International Environmental Modelling and Software Society (iEMSs)
2012 International Congress on Environmental Modelling and Software
Managing Resources of a Limited Planet, Sixth Biennial Meeting, Leipzig, Germany
R. Seppelt, A.A. Voinov, S. Lange, D. Bankamp (Eds.)
http://www.iemss.org/society/index.php/iemss-2012-proceedings

Comparing Modelling Solutions At Submodel Level: A Case On Soil Temperature Simulation.

Simone Bregaglio[1], Marcello Donatelli[2,3], Roberto Confalonieri[1], Roberto De Mascellis[4], Marco Acutis[1]

[1] *University of Milan, Department of Plant Production, via Celoria 2, Milan, Italy*
[2] *Agriculture Research Council, Research Centre for Industrial Crops, Bologna, Italy*
[3] *European Commission, Joint Research Centre, Institute for Environmental Sustainability MARS Unit, AGRI4CAST Action, Ispra, Italy*
[4] *National Research Council, Institute for Agricultural and Forest Systems in the Mediterranean, Ercolano, Italy*
simone.bregaglio@unimi.it

Abstract: What is commonly referred as either crop or cropping system simulation model is a set of interlinked mathematical representations of approaches which are abstractions of a single biological or physical process. The methodology for their evaluation has evolved in time, but it has always targeted models as unique and immutable, except for versioning, discrete units. This paper has explored aspects both of model composition in the perspective of evaluating alternate modelling approaches, and of modelling solutions evaluation. Soil temperature was chosen as case study, evaluating nine modelling solutions against a multi-year, multi-location database of field recorded time series of data. Multi-metric indices were also developed to quantify different aspects of model performance and to get a better insight on the impact of sub-model replacement. Results showed that the hybrid solution implementing the cascading model (soil water redistribution), Parton's approach (surface temperature), and SWAT (temperature along the soil profile) led to the best compromise between agreement and robustness under the explored conditions. The model libraries used to run the analysis, in form of extensible model components, are freely available for download, and they allow for further extension of the composition exercise.

Keywords: Model composition, model comparison, model development, composite indicators

1 INTRODUCTION

The term *model* has overloaded meanings: from a physical duplicate of a part of the real word to its abstraction, the latter often represented via mathematical equations which are meant to capture the traits of its behavior with respect to a specific objective of analysis. The term model is also overloaded with respect to its specific structure: models range from very complex formalizations to a single equation, very often a model being a composition of many sub-models. Biophysical models in agriculture are no exception: what is commonly referred as either crop or cropping system simulation model is a set of interlinked mathematical representations of approaches which are abstractions of a single biological or physical process. They are called models, instead of the possibly more appropriate term modelling solutions (MS), mostly because of the way they appear to the final user, who might even use them as black boxes driven by a graphical user interface. The methodology for the evaluation of such MS has evolved in time, considering different metrics [e.g., Bellocchi et al. 2002; Confalonieri et al. 2010a; Bregaglio et al. 2011]. In all cases evaluation has targeted MS as immutable, except for versioning, discrete units. However, the core of widely known MS relevant to the plant-soil system, such as APSIM [Keating et al. 2002], CropSyst [Stockle et al. 2003], DSSAT [Jones et al. 2003], STICS [Brisson et al. 2003] share many modelling approaches. Even if a “hybridization” across MS has occurred at least when a new one was firstly built, importing and testing alternate options for the simulation of a single process has never been a standard working methodology in model development. Evaluating MS is perfectly adequate for operational use of available discrete simulation packages, but it offers a confused picture if used to support model development, because it often does not provide clear indications of what is “best” in model comparison, and why. This is not a minor issue, because even if building a MS is per se science, the research which can be more easily abstracted into modelling produces results mostly at process (i.e., sub-model) level. Although the need for a finer granularity of model units, at least to avoid duplication, is a declared goal of the modelling community since many years [e.g., Argent 2004], technological bottlenecks have precluded model reuse for years, which is at the basis of the concrete opportunity of model hybridization as a development practice. One step forward to overcome technological constraints within a development environment has been given by modelling platforms, leveraging on object oriented programming to assist in discretizing the implementation of concepts into software units. The adoption of component oriented programming to build discrete, cross modelling platforms, software units [Donatelli et al. 2006a, 2006b], making available modelling platform-independent, multi-model, software components, has further facilitated the composition of modelling approaches.

