Modelling and monitoring forest evapotranspiration. Behaviour, concepts and parameters
Dekker, S.C.

Citation for published version (APA):

General rights
It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations
If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: http://uba.uva.nl/en/contact, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.
1. GENERAL INTRODUCTION

Mathematical models are univocal descriptions of our concepts. They represent our perception of the true world and they are essential tools in hydrological and ecological studies to assess ecosystem responses during changes of environmental conditions or to assess the behaviour of the system. Confidence in these models is gained by comparing model results with observations. To achieve this confidence, a variety of tests with different purposes and terminologies, but all dealing with the comparison of model results with observations, are nowadays accepted. As a result, modellers claim that a model test is performed while any reference to the criteria is mostly not given.

1.1 TERMINOLOGY

In the last decade, a debate in literature over model testing in ecology and earth sciences has started (e.g. Janssen and Heuberger, 1995; Konikow and Bredehoef, 1992; Oreskes et al., 1994; Rykiel, 1994; Rykiel, 1996). Rastetter (1996) pointed out that the essence of the debate is the problem of induction (Popper), which is the problem of extrapolating from the specific to the general. No tests can establish the general validity of the model. Main reasons that we cannot establish the truth are (i) that some parameters or variables can only be established on a specific scale and therefore are incompletely known, (ii) that model concepts are simplicities of the true world and are developed with different perceptions and different aims and (iii) that all variables and observations are measured in a specific context with their own assumptions and inferences.

A summary of the different purposes and terminologies of the model tests, used in the debate, is given here. The authors agree with the definition of calibration, as the process to estimate model parameters and constants to improve the agreement between model output and observations (Janssen and Heuberger, 1995; Konikow and Bredehoef, 1992; Oreskes et al., 1994; Rykiel, 1996). However, the purpose of calibration is not clear at all. A good match does not prove the validity of the model because the solution can be non-unique (Konikow and Bredehoef, 1992; Oreskes et al., 1994) and the model can compensate calibration errors due to a wrong parameterisation (Konikow and Bredehoef, 1992). Therefore, Janssen and Heuberger (1995) suggest a calibration process, in which the evaluation of the model is incorporated. They discerned three major aspects: (i) the assessment of the ability of the model to reproduce the system behaviour, (ii) the assessment of the suitability of the model for the intended use, and (iii) the assessment of
the robustness of the estimated model parameters for different parts of the data set. They pointed out that the uncertainty in the model parameters should be adequately accounted for in further model applications.

In contrast to calibration, many different definitions of verification and validation are proposed in the used literature and references therein. Based on definitions in dictionaries, verification means 'the act to prove to be true or accurate or to ascertain the accuracy of truth'. Validation, as defined in the dictionary, means 'the establishment of legitimacy, in terms of arguments and methods'. A first reason that many different definitions exist is that the verification and validation are synonyms in ordinary language and denote both the establishment of truth. Oreskes et al. (1994) use the definitions from the dictionary and point out that verification (truth) is only possible in closed systems in which all components of the system are established independently and are known to be correct. Because natural systems are never closed, verification is impossible. Because of the synonyms in ordinary language, the same discussion about the establishment of the truth was found for validation.

A second, maybe more important reason that causes the confusion about the terminologies of verification and validation is that different purposes can be found why model results are compared with measurements and why a model must be verified or validated. For instance, one intended goal of verification or validation is to gain confidence in the model's ability to make reliable predictions (Konikow and Bredehoeft, 1992). Another goal is to establish the truth of the model concept, in the perspective that models are hypotheses, which can only be falsified. Because models are developed with different purposes, different perceptions and in different contexts, a model concept can be 'true'in the context of one perception.

Due to the impossibility of establishing the truth, Konikow and Bredehoeft (1992) and Oreskes et al. (1994) pose that verification and validation are impossible. Rykiel (1996) pointed out that validation is a process that can be decomposed in several components. As a result, the terms verification and validation are misleading and should be abandoned in favour of more meaningful terms. A more technical definition of verification is a demonstration that the modelling formalism is correct. Konikow and Bredehoeft (1992) and Oreskes et al. (1994) use for this definition 'verification of numerical solutions'. Konikow and Bredehoeft (1992) pose terms as sensitivity testing, benchmarking or history matching. Oreskes et al. (1994) re-use the term confirmation, which was proposed by the logical positivists. A model can be confirmed by observations, if these observations can be
shown to be true. Rykiel (1996) uses the term credibility and qualification, in which credibility is a sufficient degree of belief in the model for its intended purpose. Therefore, credibility is a subjective qualitative judgement, and cannot be quantified in any absolute sense. Qualification assesses the domain over which a model may properly be used.

1.2 MODEL BEHAVIOUR, CONCEPTS AND PARAMETERS

From the above discussion, it is clear that models cannot be used to establish the truth. Nevertheless, many other purposes consist to use and develop models. From a scientific point of view, models can be used to improve the insight in the processes, to extrapolate in time and space or to determine variables, which cannot be directly measured. To achieve these goals, confidence must be gained in the model concepts and model parameters. In figure 1.1, an outline is given to find out how to gain this confidence. The start of this outline is always the comparison of the model behaviour with the system behaviour. With this comparison model concepts or values of model parameters can be evaluated. In this thesis different methodologies are developed and used to improve the understanding of the model concepts in terms of cause-effect relationships and to improve the interpretation of the model parameters in terms of system properties.

