Jul 1st, 12:00 AM

A Comparison of Methods for Providing Solar Radiation Data to Crop Models and Decision Support Systems.

M. Rivington

K. B. Matthews

K. Buchan

Rivington, M.; Matthews, K. B.; and Buchan, K., "A Comparison of Methods for Providing Solar Radiation Data to Crop Models and Decision Support Systems." (2002). *International Congress on Environmental Modelling and Software*. 232.
http://scholarsarchive.byu.edu/iemssconference/2002/all/232

A Comparison of Methods for Providing Solar Radiation Data to Crop Models and Decision Support Systems.

M. Rivington, K.B. Matthews, and K. Buchan.
Macaulay Land Use Research Institute, Craigiebuckler, Aberdeen, UK. (m.rivington@macaulay.ac.uk)

Abstract: Models are increasingly being used to represent land uses within decision support systems. Crop simulation models often require climate data as input variables. Whilst precipitation and temperature data are usually available, there is often a dearth of representative solar radiation data in most countries. An illustration is made of the spatial distribution of meteorological stations with records of solar radiation and / or sunshine duration in the UK. Methods are available to estimate solar radiation using meteorological data, or by conversion from sunshine duration. In the absence of site-specific data, the nearest meteorological station data is often used to run models at a particular location. The aim of this paper is to determine an appropriate source of solar radiation data, given the range of meteorological data available, for the site-specific application of a crop model (CropSyst). Three methods of providing solar radiation data were tested: conversion from sunshine duration; the two nearest meteorological stations; and the Campbell-Donatelli model. Generic simulations of spring barley were run within CropSyst for 13 separate years, using the three sources of solar radiation data for three neighbouring locations in southern England. Crop yield output was compared with results derived from observed solar radiation. For the three locations tested, the order of most suitable data source was: conversion of sunshine duration; nearest meteorological station; and the Campbell-Donatelli model. There is a significant effect on model results arising from the data source. The results demonstrate that DSS employing crop models should use an appropriate source of solar radiation data. The results are discussed in the context of utilising CropSyst within the Land Allocation Decision Support System (LADSS), a spatial multiple-objective land-use planning tool for considering farm-scale environmental, social and economic trade-offs.

Keywords: Solar radiation; climate data; CropSyst; crop model; decision support system; LADSS.

1. INTRODUCTION

In Britain, as elsewhere, the availability of solar radiation data from meteorological stations is more restricted than for precipitation and temperature. Sunshine duration is more widely recorded. Many crop simulation models have a minimal requirement for data on daily precipitation, temperature and solar radiation as input variables. The lack of solar radiation data, synoptically synchronised with other data types, presents difficulties when a crop model is to be applied to a particular location. The problem is how to define what are spatially and temporally representative climate data to be used within crop models. Model accuracy can be significantly affected by the issues relating to missing or non-spatially representative data (Hoogenboom 2000). This paper investigates a range of sources for supplying solar radiation data to CropSyst, a cropping systems simulation model (Stöckle and Nelson 1998). CropSyst is to be used to represent crop-based land-uses within the Land Allocation Decision Support System (LADSS) (Matthews et al. 1999). This work has been undertaken as a first step in validating a complex decision support system, by focusing on key biophysical input data. The results are discussed in terms of the choice of data source on CropSyst output and the impacts on decision support.

2. PREDICTION OF SOLAR RADIATION

Cloudiness, atmospheric transmissivity, latitude and orientation of the Earth relative to the Sun, time of day, slope and aspect of the surface determine the spatial and temporal distribution of irradiance incident on a surface. A number of methods exist for conversion of sunshine duration to solar radiation. Each has a range of data input requirements.