Soil temperature (ST) is a state of the system of primary importance to simulate several processes. In fact, there is a huge literature about the importance of ST in driving the biochemical and physical processes involved with both the productivity and the sustainability of agricultural lands [e.g., Belviso et al. 2010]. In spite of the importance of ST data for the understanding of such a variety of processes, the availability of measured ST data is mostly limited to research sites. In any case, being ST internal to the system of interest and driven by changes of several states, a measurement from a reference site cannot be used as external driving force, as it is done with air temperature. The need for reliable ST data led to the development of models for their estimation, characterized by different degree of adherence to the physical processes involved [e.g., Campbell 1985; Elias et al. 2004]. Among the inputs needed by such models, soil water content along the soil profile plays a major role, and it represents – in addition – another variable usually not present in large-area databases, therefore often requiring simulation with field parameterization as well [Basile et al. 2003]. The simulation of ST hence requires a MS to account for the processes involved, each that can be simulated with various approaches. The objectives of this paper are then (i) to discuss the methodology for hybridization and evaluation of MS and (ii) to present a case based on nine MS resulting from three widely known simulation packages.

2 MATERIALS AND METHODS

2.1 Simulation experiment design

Nine MS were built by combining (i) three different approaches for the simulation of soil water dynamics (ii), two surface temperature models and (iii) two models for the simulation of ST along the profile (Figure 1). The simulation of crop growth and development was carried out using the generic crop simulator CropSyst, whereas evaporation was simulated with the approach proposed by Ritchie et al. [1972] and implemented in the EPIC model; the EPIC approach for root water uptake was also used. The approaches tested for the simulation of soil water redistribution among soil layers are: cascading (SW_c), cascading with travel time (SW_{ctt}), and an approach based on a finite difference solution of the Richards' equation (SW_R). The SW_c approach (also known as ‘tipping bucket’) is the most simplified and assumes that water can move only downward through the soil profile, filling up the layers until field capacity is reached, with the fraction of water exceeding this threshold moving to the deeper layer. The SW_{ctt} approach is a modification of cascading in which water movements are reduced by soil hydraulic conductivity, thus allowing water contents to be higher than field capacity. This approach is adopted in various simulation models [e.g., SWAT, Neitsch et al. 2002]. The SW_R approach [Richards 1931] is based on the physical concept that water flux between two points is driven by the pressure gradient between the points and it is function of the hydraulic conductivity. The approaches used for the simulation of soil surface temperature are the one proposed by Parton (SST_P) [Parton 1984] and the one implemented in the SWAT (SST_{SWAT}) model [Neitsch et al. 2002]. The two approaches tested for the simulation of ST along the soil profile are those proposed by Campbell (ST_C) [1985] and the one implemented in the SWAT model (ST_{SWAT}) [Carslaw and Jaeger 1959]. A MS is hence composed by (i) a soil water redistribution model ($SW_{??}$), (ii) a soil surface temperature model ($SST_{??}$) and (iii) a soil temperature model ($ST_{??}$). All the approaches for the simulation of soil hydrology and temperature are implemented in the software components UNIMI.SoilW and UNIMI.SoilT, respectively, freely downloadable at http://agsys.cra-cin.it/tools/.

