

## Introduction

Statistical analysis of simulation results is required to evaluate how well the model results correspond to experimental data. Often times the evaluation of a model's performance is limited to a comparison between a few simulated variables and observed phenomena. Although cropping system models produce several outputs, it is common to evaluate individual outputs separately by means of one or more indices or test statistics. Such simple comparisons may be misleading, because a comprehensive evaluation of the model's performance requires an integrated evaluation of a set of major outputs. An integrated approach for evaluation based on fuzzy theory was proposed by Bellocchi et al. An extension of such a methodology is proposed here, accounting for a number of outputs from maize cultivations in the Pisa Valley, Italy.

## Simulations

The **CropSyst** model was evaluated for its ability to simulate above ground maize biomass and two soil variables at 0-0.60 m depth (i.e., water and  $\text{NO}_3\text{-N}$ ) in response to different types of tillage (conventional, no tillage) and nitrogen fertilisation (0, 300 kg N  $\text{ha}^{-1}$ ) on a silty soil at Pisa, Central Italy (lat.: 43.67 North, long.: 10.17 East, elev.: 6 m a.s.l.). Experimental data sampled during the growing seasons 1994 and 1995 were used for this purpose.



## Data processing

Model outputs such as soil water content (SWC,  $\text{m}^3 \text{m}^{-3}$ ), soil  $\text{NO}_3\text{-N}$  content (SNC,  $\text{kg ha}^{-1}$ ), and above ground biomass (AGB,  $\text{kg ha}^{-1}$ ) were compared against observed values. The agreement between simulations and observations was evaluated through the relative percent root mean squared error (**RRMSE**). An integrated index (**MVPI**: multiple variable performance index) was then created aggregating the **RRMSE** values computed for each output into one index, according to a fuzzy-based procedure. Different expert weights ( $B_i$ , i referred to as the i-th output variable) were given to different outputs.

	<b>MPVI (0, best - 1, worst)</b>			
	Basic indices	Expert weights	Boundaries Fav.	Unfav.
<b>RRMSE<sub>SWC</sub></b>	1.0	20	40	
<b>RRMSE<sub>SNC</sub></b>	0.5	20	40	
<b>RRMSE<sub>AGB</sub></b>	2.0	20	40	

## Results

The results of the comparisons between observed and simulated outputs for different management options are summarized in the table below. In general, the  $\text{NO}_3\text{-N}$  dynamics appeared more difficult to simulate when nitrogen fertilization was applied. However, the satisfactory estimate of other variables tended to compensate the unsatisfactory model response with  $\text{NO}_3\text{-N}$  under no tillage.

output	<b>RRMSE</b>		<b>MVPI</b>	
	no nitrogen fertilisation	nitrogen fertilisation	no nitrogen fertilisation	nitrogen fertilisation
<b>conventional tillage</b>				
Soil Water Content	10.50%	10.00%		
Soil $\text{NO}_3\text{-N}$ Content	26.60%	31.80%	0.2165	0.2996
Above Ground Biomass	26.60%	27.70%		
<b>no tillage</b>				
Soil Water Content	12.60%	8.60%		
Soil $\text{NO}_3\text{-N}$ Content	27.80%	44.10%	0.2831	0.2207
Above Ground Biomass	28.20%	25.20%		

## Conclusions

1. This work shows examples of an integrated approach to assess model performance. The methodology was applied to the simulation results of some common variables under contrasting experimental conditions. In particular, the expert system proposed here takes into account three output variables of major interest in cropping systems modelling.
2. We had to decide which attributes have more weight in determining the outcome of the reasoning. Larger importance was given to a model's ability to simulate crop biomass (as it is associated with the yield prediction), whereas possible failure in the simulation of soil variables was considered less stringent. Lower relevance was given to nitrogen balance than water balance, the latter interacting with chemical budget (either nitrogen, pesticides, or salinity). These considerations reflect our subjective judgment, and would have been different if the primary goal of our simulation was, for instance, nitrogen dynamics.
3. The results obtained are the consequence of the choices we made regarding the selection of limits and weights, which may significantly change the value of the outcome. The sensitivity of the outcome to weights can be recognized, for instance, with nitrogen fertilization under no tillage, producing better **MVPI** (0.2207) than the condition with no fertilization (0.2831), owing a better response (**RRMSE**=25.2%) associated with the most influential variable (i.e. crop biomass). A wider examination is therefore required for a general consensus about either weights applied to each output or the limits applied to **RRMSE**.
4. The fuzzy-based methodology proposed here represents a pragmatic approach towards a satisfactory solution for a comprehensive evaluation of model performance. Analysis of this type can give some protection against partial conclusions and indicate whether the model as a whole is conceptually consistent and related to experimental evidence. The method is flexible and can be proficiently applied to aggregating more output variables.
5. Provisions for the computation of integrated indices are implemented in the dynamic link library **IRENE\_DLL**, freely downloadable from the site <http://www.isci.it/tools>. The installation package includes an example of how use the DLL for aggregating variables within a Microsoft Excel spreadsheet.

## Project SIPEAA

Software Tools for Eco-compatible Farm Planning

<http://www.sipeaa.it>