2.1 Prediction Using Sunshine Duration.

Johnson et al. (1995) and Woodward et al. (2001) use sunshine duration to predict solar radiation. The model accounts for latitude, solar declination and elevation, day length and atmospheric transmissivity on a daily basis and has only daily sunshine duration (hours) as input. Total daily irradiance (J_o) is given by:

$$hJ_0 = sun.J_{0,s} + h.J_{0,d} \qquad (1)$$

where h is **day length**, *sun* is the sunshine duration, $J_{0,s}$ is the direct beam and $J_{0,d}$ the diffuse components. Daylength in hours is calculated by:

$$h = \frac{24}{\Pi}\cos^{-1}(-\tan\lambda\tan\delta) \qquad (2)$$

where λ is the **latitude** and δ is the **solar declination** in radians. Solar declination for each day of the year is given by:

$$\delta = -0.4084\cos\left(2\Pi\frac{d+10}{365}\right) \qquad (3)$$

where d is the Julian day of year. $J_{0,s}$ is given by:

$$J_{0.s} = 1367\frac{2p}{\Pi}\sin\phi\left(\tau^{1/\sin\phi}\right) \qquad (4)$$

where p is the fraction of radiation in full spectrum sunlight (here 1 is used) and 1367 is the solar constant (J m^{-2} s^{-1}). τ is the atmospheric **transmissivity** (see section 2.2). ϕ is the **solar elevation** at noon:

$$\sin\phi = \sin\lambda\sin\delta + \cos\lambda\cos\delta \qquad (5)$$

The diffuse portion of total irradiance is represented by $J_{0,d}$ (cloud conditions and from blue sky scattering simultaneously), can be calculated by:

$$J_{0,d} = J_{0,p}\left(f_{blue}(1-c) + f_{cloud}c\right) \qquad (6)$$

where c is the average daily fraction of cloud cover, given by: 1 – (sun duration / h), with $0<c>1$. $J_{0,p}$ is the total clear sky mean daily irradiance:

$$J_{0.p} = 1367\frac{p}{\Pi}\sin\phi\left(1+\tau^{1/\sin\phi}\right) \qquad (7)$$

The values of f_{blue} and f_{cloud} represent the relative different radiation intensities under blue sky and cloud conditions, respectively:

$$f_{blue} = \frac{1-\tau^{1\sin\phi}}{1+\tau^{1\sin\phi}} \qquad (8)$$

$$f_{cloud} = FF.f_{blue} \qquad (9)$$

Woodward et al. (2001) determined an empirical fixed factor (FF) of 1.11 for New Zealand. To optimise FF for the three test sites, daily FF values were fitted for each day per year for each complete year within the data set. The average of all years can then be used.

This method imposes a base-line amount of diffuse radiation, variable with h, such that an input of 0 sunshine hours will still produce a value of irradiance.

2.2. Transmissivity

The method detailed in Woodward et al. (2001) finds the clear sky transmissivity (τ) by:

$$\tau = 0.64 + 0.12\cos\left(2\Pi\frac{d-174}{365}\right) \qquad (10)$$

However, to produce site-specific τ values for the three test sites, we applied an alternative method by determining the extraterrestrial (potential) solar radiation (*Ra*) (Campbell and Norman 1998):

$$\frac{1}{\Pi}Ksc.dr\left[\omega_s\sin\lambda\sin\delta + \cos\lambda\cos\delta\sin\omega_s\right] \qquad (11)$$

where *Ksc* is the solar constant, here taken as 118.08 (MJ m^{-2} day^{-1}). *dr* is the inverse relative distance of the Earth to the Sun, given by:

$$dr = 1 + 0.0334\cos\left(\frac{2\Pi}{365}d\right) \qquad (12)$$

where d is the Julian day. The sunset hour angle (ω_s) in radians is calculated as:

$$\omega_s = \arccos[-\tan\lambda\tan\delta]$$
(13)

where λ is the latitude (radians) and δ is the solar declination. Having found *Ra* for each site, transmissivity was determined by averaging the seven highest atmospheric transmissivities between day of year 120 to 240 (Bechini et al. 2000):

$$\tau = obs / Ra \qquad (14)$$

where *obs* is the observed solar radiation at each site.

