn, and R. Hornberger, 1879. Beobachtungen über das Wachsthum der Maispflanze. Landwirthschaftliche Jahrbücher 8: 617-622.

Miller, M.F. and F.L. Duley, 1925. The effect of a varying moisture supply upon the development and composition of the maize plant at different periods of growth. University of Missouri, College of Agriculture, Agricultural experiment station, Research Bulletin 76.

Nösberger, J., 1971. Einfluß der Bestandesdichte auf die Ertragsbildung bei Mais. I Teil: Effekte der Bestandesdichte und der Lichtintensität auf die Ertragsbildung. Zeitschrift für Acker- und Pflanzenbau-Journal of Agronomy and Crop Science 133: 215-232.

Sayre, J.D., 1948. Mineral accumulation in corn. Plant Physiology 23(3): 267-281.

Thom, E.R. and B.R. Watkin, 1978. Effect of rate and time of fertiliser nitrogen application on total plant, shoot, and root yields of maize. New Zealand Journal of Experimental Agriculture 6: 29-38.

Tollenaar, M., 1989. Response of dry matter accumulation in maize to temperature: I. Dry matter partitioning. Crop Science 29: 1239-1246.

Wagner, H., 1932. Beiträge zum Wachstumsverlauf landwirtschaftlicher Nutzpflanzen Mitteilung III (Mais, Buchweizen, Senf, Tabak und Mohn). Zeitschrift für Pflanzenernährung, Düngung und Bodenkunde, Teil A. 26: 8-57.

Shoot-root partitioning (BSH)

Allmaras, R.R. and W.W. Nelson, 1971. Corn root configuration as influenced by some interrow variants of tillage and straw mulch management. Soil Science Society America Proceedings 35: 974-980.

Allmaras, R.R., W.W. Nelson, and W.B. Voorhees, 1975. Soybean and corn rooting in southwestern Minnesota: II. Root distributions and related water inflow. Soil Science Society of America Proceedings 39: 771-777.

Alexander, K.G. and M.H. Miller, 1991. The effect of soil aggregate size on early growth and shoot-root ratio of maize. Plant and Soil 138: 189-194.

Anderson, E.L., 1988. Tillage and N fertilization effects on maize root growth and root-shoot ratio. Plant and Soil 108: 245-251.

Beauchamp, E.G. and D.J. Lathwell, 1967a. Effect of changes in root zone temperature on the subsequent growth and development of young corn plants. Agronomy Journal 59(2): 189-193.

Beauchamp, E.G. and D.J. Lathwell, 1967b. Root-zone temperature effects on the early development of maize. Plant and Soil 26: 224-234.

Breuer, C.M., R.B. Hunter, and L.W. Kannenberg, 1976. Effects of 10- and 20- hour photoperiod treatments at 20 and 30°C on rate of development of a single-cross maize hybrid. Canadian Journal of Plant Science 56: 795-798.

Brimmer, C. and J. König, 1877. Trockengewichts-Bestimmungen beim Mais in siebentägigen Vegetations-Perioden. Landwirthschaftliche Jahrbücher 6: 839-848.

Eghball, B. and J.W. Maranville, 1993. Root development and nitrogen influx of corn genotypes grown under combined drought and nitrogen stresses. Agronomy Journal 85:147-152.

Fairey, N.A. and T.B. Daynard, 1978. Quantitative distribution of assimilation in component organs of maize during reproductive growth. Canadian Journal of Plant Science 58:709-717.

Foth, H.D., 1962. Root and top growth of corn. Agronomy Journal 54: 49-52.

Foth, H.D., K.L. Kinra, and J.N. Pratt, 1960. Corn root development. Michigan State College of Agriculture and Applied Science, Agricultural Experiment Station Quarterly Bulletin 43: 2-13.

Frossard, J.-S. and M. Crocombette, 1985. Effet de la température des racines sur leur respiration et sur la croissance des plantules de deux hybrides de maïs. Agronomie 5(8): 719-725.

Hunter, R.B., M. Tollenaar, and C.M. Breuer, 1977. Effects of photoperiod and temperature on vegetative and reproductive growth of a maize hybrid. Canadian Journal of Plant Science 57: 1127-1133.

Jungk, A. and S.A. Barber, 1974. Phosphate uptake rate of corn roots as related to the proportion of the roots exposed to phosphate. Agronomy Journal 66: 554-557.

Knoll, H.A., D.J. Lathwell, and N.C. Brady, 1964. The influence of root zone temperature on the growth and contents of phosphorus and anthocyanin of corn. Soil Science Society of America Proceedings: 400-403.