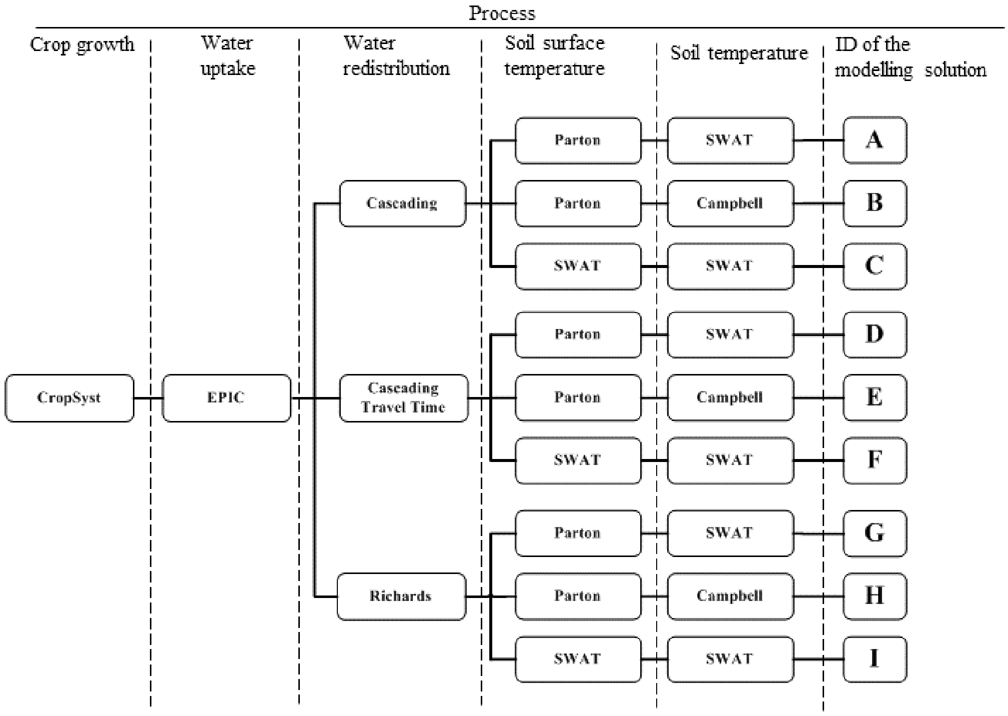

Figure 1. Modelling solutions developed and evaluated.

2.2 Data set description

ST data used for this study were collected in four experimental sites, three located in Northern Italy (Ghisalba, Mantova and Lodi), and one in Southern Italy (Eboli). The northern sites are placed in the Po valley, while the other is in the alluvial plain of the Sele river. These datasets are characterized by a large soil texture variability, with Eboli presenting a clay soil, Ghisalba a silty-loam soil, Lodi a typical loam texture, and Mantova a silty-clay-loam soil (USDA classification). For all the years and datasets, maize was cultivated in standard growing conditions according to the management practices typical of the study areas. ST measurements were carried out using probes placed at three different depths, aiming at exploring the regions where microbial activity is higher. Measurements were carried out using a PT100 sensor connected to a Cr10 (Campbell Scientific Ltd, Utah, USA) data logger; four measurements per day were recorded; daily average was extrapolated from them and it was used as reference for MS comparison.

2.3 Model outputs evaluation

The evaluation of the nine MS was carried out by developing an ad hoc fuzzy-based, modular indicator (I_{ST}) (Figure 2). I_{ST} partially extends the structure of the indicators proposed by Bregaglio et al. [2010] for evaluating different approaches for generating hourly air relative humidity and by Confalonieri et al. [2010b] for the assessment of hydrological models, with the latter considering the vectors of measured and simulated values for the different soil layers as not independent. Detailed description about the fuzzy-based aggregation procedure is provided by Bellocchi et al. [2002]. I_{ST} allows to consider: (i) the accuracy of the MS in fitting measured data (module Accuracy), (ii) the correlation between measured and simulated ST matrices (module Correlation), and (iii) the ability of the MS to maintain the same degree of accuracy under different conditions of applications (module Robustness). The three modules are composed by one or more simple metrics The aggregation of the metrics in their respective module and of the modules in the final indicator I_{ST} is achieved via a 4-stage design inferring system of fuzzy based rules, as described in Figure 2. The discussion of the results follows a top-down logic, considered particularly suitable for composite indicators since the final-level one (i.e., I_{ST}) provides an overall overview of the MS performances, whereas the values at lower levels of aggregation (modules and simple metrics) clarify the MS behaviour with respect to the specific features evaluated.