2.2 Prediction of Solar Radiation from Air Temperature

The Campbell-Donatelli model (CD) estimates daily global solar radiation by estimating the extraterrestrial (potential) radiation *Ra* and multiplying by the atmospheric transmissivity coefficient (tt_i) (Donatelli and Campbell 1998). The CD model is part of a suite of models contained within the RadEst global solar radiation estimation tool (beta v3.00)(ISCI 2001), were estimated radiation = tt_ipotential radiation. In the CD model, the value of tt_i is determined by:

$$tt_i = \tau\left[1-\exp\left(-b.f\left(T_{avg}\right)\left(\Delta T^2\right)f\left(T_{\min}\right)\right)\right] \quad (15)$$

Where:
τ = clear sky transmissivity,
$\Delta T = Tmax_i - (Tmin_i + Tmin_{i+1}) / 2$
$f(Tavg) = 0.017*\exp(\exp(-0.053*Tavg))$
$Tavg = (Tmax_i + Tmin_i) / 2$, f(Tavg)
$f(Tmin) = \exp(Tmin_i / Tnc)$

The data input requirements are T*max* and T*min*, the daily (*i*) maximum and minimum air temperatures. T*nc* and *b* are empirical parameters. For optimal model performance, location-specific parameters (τ, T*nc* and *b*) need to be determined by fitting to observed data. The τ, *Tnc* and *b* parameters were iteratively optimised for each location using the RadEst tool.

3. METHODS AND MATERIALS

3.1 Meteorological Stations

Three meteorological stations, Rothamstead, Bracknell and Wallingford, in southern England, were selected as they hold long-term records (1969 to 1998) for precipitation, maximum and minimum air temperature, solar radiation and sunshine duration. The distances (km) between sites are: Rothamstead to Bracknell = 55; Rothamstead to Wallingford = 57; Bracknell to Wallingford = 33. Site elevations (m a.s.l.) are: Rothamstead, 128; Bracknell, 74; and Wallingford, 48. Stations in Britain recording solar radiation and sun duration were mapped to determine the spatial distribution of available data.

3.2 Climate Data

Observed sun duration data from the three sites were converted to solar radiation values (MJ m^{-2} day^{-1}) using the method detailed in section 2.1. Estimated values were compared with observed solar radiation by regression analysis. The method used here to calculate τ was compared to that used by Woodward et al. (2001) (10) by comparison of root mean square error (RMSE). Values of solar radiation were calculated using the CD model within RadEst, based on observed max and min temperature. The following climate data sets were created for each site:

- Observed precipitation, max and min air temperature (Basic).
- Observed precipitation, max and min air temperature, solar radiation (Complete).
- Observed Basic + solar radiation converted from sun duration (from the same site).
- Observed Basic + solar radiation calculated by the CD model.

3.3 Simulation Testing

A standardised CropSyst scenario was created for a spring barley crop, such that the initial water and nitrogen were not limiting factors. CropSyst simulations were run for each site for each year where the years' climate data set did not contain any missing values, using the following sources:

- Observed Complete at the site (Obs)
- Observed Complete from nearest met station (NMS1).
- Observed Complete from second nearest met station (NMS2).
- Observed Basic at the site + solar radiation converted from sun duration (Sun)
- Observed Basic at the site + solar radiation calculated by the CD model (CD)

For simulations run with the Sun and CD data only the solar radiation varied, for the NMS, precipitation, temperature and solar radiation were the only variables. Crop physiological parameters were set according to previous (unpublished) calibration exercises, to produce an average yield of 5.5 tonnes/ha. All initialisation values were constant for each simulation. Planting data was always 16th March (day 75). The effects of data source on estimated yields were analysed using absolute differences between total accumulated yield (t/ha), standard deviation and RMSE for individual estimates per year. The significance of differences between means for each data source were analysed using a Monte Carlo sample difference test (Noreen 1989).

4. RESULTS

4.1 Spatial Distribution of Solar Radiation Data.

In Britain there are 68 meteorological stations with measured solar radiation, where as there are 1261 stations with measured sunshine duration with 29% and 12% having records for less than 5 years respectively. In the period 1969-98, there were 13 corresponding years where Rothamstead, Bracknell and Wallingford had complete data.

