Available online at www.sciencedirect.com

SCIENCE DIRECT®

Agricultural Systems 88 (2006) 451–471

AGRICULTURAL SYSTEMS

www.elsevier.com/locate/agsy

Evaluating uncertainty introduced to process-based simulation model estimates by alternative sources of meteorological data

M. Rivington [a,*], K.B. Matthews [a], G. Bellocchi [b], K. Buchan [a]

[a] *Macaulay Institute, Craigiebuckler, Aberdeen AB15 8QH, Scotland, United Kingdom*
[b] *Research Institute for Industrial Crops, Via Corticella, 133-40128 Bologna, Italy*

Received 31 March 2004; received in revised form 17 February 2005; accepted 13 July 2005

Abstract

Models that represent biophysical processes in hydrology, ecology and agricultural systems, when applied at specific locations, can make estimates with significant errors if meteorological input data are not representative of the sites. This is particularly important where the estimates from the models are used for decision support, strategic planning and policy development, due to the impacts of introduced uncertainty. This paper investigates the impacts of meteorological data sources on a cropping systems simulation model's estimate of crop yield, and quantifies the uncertainty that arises when site-specific weather data are not available. In the UK, as elsewhere, many meteorological stations record precipitation and air temperature, but very few also record solar radiation, a key driving input data set. The impacts of using on-site observed precipitation and temperature with estimated solar radiation, and off-site entirely observed meteorological data was tested on the model's yield estimates. This gave two scenarios: on-site observed versus partially modelled data; and on-site observed versus substitute data from neighbouring sites, for 24 meteorological stations in the UK.

The analysis indicates that neighbouring meteorological stations can often be an inappropriate source of data. Of the 24 stations used, only 32% of the nearest neighbours were able to provide the best substitute off-site data. On-site modelled data provided better results than

* Corresponding author. Tel.: +44 0 1224 498200; fax: +44 0 1224 311556.
E-mail address: m.rivington@macaulay.ac.uk (M. Rivington).

doi:10.1016/j.agsy.2005.07.004

observed off-site data. The results demonstrate that the range of alternative data sources tested are capable of having both acceptable and unacceptable impacts on model performance across a range of assessment metrics, i.e. on-site data sources each produced yield over- or under-estimate errors greater than 2 t ha^{-1}. A large amount of uncertainty can be introduced to the model estimates due to the data source. Therefore, the applications of models that represent biophysical process where meteorological data are required, need to include the quantification of input data errors and estimate of the uncertainty that imperfect data will introduce.

Keywords: Meteorological data; Solar radiation; Simulation model; Uncertainty; Error propagation; Decision support; CropSyst

1. Introduction

There often exists a significant difference between the sophistication of models developed to represent natural processes and our ability to provide the required biophysical input data at a particular place of model application (PoMA). It may be desirable to apply a detailed model for a site-specific case study where biophysical data does not exist, hence the model becomes redundant if the quality of the input data leads to unreliable estimates. The lack of location-specific input data for spatially and temporally variable entities means that a model's site-specific estimates have potentially large and unquantified uncertainties.

There is a serious limit on the application of agricultural, hydrological and ecosystem models if weather data are not directly available (Hoogenboom, 2000; Bechini et al., 2000). The weather is one of the primary driving variables in biophysical processes and in determining human intervention through management responses. This is particularly the case in farming systems. The influence of the weather on biological processes tends to be non-linear (i.e. Nonhebel, 1994a), and is dependent on the correlations between individual weather variables. Models that represent multiple entities with complex biophysical interactions between them therefore require meteorological data that maintains appropriate values and correlations between variables.

Inappropriate choices of data source can have significant impacts on model estimates (Rivington et al., 2002), introducing uncertainties which manifest themselves as incorrect estimations of magnitudes, absolute values, relative timing and synchronisation. Appropriate location-specific data are also essential for model calibration and parameterisation, as non-representative data will result in unsuitable parameters for the PoMA. The consequences of data source choices are particularly important when models are used as components for decision support systems (DSS), strategic planning and policy development (i.e. in climate change impact studies; Rivington et al., 2004). Errors may be propagated through the model, leading to incorrect conclusions, recommendations or policy formulation. Using model estimates for decision support requires that the quality of model estimates is assessed in advance, or that the DSS outcomes be made insensitive to the prediction uncertainty (Norton,