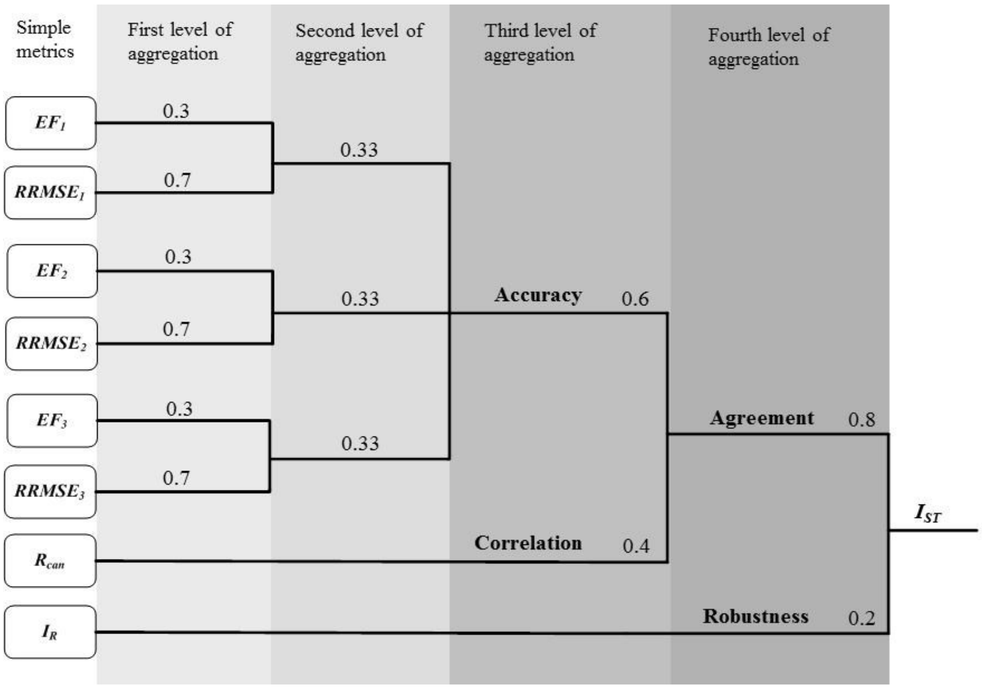

Figure 2. Design of the assessment method.

3 RESULTS

3.1 Model composition

When considering model composition, which is at the base of model development using different and alternate approaches, some terms have overloaded meanings, and at times confounded concepts can impact the activity. Firstly, two models are alternate options for simulating a process primarily if they estimate the same output. Other factors, that is, for now, number of parameters and even inputs required, are aspects relevant to the MS resulting from the use of a specific approach. One controversial aspect is the impact of estimating parameters for MS with a diverse number of parameters. When and if parameters can be estimated, as it is often needed for instance for hydrological parameters in complex soil water models, another modelling layer is used. As far as such layer is available, its effectiveness impacts on the overall evaluation of the MS against reference data, and it should not be considered per se an absolute constraint in using the more complex approach. On the other hand, there is no such a thing as “universal validation” of biophysical models, being model evaluation specific primarily to the context represented by reference data. Another controversial aspect in the model composition procedure of simulation models relates to the time step of the modelling approaches being composed. The best method to handle this problem is, in theory, the adoption of the shortest time step among the ones of the processes composed but, in practice, the time step chosen in available simulation packages was determined by input data availability (e.g., daily weather data), also to handle the complicacy of code implementation. Provided that the time step of the MS is adequate with respect to the analysis to be done as above, it determines when integration of state variables occurs, and, in implementation terms, when different parts of the MS communicate. Consequently, models at fine granularity can be composed if a given output matches exactly the definition of the input required by another model. An evaluation should be done, prior to its use, on the correctness of the first model in producing the output, but once it is accepted that the model provides the specific output, using that model should not be questioned in model composition, and it should be evaluated, in an operational perspective and as discussed above, in the frame of the MS with respect to parameters and inputs availability.

3.2 Soil temperature modelling solutions comparison

The values of I_{ST}, of the modules and of the simple metrics obtained by the nine MS are reported in Table 1. The processes that had the strongest impact on the simulation of ST among the layers resulted: (i) the simulation of surface ST and, as expected, (ii) the simulation of ST across the soil profile. The choice of the approach for the simulation of soil water content, instead, had a lower – although not negligible – influence on the output evaluated. In fact, the MS implementing the same couple of models for the simulation of surface ST and ST among the layers (i.e., A, D and G; B, E and H; C, F and I) resulted similar.

By considering only the first layer, which is the nearest to the soil surface, it is noticeable that the SST_P approach seems to be more accurate than the SST_{SWAT} one in reproducing surface ST (mean *RRMSE* = 13.92 and 16.60, respectively).