System behaviour – Model behaviour

A model concept or values of model parameters can only be evaluated by comparing model results with measurements. As a consequence, we must always link the system behaviour, e.g. the measurements, to the model behaviour. Model results are compared to measurements to confirm the model concept or the value of the model parameter. However, confirmation of a model by measurements can be very easy and is dependent on the range and kind of the measurements. The result of confirmation is often a statement as average, well or good. To make confirmation more valuable, Reckhow (1983) point out that ‘the modeller must apply (i) a variety of tests, e.g. using the same variations in conditions as the calibration was performed, (ii) a statistical criterion for goodness of fit and (iii) an error analysis in both the predictions and observations’. Nevertheless, confirmation is a subjective measure and a good model result, only enhances our confidence in the model concept or the model parameters.
Figure 1.1: Outline to find out how to gain confidence in models: Model concepts or values of model parameters are always evaluated by comparing model results to measurements, e.g. comparing the model behaviour to the system behaviour. The understanding of the model concept can be improved by a focus on cause-effect relationships and the interpretation of the model parameters can be improved in terms of system properties.

As a result of the subjective judging of the confirmation step, the same results can either enhance the confidence in the model concept or model parameters or can stimulate the development of new model concepts or new model parameterisations. To improve this stimulation, we must not focus on similarities but rather on discrepancies (e.g. falsification, or an analysis of residuals) between model results and measurements.

Model concepts – Cause-Effect relationships

Several model concepts, using different processes, can give equal results. The choice of the processes and variables, included in the model concept, are related to the modellers own perception and to the specific aim of the model. As shown in Figure 1.1, a model concept can be improved by incorporating cause-effect relationships. With a focus on
discrepancies, the residuals between model results and measurements can be compared with input variables to identify missing variables or processes. These missing variables and processes with identifiable physical basis can give information on cause-effect relationships. If two or more model concepts are available, the discrepancies between the model results can also improve our understanding of the processes.

Model Parameters – System properties

In general, models contain parameters, which need to be identified. In many cases the parameters cannot be measured independently and can only be calibrated by a comparison of model results and measurements. The aim of calibration is the fit. However, a good fit does not guarantee the uniqueness of the parameter values and does not contribute to the interpretation of the model parameter in terms of system properties. Only a unique parameter estimate with high accuracy can contribute to the understanding of the system and can be used for extrapolation in time and space. With transfer functions these parameter estimates can be linked to system properties.

The parameter identification methodologies presented in this thesis will focus on the uniqueness of the parameters. Classical parameter identification approaches aim to find an optimal model-to-data fit by minimising the total data set with one objective function, for instance the Sum of Squared Errors (SSE). A major problem of parameter identification is that systematic model errors can be compensated by calibration errors in which parameters become non-unique fit-parameters without any physical meaning. The remaining residuals, between model results and measurements, are caused by random and systematic measurement errors and model inaccuracies and may contain information to improve the parameter estimates. With residual analysis, patterns can be explored to trace systematic effects due to wrong model parameter estimates. If fit-parameters are identified by calibration, than parameter estimates can vary by using different objective functions (Janssen and Heuberger, 1995). It is also known that the identification of the parameters is dependent on the range and distribution of the data (e.g. (Gupta and Sorooshian, 1985; Gupta et al., 1998; Kuczera, 1982; Musters and Bouten, 2000; Sorooshian et al., 1983; Yapo et al., 1998)) and dependent on extreme values (Finsterle and Najita, 1998; Legates and McCabe, 1999). This means that parameter identification problems will not simply disappear with the availability of more measurements. It also means that relevant information must be extracted from the total data set to identify the parameters. Once these conditions are selected, parameter values and accuracies can be estimated. The
accuracy of the parameter value is dependent on both model and measurement errors.

Parameter identification will suffer less from the problems of non-uniqueness by using independent parameters. These parameters can either be derived from literature or by calibration only by using another type of measurements than used with the model evaluation.

1.3 EVAPOTRANSPIRATION

In this thesis, several methodologies are developed and used to improve the understanding of forest evapotranspiration model concepts and to improve the interpretation of the model parameters.

The energy and water exchange at the earth surface play an important role in climate and climate change research (Shuttleworth, 1995). So-called Soil Vegetation Atmosphere Transfer (SVAT) processes describe this exchange and are incorporated in atmospheric Global Circulation Models (GCM) and global change models. The grid-sizes of these global models are in the order of 100-300 km².

The major issues in SVAT research deal with (1) plot scale research on SVAT processes and (2) how to scale these SVAT processes to regional scales and to global change time scales. Scaling in space can be done by aggregation of parameters or by aggregation of model output (e.g. Rastetter et al., 1992), (Kabat et al., 1997), (Heuvelink and Pebesma, 1999). With scaling in time more feedback mechanisms must be taken in the plot scale model, such as growth and nutrient availability. A major problem in model evaluation is that the evaluation measurements are collected at smaller spatial and temporal scales than the model predictions.