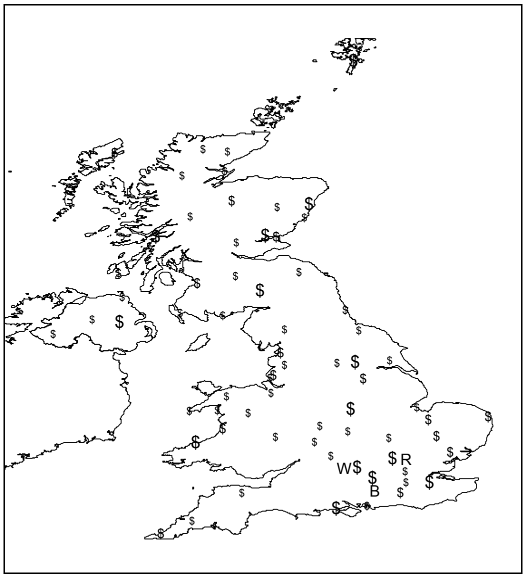

Figure 1. Locations of meteorological stations with measured solar radiation and length of record (years). ▲ 1-10; ▲ 11-21; ▲ 22-48. R = Rothamstead; B = Bracknell; W = Wallingford.

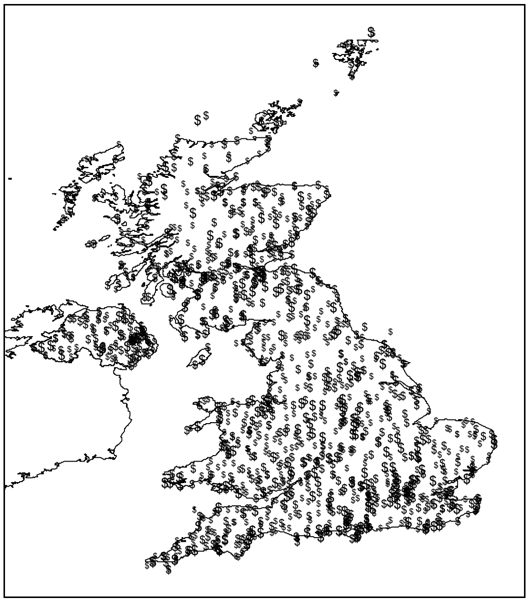

Figure 2. Locations of meteorological stations with measured sunshine duration and length of record (years). ▲ 3-13; ▲ 14-30; ▲ 31-103.

4.2 CropSyst Yield Output

Conversion of sunshine duration to solar radiation showed good regression fits. The minimum, mean and maximum R^2 for Rothamstead were 85.1, 91.2 and 94.2 (n = 25 years); Bracknell 90.5, 92.5 and 95.3 (n = 25 years) and Wallingford 85.3, 90.7 and 93.6 (n = 20 years) respectively. The method used to calculate transmissivity (Campbell and Norman 1998, Bechini et al 2000), produced similar results to that of Woodward et al. (2001) (10). RMSE for Rothamstead was 2.50 and 2.54, respectively, and for Bracknell 2.33 for both methods.

Yield estimates made by CropSyst are shown in Tables 1-3. **Total yield** is that accumulated for all 13 matching years. **Difference** is between observed and estimated total accumulated yield (13 years). **Max error** is the largest difference between observed and estimated yields per year. **StDev** is the standard deviation between individual years. **P** values are for means being equal.

	Obs	NMS1	NMS2	Sun	CD
Total yield	74.88	76.88	74.11	69.85	77.35
Difference		1.99	-0.77	-5.03	2.47
Max error		1.52	1.26	-1.55	1.84
Mean	5.76	5.91	5.70	5.37	5.95
StDev	0.75	0.41	0.38	0.81	0.23
RMSE		2.17	2.36	2.36	2.91
P value		0.527	0.813	0.220	0.415

Table 1. Yields (t/ha) for Rothamstead: NMS1 is Bracknell; NMS2 is Wallingford (n = 13).