More in detail, the solution SST_P + ST_{SWAT} proved to be the best one in reproducing measured data under the explored conditions, with an average value of I_{ST} of 0.144 and a limited variability due to the different approaches for the simulation of soil water redistribution. This led MS A, D and G to achieve the best values for I_{ST}, with MS A – including the SW_C approach – ranked 1st (I_{ST} = 0.125). The hybridization between the SST_P and ST_C led the related MS (i.e., B, E and H) to be ranked from 4th to 6th (average I_{ST} = 0.244). Within these MS, the best results (I_{ST} = 0.241) was achieved by the one adopting the SW_R approach for water redistribution across the soil profile. Finally, according to I_{ST}, the couple of models for surface ST and ST across the layers leading to the worst results was SST_{SWAT} + ST_{SWAT} (MS C, F, I), with an average I_{ST} equal to 0.284. Like for the solutions adopting the couple SST_P + ST_C, the hydrological model assuring the better performance was SW_R.

Moving to the second-level aggregated measures, the values obtained by the MS are better for the Agreement module with respect to Robustness one. In particular, it emerges that the best performances obtained by MS A are mainly due to its robustness, i.e., its ability of maintaining stable the degree of efficiency in reproducing measured data across different conditions, since the same MS was ranked 3rd according to the Agreement module. The SW_R approach allowed MS G to obtain the best Agreement value, although its Robustness penalized it during the computation of I_{ST}. MS B, E and H achieved the worst Robustness values, thus allowing to consider the performances of the hybrid SST_P + ST_C as more dependent upon the particular agro-meteorological conditions in which they are applied.

Table 1. Results of the simple metrics in each data set, average values of modules and final indicator. Italics= best result; grey =best result per metric.

MS	Site	*EF*	*RRMSE*	R_{can}	I_r	Accuracy	Correlation	Robustness	Agreement	I_{ST}
A	Mantova	0.290	9.462	0.942	3.305	0.363	0.098	0.425	0.131	0.125
	Lodi	*0.933*	11.810	0.988						
	Ghisalba	0.855	15.560	*0.993*						
	Eboli	0.841	*8.519*	0.991						
B	Mantova	-0.189	11.940	0.946	5.211	0.423	0.087	0.950	0.163	0.244
	Lodi	*0.913*	13.664	0.990						
	Ghisalba	0.808	18.567	*0.995*						
	Eboli	0.903	*8.668*	*0.995*						
C	Mantova	-0.107	13.089	0.938	4.255	0.471	0.182	0.756	0.250	0.319
	Lodi	0.869	16.197	0.989						
	Ghisalba	0.725	21.834	*0.990*						
	Eboli	*0.876*	*8.695*	0.981						
D	Mantova	0.205	9.780	0.944	3.555	0.356	0.100	0.522	0.128	0.157
	Lodi	*0.936*	11.587	0.987						
	Ghisalba	0.865	15.044	*0.992*						
	Eboli	0.863	*8.100*	0.991						
E	Mantova	-0.164	11.872	0.947	5.102	0.427	0.091	0.935	0.167	0.247
	Lodi	*0.915*	13.528	0.988						
	Ghisalba	0.797	19.164	0.994						
	Eboli	0.897	*8.903*	*0.995*						
F	Mantova	0.056	12.272	0.931	3.523	0.467	0.199	0.509	0.259	0.273
	Lodi	0.869	16.194	*0.989*						
	Ghisalba	0.723	21.914	*0.989*						
	Eboli	*0.877*	*8.697*	0.981						
G	Mantova	0.207	9.773	0.957	3.556	0.343	0.095	0.522	0.119	0.151
	Lodi	*0.937*	11.527	0.988						
	Ghisalba	0.889	13.983	0.991						
	Eboli	0.852	*8.293*	*0.992*						
H	Mantova	-0.226	12.129	0.948	5.441	0.418	0.088	0.975	0.160	0.241
	Lodi	*0.916*	13.403	0.990						
	Ghisalba	0.816	18.648	0.994						
	Eboli	0.899	*8.718*	*0.995*						
I	Mantova	0.003	12.480	0.941	3.625	0.464	0.165	0.549	0.234	0.259
	Lodi	0.866	16.493	*0.988*						
	Ghisalba	0.764	20.598	*0.988*						
	Eboli	0.290	9.462	0.986						