This study deals with plot scale research of forest evapotranspiration processes. Evaporation of intercepted rain is an important hydrological process in forests. Water budget studies show that the evaporation of intercepted rain amounts 10 - 50 % of the total rainfall (e.g. Calder, 1998; Wijk et al., 2000). In general, the model concept of a water bucket of stored water in the canopy that can evaporate or drain is rather well understood. Evaporation of intercepted rainfall is normally considered to be a physical process by using the energy balance and aerodynamic transport equations of Penman (1948). In most studies, the water retention characteristics of the canopy are not known, while evaporation, canopy water storage and drainage are dependent to it. To estimate these processes, the model is calibrated to measurements. In most studies only throughfall is
measured, while the other processes are derived from the calibrated values.

In contrast to evaporation, there is no consensus about the concepts of the process of transpiration. From all different types of process oriented forest transpiration models, four different perspectives were found: the cooling of leaves, the assimilation of CO₂, the energy balance (combined with bulk stomatal conductance) and the water balance. Transpiration of forests can be measured at different spatial scales. At the leaf level, porometers and gas-exchange chambers are used to find plant-physiological mechanisms under changes of environmental conditions. At the tree level, two techniques are generally used: sapflow (Köstner et al., 1998) and soil water content measurements (Musters et al., 2000). At the stand level, eddy-correlation techniques are used and for larger areas remote sensing techniques can be useful for obtaining information of parameters for land-surface interactions (Running et al., 1989). Due to the different model concepts, problems related to transpiration are even broader than with the process of evaporation of intercepted water.

1.4 ORGANISATION OF THE THESIS

The chapters 2 to 6 are integral copies of manuscripts that are published, submitted or will be submitted in relevant scientific journals. In each chapter information on models, measurements and research site, relevant for that manuscript is given. Consequently, duplication sometimes occurs.

All half-hourly micro-meteorological measurements used in this thesis for both 1989 and 1995 were measured by the KNMI (Bosveld et al., 1998; Bosveld 1999). All soil water, throughfall and water storage measurements were measured by the UvA (Bouten et al., 1996; Tikta and Bouten, 1994)

In chapter 2, three forest transpiration model concepts are compared: leaf cooling, CO₂ assimilation and the combined energy and water balance. The purpose of the chapter is to find similarities and discrepancies for transpiration fluxes of half hourly periods and to find improvements of descriptions of forest transpiration processes.

Chapter 3 describes the gas-exchange of CO₂ and H₂O at the leaf and stand scale. Photosynthesis measurements with gas exchange chambers are used to calibrate the Farquhar/Ball leaf scale model. This calibrated CO₂ leaf model is scaled up to the canopy level by a three-dimensional light interception model in order to estimate CO₂ photosynthesis, transpiration and water use efficiency. Modelled canopy transpiration is
independently confirmed to sapflow measurements. So independent parameters are used to simulate stand fluxes. Finally the residual, between model results and measurements are used to identify variables and processes, which were not considered in the original model.

In chapter 4 and 5, the information content of measurements is used to identify unique parameters with high accuracy. In this thesis, the Parameter Identification Method based on Localisation of Information (PIMLI) was further developed and was partly based on the work of Musters and Bouten (2000) and Vrugt et al. (2000). Different objective functions with high information content are used by PIMLI to identify the various parameters. The selected conditions can be used to improve the physical meaning of the various parameters. In chapter 4, forest transpiration is modelled with the Single Big Leaf (SBL) model concept, based on the Penman-Monteith equation. The model contains many calibration parameters and mathematical forms of response functions. With calibration, the model parameters are optimised to fit the latent heat eddy correlation measurements. However, time series of environmental conditions determining forest transpiration contain periods with coupled conditions and redundant information while other conditions are hardly measured. In this chapter, measurements with high information content are selected by PIMLI. The accuracy and parameter estimates are calculated by using only these selected measurements. The aim of chapter 5 is to identify model parameters of a rainfall interception model by using throughfall and canopy storage measurements. Throughfall, canopy storage and evaporation processes are all dependent of each other. With PIMLI, conditions are selected with highest information yielding unique parameters with high accuracy. As soon the selection criteria are known to identify the parameters, true measurements were used.

In chapter 6, an analysis of the residuals between model results and measurements is performed with Artificial Neural Networks (ANNs). Random and systematic measurement errors and model inaccuracies cause these residuals. ANNs are used to explore patterns in the residuals to find model inaccuracies. Only systematic errors with an identifiable physical basis are used to further improve the existing SBL model. Model improvement may consist of incorporation of additional environmental variables, not considered in the original model or an improved model parameterisation.

Finally, in chapter 7, some remarks are given about modelling and monitoring and some suggestions are made for future forest evapotranspiration research to improve the understanding of the cause-effect relationships and to improve the interpretation of the parameters in terms of system properties.
REFERENCES