Kreusler, U., A. Prehn, and G. Becker, 1877. Beobachtungen über das Wachsthum der Maispflanze. Landwirthschaftliche Jahrbücher 6: 759-785.

Kreusler, U., A. Prehn, and G. Becker, 1877. Beobachtungen über das Wachsthum der Maispflanze. Landwirthschaftliche Jahrbücher 6: 787-800.

Kreusler, U., A. Prehn, and R. Hornberger, 1878. Beobachtungen über das Wachsthum der Maispflanze. Landwirthschaftliche Jahrbücher 7: 536-563.

Kreusler, U., A. Prehn, and R. Hornberger, 1879. Beobachtungen über das Wachsthum der Maispflanze. Landwirthschaftliche Jahrbücher 8: 617-622.

Lubet, E. and C. Juste, 1985. Cinétique de la production de matière sèche et de prélèvement d'éléments nutritifs par une culture irriguée de maïs à haute potentialité de rendement.

Agronomie 5(3): 239-250.
Mackay, A.D. and S.A. Barber, 1984. Soil temperature effects on root growth and phosphorus uptake by corn. Soil Science Society of America Journal 48: 818-823.
Maizlish, N.A., D.D. Fritton, and W.A. Kendall, 1980. Root morphology and early development of maize at varying levels of nitrogen. Agronomy Journal 72: 25-31.
Mengel, D.B. and S.A. Barber, 1974. Development and distribution of the corn root system under field conditions. Agronomy Journal 66: 341-344.
Miller, M.H., G.K. Walker, M. Tollenaar, and K.G. Alexander, 1989. Growth and yield of maize grown outdoors hydroponically and in soil. Canadian Journal of Soil Science 69: 295-302.
Neubauer, C., and Teichler, 1877. Bestimmung der Trockengewichts-Zunahme bei Mais in den verschiedenen Wachsthums-perioden. Landwirthschaftliche Jahrbücher 6: 827-837.
Nösberger, J., 1971. Einfluß der Bestandesdichte auf die Ertragsbildung bei Mais. I. Teil: Effekte der Bestandesdichte und der Lichtintensität auf die Ertragsbildung. Zeitschrift für Acker- und Pflanzenbau -Journal of Agronomy and Crop Science 133: 215-232.
Porter, O.A. and J.T. Moraghan, 1975. Differential response of two corn inbreds to varying root temperature. Agronomy Journal 67: 515-518.
Schenk, M.K. and S.A. Barber, 1980. Potassium and phosphorus uptake by corn genotypes grown in the field as influenced by root characteristics. Plant and Soil 54: 65-76.
Sibma, L., 1987. Ontwikkeling en groei van maïs onder nederlandse omstandigheden Gewassenreeks 1. Wageningen: Pudoc, 57p.
Spencer, J.T., 1940. A comparative study of the seasonal root development of some inbred lines and hybrids of maize. Journal Agricultural Research 61(7): 521-538.
Tollenaar, M., 1989. Response of dry matter accumulation in maize to temperature: I. Dry matter partitioning. Crop Science 29: 1239-1246.
Tollenaar, M. and W. Migus, 1984. Dry matter accumulation of maize grown hydroponically under controlled-environment and field conditions. Canadian Journal Plant Science 64: 475-483.
Turner, T.W., 1922. Studies of the mechanism of the physiological effects of certain mineral salts in altering the ratio of top growth to root growth in seed plants. American Journal Botany 9(8): 415-445.
Warncke, D.D. and S.A. Barber, 1974. Root development and nutrient uptake by corn grown in solution culture. Agronomy Journal 66: 514-516.

Specific leaf area

Brimmer, C. and J. König, 1877. Trockengewichts-Bestimmungen beim Mais in siebentägigen Vegetations-Perioden. Landwirthschaftliche Jahrbücher 6: 839-848.
Gardner, F.P., R. Valle, and D.E. McCloud, 1990. Yield characteristics of ancient races of maize compared to a modern hybrid. Agronomy Journal 82: 864-868.
Kiesselbach, T.A., 1963. The use of water by crop plants. Agricultural Experiment Station special report 11. Lincoln: University of Nebraska, 126 p.
Kreusler, U., A. Prehn, and G. Becker, 1877. Beobachtungen über das Wachsthum der