	Obs	NMS1	NMS2	Sun	CD
Total yield	76.88	74.11	74.88	76.05	72.99
Difference		-2.77	-1.99	-0.82	-3.99
Max error		-0.54	-1.52	-0.44	-1.12
Mean	5.91	5.70	5.76	5.85	5.61
StDev	0.41	0.38	0.75	0.46	0.40
RMSE		0.94	2.17	0.67	1.70
P value		0.184	0.527	0.719	0.366

Table 2. Yields (t/ha) for Bracknell: NMS1 is Wallingford: NMS2 is Rothamstead (n = 13).

	Obs	NMS1	NMS2	Sun	CD
Total yield	74.11	76.88	74.88	75.40	70.17
Difference		2.77	0.77	1.29	-3.94
Max error		0.54	-1.26	0.28	-0.64
Mean	5.70	5.91	5.76	5.80	5.40
StDev	0.38	0.41	0.75	0.41	0.39
RMSE		0.94	2.36	0.50	1.35
P value		0.185	0.813	0.524	0.405

Table 3. Yields (t/ha) for Wallingford: NMS1 is Bracknell; NMS2 is Rothamstead (n = 13).

The ability of the different data sources to produce yield estimates that match those from observed climate data vary depending on the measure used. (Tables 1-3). The order of best fit per measure varies between the three locations (Table 4). For Rothamstead, the best overall fit was achieved using the NMS2 (Wallingford) data. For Bracknell, Sun data produced the best fit. At Wallingford, the best fit for difference between accumulated totals and mean of yearly yields came from the NMS2 (Rothamstead) data, whilst Sun gave the smallest maximum errors and RMSE. Sun produced the best result, 9.41 t/ha, for minimising the sum of absolute difference in yield for all years and sites, whilst NMS1, NMS2 and CD gave 11.2, 19.79 and 15.45 t/ha respectively.

		Order of best fit		
		Roth'd	Brack'l	Wall'd
Difference between totals	1	NMS2	Sun	NMS2
	2	NMS1	NMS2	Sun
	3	Sun	NMS1	NMS1
	4	CD	CD	CD
Smallest maximum error	1	NMS2	Sun	Sun
	2	NMS1	NMS1	NMS1
	3	Sun	CD	CD
	4	CD	NMS2	NMS2
Mean and P value (means are equal)	1	NMS2	Sun	NMS2
	2	NMS1	NMS2	Sun
	3	CD	CD	NMS1
	4	Sun	NMS1	CD
RMSE	1	NMS1	Sun	Sun
	2	Sun + NMS2	NMS1	NMS1
			CD	CD
	4	CD	NMS2	NMS2

Table 4. Order of best fit for each data source for different measures (from Tables 1-3).

	Error (t/ha)			
	NMS1	NMS2	Sun	CD
Largest over estimation	1.52	1.26	0.34	1.84
Largest under estimation	-0.54	-1.52	-1.55	-1.12
Accumulated difference	11.02	19.79	9.41	15.45
Years with error being > 1 t/ha	2	6	2	4

Table 5. Yield over and under-estimation errors per data source (t/ha).

CD gave the largest maximum error in individual years' estimates of yield (1.84 t/ha, at Rothamstead). With the exception of one estimate (-1.55 t/ha) Sun had the smallest maximum errors. The CD model maintained the smallest standard deviation, whilst Sun had the best fit of standard deviations to those of the observed. NMS2 and Sun maintained the best fit for means. No sources had means that were significantly different.

4 DISCUSSION

Though Bracknell and Wallingford are closer to each other (33 km) than to Rothamstead, there were closer similarities between simulations for Wallingford and Rothamstead (57km). Simulations for the Wallingford site using Bracknell's climate data produced less satisfactory results than when using Rothamstead's data. The nearest meteorological station does not, therefore, necessarily produce the closest fitting results. Selecting a suitable climate data source is difficult in locations where the climate is more strongly influenced by maritime and topographical factors.