3.3 Model hybridization and model development

Comparing whole MS as finalized products for cropping system simulation has a clear role given a specific context, unchangeable modelling resources at a specific time, and assuming that the reference data to test the MS are adequate to limit the possible effect of misuse of calibration, degrading process based models to fully empirical models. However, it provides a very weak link to the evaluation of specific modelling approaches which are produced by research. Testing MS which differ by one modelling approach allows estimating its contribution to the predictive power in the specific context of interest. The reason for a specific performance in a comparison may very well be the result of the specific parameterization, but the important aspect is that by replacing a specific sub-model, the observed impact will be due to the sub-model in the specific context. The impact of model hybridization in fostering model development, both in terms of optimization of choices for process simulation, and of extension of model representativeness with respect to the real system, appears very relevant.

4. Conclusions

This paper has explored aspects of composition of models in the perspective of evaluating alternate modelling approaches, and of MS evaluation based on widely known simulation packages. The comparison run must be considered as a proof of concept of model composition and hybridization across MS. The results of this comparison should hence be read with respect to the database of reference data used, including the possible constraints in the availability of physically measured parameters. The usefulness of considering a broad range of metrics in model evaluation, although not providing statistical significance, allows getting an articulated insight on model performance. The subjective judgment in defining index and module weights, rather than being a weakness, may as well be the strength of the procedure. In fact, it can allow targeting the development of a new index to specific goals. The different ranks obtained by the nine MS when the ranking criterion (i.e., the metric) changed show evidence of differences in model performance which cannot be highlighted by a single metric, providing insight in the consequences of model hybridization.

The comparison of the hybrid MS tested during this work demonstrated that – under the explored conditions – the SST_P is the one guaranteeing the best performances for the simulation of surface temperature, whereas the ST_{SWAT} model resulted the most reliable among the approaches for ST in the soil profile. Moreover, the fact that the ST_{SWAT} led to the best and to the worst results when coupled with the SST_P and the SST_{SWAT} approaches, respectively, allows concluding that (i) the surface ST is the process with the highest impact on the output evaluated and (ii) a hybrid solution can be better than the pure 'parent' ones (SST_{SWAT} - ST_{SWAT} in this case). The similar level of overall agreement presented by the MS implementing different approaches for soil water redistribution and the highest robustness achieved by the MS implementing SW_c seem to suggest the use of this simplest approach for soil water dynamics, especially if the estimation of ST is needed within large area operational studies, where specific calibration cannot be performed.

The availability of software components implementing libraries of models for the simulation of the different biophysical processes represents a clear step towards the improvement of simulation modelling science. The approach strongly stimulates modellers to look beyond the standard procedure of comparing models as a whole to identify the most suitable in the target context of use for the MS. In fact, it allows building suitable MS for specific objectives and conditions of applications, also via the hybridization of existing models. The development of MS, from composition to evaluation, should include the capability to evaluate via comparison the contribution of models at fine granularity, improving the transfer from research to the development of integrated model tools.

REFERENCES

Argent, R.M., An overview of model integration for environmental applications – components, frameworks and semantics, *Environmental Modelling and Software, 19,* 219-234, 2004.

Basile, A., G. Ciollaro, and A. Coppola, Hysteresis in soil-water characteristics as a key to interpreting comparisons of laboratory and field measured hydraulic properties, *Water Resources Research, 39,* 1355-1377, 2003.

Bellocchi, G., M. Acutis, G. Fila and M. Donatelli, An indicator of solar radiation model performance based on a fuzzy expert system, *Agronomy Journal*, *94,* 1222-1233, 2002.

Belviso, C., F. Cavalcante, P. Ragone, and S. Fiore, Immobilization of Ni by synthesizing zeolite at low temperatures in a polluted soil, *Chemosphere, 78,* 1172-1176, 2010.

Bregaglio, S., M. Donatelli, R. Confalonieri, M. Acutis, and S. Orlandini, An integrated evaluation of thirteen modelling solutions for the generation of hourly values of air relative humidity, *Theoretical and Applied Climatology, 102,* 429-438, 2010.