Maispflanze. Landwirthschaftliche Jahrbücher 6: 759-785.
Kreusler, U., A. Prehn, and G. Becker, 1877. Beobachtungen über das Wachsthum der Maispflanze. Landwirthschaftliche Jahrbücher 6: 787-800.
Kreusler, U., A. Prehn, and R. Hornberger, 1878. Beobachtungen über das Wachsthum der Maispflanze. Landwirthschaftliche Jahrbücher 7: 536-563.
Kreusler, U., A. Prehn, and R. Hornberger, 1879. Beobachtungen über das Wachsthum der Maispflanze. Landwirthschaftliche Jahrbücher 8: 617-622.
Tollenaar, M., 1989. Response of dry matter accumulation in maize to temperature: I. Dry matter partitioning. Crop Science 29: 1239-1246.
Mackay, A.D. and S.A. Barber, 1984. Soil temperature effects on root growth and phosphorus uptake by corn. Soil Science Society of America Journal 48: 818-823.
Miller, M.F. and F.L. Duley, 1925. The effect of a varying moisture supply upon the development and composition of the maize plant at different periods of growth. University of Missouri, College of Agriculture, Agricultural Experiment Station, Research Bulletin 76.
Nösberger, J., 1971. Einfluß der Bestandesdichte auf die Ertragsbildung bei Mais. I. Teil: Effekte der Bestandesdichte und der Lichtintensität auf die Ertragsbildung. Zeitschrift für Acker- und Pflanzenbau -Journal of Agronomy and Crop Science 133: 215-232.

Initial leaf area growth (RGRL and LAINI)

Brimmer, C. and J. König, 1877. Trockengewichts-Bestimmungen beim Mais in siebentägigen Vegetations-Perioden. Landwirthschaftliche Jahrbücher 6: 839-848.
Kiesselbach, T.A., 1963. The use of water by crop plants. Agricultural Experiment Station special report 11. Lincoln: University of Nebraska, 126 p.
Kreusler, U., A. Prehn, and G. Becker, 1877. Beobachtungen über das Wachsthum der Maispflanze. Landwirthschaftliche Jahrbücher 6: 759-785.
Kreusler, U., A. Prehn, and G. Becker, 1877. Beobachtungen über das Wachsthum der Maispflanze. Landwirthschaftliche Jahrbücher 6: 787-800.
Kreusler, U., A. Prehn, and R. Hornberger, 1878. Beobachtungen über das Wachsthum der Maispflanze. Landwirthschaftliche Jahrbücher 7: 536-563.
Kreusler, U., A. Prehn, and R. Hornberger, 1879. Beobachtungen über das Wachsthum der Maispflanze. Landwirthschaftliche Jahrbücher 8: 617-622.
Verheul, M.J., 1992. Seedling growth of maize genotypes under chilling conditions. PhD-thesis, Eidgenössische Technische Hochschule Zürich, 101 p.

Light interception function (FINT)

Behaeghe, T., J. Stryckers, J. Traets, D. Reheul, G. Eeckhout, E. van Bockstaele, and A. de Baets, 1982. Beknopt verslag van de resultaten 1980-1982. Mededeling nr. 1982/2 van de leerstoel voor plantenteelt, Rijksuniversiteit Gent, pag. 23.
Brimmer, C., and J. König, 1877. Trockengewichts-Bestimmungen beim Mais in siebentägigen Vegetations-Perioden. Landwirthschaftliche Jahrbücher 6:839-848.
Gardner, F.P., R. Valle, and D.E. McCloud, 1990. Yield characteristics of ancient races of

maize compared to a modern hybrid. Agronomy Journal 82: 864-868.
Kiesselbach, T.A., 1963. The use of water by crop plants. Agricultural experiment station special report 11. Lincoln: University of Nebraska, 126 p.
Kreusler, U., A. Prehn, and G. Becker, 1877. Beobachtungen über das Wachsthum der Maispflanze. Landwirthschaftliche Jahrbücher 1877 6:759-785.
Kreusler, U., A. Prehn, and G. Becker, 1877. Beobachtungen über das Wachsthum der Maispflanze. Landwirthschaftliche Jahrbücher 6: 787-800.
Kreusler, U., A. Prehn, and R. Hornberger, 1878. Beobachtungen über das Wachsthum der Maispflanze. Landwirthschaftliche Jahrbücher 7: 536-563.
Kreusler, U., A. Prehn, and R. Hornberger, 1879. Beobachtungen über das Wachsthum der Maispflanze. Landwirthschaftliche Jahrbücher 8: 617-622.
Neubauer, C., and Teichler, 1877. Bestimmung der Trockengewichts-Zunahme bei Mais in den verschiedenen Wachsthums-Perioden. Landwirthschaftliche Jahrbücher 6: 827-837.
Thom, E.R. and B.R. Watkin, 1978. Effect of rate and time of fertiliser nitrogen application on total plant, shoot, and root yields of maize. New Zealand Journal of Experimental Agriculture 6: 29-38.

Light use efficiency (LUE)

Andrade, F.H., S.A. Uhart, and A. Cirilo, 1993. Temperature affects radiation use efficiency in maize. Field Crops Research 32: 17-25.
Cirilo, A.G. and F.H. Andrade, 1994. Sowing data and maize productivity; I. Crop growth and dry matter partitioning. Crop Science 34: 1039-1043.
Cooper, P.J.M., 1979. The association between altitude, environmental variables, maize growth and yields in Kenya. Journal Agricultural Science Cambridge 93: 635-649.
Gallo, K.P., C.S.T. Daughtry, and C.L. Wiegand, 1993. Errors in measuring absorbed radiation and computing crop radiation use efficiency. Agronomy Journal 85: 1222-1228.
Gosse, G., C. Varlet-Grancher, R. Bonhomme, M. Chartier, J.-M. Allirand, and G. Lemaire, 1986. Production maximale de matière sèche et rayonnement solaire intercepté par un couvert végétal. Agronomie 6(1): 47-56.
Kiniry, J.R., 1994. Radiation use efficiency and grain yield of maize competing with Johnsongrass. Agronomy Journal 86: 554-557.
Kiniry, J.R., C.A. Jones, J.C. O'Toole, R. Blanchet, M. Cabelguenne, and D.A. Spanel, 1989. Radiation use efficiency in biomass accumulation prior to grain-filling for five grain-crop species. Field Crops Research 20: 51-64.
Major, D.I., B.W. Beasley, and R.I. Hamilton, 1991. Effect of maize maturity on radiation use efficiency. Agronomy Journal 83: 895-903.
Muchow, R.C., 1989. Comparative productivity of maize, sorghum and pearl millet in a semi-arid tropical environment. I. Yield potential. Field Crops Research 20: 191-205.
Muchow, R.C., 1994. Effect of nitrogen on yield determination in irrigated maize in tropical and subtropical environments. Field Crops Research 38: 1-13.
Muchow, R.C. and R. Davis, 1988. Effect of nitrogen supply on the comparative productivity of maize and sorghum in a semi-arid tropical environment II. Radiation interception and

biomass accumulation. Field Crops Research 18: 17-30.
Otegui, M.E., M.G. Nicolini, R.A. Ruiz, and P.A. Dodds, 1995. Sowing date effects on grain yield components for different maize genotypes. Agronomy Journal 87: 29-33.
Westgate, M.E., F. Forcella, D.C. Reicosky, and J. Somsen, 1997. Rapid canopy closure for maize production in the northern US corn belt; Radiation use efficiency and grain yield. Field Crops Research 49: 249-258.

Curriculum vitae

Klaas Metselaar werd geboren op 5 augustus 1962 te Noord-Scharwoude. In 1981 behaalde hij het Abitur aan het Neusprachliches Gymnasium Hölty in Celle (Dtsl.). Vervolgens studeerde hij Bodemkunde en Bemestingsleer aan de Landbouwuniversiteit in Wageningen. Na een stage bij een onderzoeksproject van het ORSTOM in Bolivia, en de afstudeervakken Meteorologie, Irrigatie en Bodemnatuurkunde rondde hij zijn studie in 1988 af.

Na ondersteunende activiteiten voor een farming systems analysis project in Indonesië (INRES) bij de vakgroep Waterbeheer vervulde hij zijn vervangende dienstplicht van 1989 tot 1991 bij het SC-DLO in Wageningen. Hij droeg daar bij aan simulatiestudies op het gebied van waterbeheer. Bij het AB-DLO in Wageningen werkte hij vervolgens binnen een project (SARP-II) gericht op kennisoverdracht met behulp van simulatie en simulatiemodellen naar Zuid-oost Azië.

Van 1992 tot 1996 werkte hij, vanuit de vakgroep Agrotechniek en Agrofysica gedetacheerd bij de Groep Landbouw-Wiskunde van DLO (het huidige CBW, onderdeel van het CPRO-DLO), aan een project getiteld "Methoden voor de systematische evaluatie van de nauwkeurigheid van voorspellende modellen". De resultaten van het onderzoek zijn beschreven in dit proefschrift.

Sinds 1 juli 1998 werkt hij bij het AB-DLO aan een onderzoeksproject gericht op de zetmeelaardappelteelt.