Bregaglio, S., M. Donatelli, R. Confalonieri, M. Acutis and S. Orlandini, Multi metric evaluation of leaf wetness models for large-area application of plant disease models, *Agricultural and Forest Meteorology, 151*, 1163-1172, 2011.

Brisson, N., C. Gary, E. Justes, R. Roche, B. Mary, D. Ripoche, D. Zimmer, J. Sierra, P. Bertuzzi, P. Burger, F. Buissière, Y.M. Cabidoche, P. Cellier, P. Debaeke, J.P. Gaudillère, F. Maraux, B. Seguin, and H. Sinoquet, An overview of the crop model STICS, *European Journal of Agronomy, 18,* 309-332, 2003.

Campbell, G.S., 1985. Soil physics with Basic. *Elsevier*, Amsterdam.

Carslaw, H.H., and J.C. Jaeger, Conduction of heat in solids, *Clarendon Press*, Oxford, 1959.

Confalonieri, R., S. Bregaglio, and M. Acutis, A proposal of an indicator for quantifying model robustness based on the relationship between variability of errors and of explored conditions, *Ecological Modelling, 221*, 960-964, 2010a.

Confalonieri, R., S. Bregaglio, S. Bocchi, and M. Acutis, An integrated procedure to evaluate hydrological models, *Hydrological Processes,* 24, 2762-2770, 2010b.

Donatelli, M., G. Bellocchi, E. Habyarimana, S. Bregaglio, and B. Baruth, AirTemperature: extensible software library to generate air temperature data, *SRX Computer Science*, 2010.

Donatelli, M., L. Carlini, and G. Bellocchi, A software component for estimating solar radiation, *Environmental Modelling and Software*, *21,* 411-416, 2006a.

Donatelli, M., G. Bellocchi, and L. Carlini, Sharing knowledge via software components: models on reference evapotranspiration, *European Journal of Agronomy, 24,* 186-192, 2006b.

Donatelli, M., and A. Rizzoli, A design for framework-independent model components for biophysical systems, in: iEMSs 2008, Sànchez-Marrè, M., Béjar, J., Comas, J. Rizzoli A., Guariso G. (Eds.) *International Environmental Modelling and Software Society (iEMSs)*, 7-10 July 2008, Barcellona, Spain, pp. 727-734, 2008.

Elias, E.A., R. Cichota, H.H. Torriani, Q. de Jong van Lier, Analytical soil-temperature model: correction for temporal variation of daily amplitude, *Soil Science Society of America Journal, 68,* 784-788, 2004.

Jones, J.W., G. Hoogenboom, C.H. Porter, K.J. Boote, W.D. Batchelor, L.A. Hunt, P.W. Wilkens, U. Singh, A.J. Gijsman, and J.T. Ritchie, The DSSAT cropping system model, *European Journal of Agronomy, 18,* 235-265, 2003.

Keating, B.A., P.S. Carberry, G.L. Hammer, M.E. Probert, M.J. Robertson, D. Holzworth, N.I. Huth, J.N.G. Hargreaves, H. Meinke, Z. Hochman, G. McLean, K. Verburg, V.O. Snow, J.P. Dimes, M. Silburn, E. Wang, S. Brown, K.L. Bristow, S. Asseng, S.C. Chapman, R.L. McCown, D.M. Freebairn, and C.J. Smith, An overview of APSIM, a model designed for farming systems simulation, *European Journal of Agronomy, 18,* 267-288, 2002.

Neitsch, S.L., J.G. Arnold, J.R. Kiniry, R. Srinivasan, and J.R. Williams, Soil and water assessment tool. User's manual. Grassland, soil and water research laboratory, *Agricultural Research Service*, Temple, Texas, 2002.

Parton, W.J., Predicting soil temperatures in a shortgrass steppe, *Soil Science, 138,* 93-101, 1984.

Richards, L.A., Capillary conduction of liquids through porous mediums, *Physics, 1,* 318-333, 1931.hg

Ritchie, J.T., Model for predicting evaporation from a row crop with incomplete cover, *Water Resources Research, 8,* 1204-1213, 1972.

Stöckle, C.O., M. Donatelli, and R. Nelson, CropSyst, a cropping systems simulation model, *European Journal of Agronomy, 18,* 289-307, 2003.