The ability of the derived data sets to produce comparable results to the complete observed data sets depended on the measure used for comparison. The implications of the impact on decision support depend on how the model output is to be used, and the objectives of the DSS. Where interest is in predicting individual years' yield estimates, each data source had a high probability of producing large maximum errors, with *Sun* having the lowest probability. If the aim is to minimise absolute error in yield estimates, then *Sun* produced the best results. Conversely, to predict long-term average yields, then the NMS were the best options for two out of three of the locations. Output from crop models and DSS therefore has to be interpreted considering the effect of errors arising from the choice of input data and how the model output is used. One approach to dealing with errors arising from the input data is to set tolerance ranges to individual output estimates. If the criteria set for the DSS cannot tolerate a single years' error in estimate over 1 t/ha, in the present case, then all data sources would be rejected. Conversely, if the DSS can tolerate an average error of +/- 0.5 t/ha, then all data sources are acceptable.

LADSS uses land-use systems simulation models (for arable crops, livestock and forestry) to make estimates of productivity at the field scale. Productivity is then converted into a financial value, which is used within an optimisation process to compare between land-uses for the same field. Therefore, any error arising in the initial estimate of productivity will have down-stream consequences on the comparison between land-uses. This problem can be compounded if the error induced by input data disadvantages one land-use but favours another.

An important consideration in the site-specific application of a complex DSS is the requirement for parameterisation of individual components of the system. To reduce the cost of using the system, it should be a goal to minimise the amount of parameterisation work needed. The advantage of the data sources tested here is that none of them required other data types or additional parameterisation, other than the use of the RadEst automated optimisation of the CD models' parameters.

The differences in estimates of yield from observed, NMS1 and NMS2 data can partly be attributed to differences in precipitation and temperature, as well as solar radiation. CropSyst uses precipitation and temperature as the basis of crop development through calculations of evapotranspiration, transpiration, soil and crop water and nitrogen balances. Thermal time determines the phenological development of the crop. Differences in precipitation and temperature data between the test sites and their respective NMS resulted in different estimated crop growth rates, soil water and nitrogen budgets, and phenological development dates. The comparison of NMS1 and NMS2 with observed data does not indicate the role of solar radiation alone in determining the appropriateness of using the NMS climate data. However, to maintain synoptic synchronisation, it was not realistic to use site-observed precipitation and temperature and substituted solar radiation from another location.

5 CONCLUSIONS

Given the spatial distribution of meteorological stations with available data and the conversion method detailed here, sunshine duration appears to be a suitable substitute for measured solar radiation. The most appropriate data source varied between sites when considering all measures of best fit. For Rothamstead the best source was NMS2 (Wallingford), for Bracknell and Wallingford it was Sun and NMS2 (Rothamstead). The nearest meteorological station may not be the most representative of the site of model application. The results indicate that when applying crop models to a particular site for which solar radiation data is unavailable, that an appropriate order of data source to maximise site representation, is:

- Observed precipitation, temperature and sun duration converted to solar radiation.
- The nearest meteorological station with either solar radiation or sun duration (assuming geographic similarity and relative proximity).
- A calculation method such as the CD model, using site-specific precipitation and temperature data, and optimised parameters, provides a useable replacement in the absence of solar radiation or sun duration.

The quality and representativeness of climate data is fundamental in the ability of crop models, whether used separately or within DSS, to make reliable estimates. The choice of solar radiation data source has a marked effect on model output. The interpretation of output from a DSS employing crop models should take into consideration the consequences of using a particular climate data source and how the DSS uses the model output.

ACKNOWLEDGEMENTS

The authors gratefully acknowledge the funding support of the Scottish Executive Environment and Rural Affairs Department.

REFERENCES

Bechini, L., G. Ducco, M, Donatelli, and A. Stein, Modelling, interpolation and stochastic simulation in space and time of global solar radiation. *Agriculture, Ecosystems and Environment* 81, 29-42. 2000.

Campbell, G.S., and J.M. Norman, *An Introduction to Environmental Biophysics (2nd Edition)*, Springer-Verlag, 147-181. 1998.

Donatelli, M. and Campbell, G.S., A simple model to estimate global solar radiation. Proceedings of the 5th European Society of Agronomy Congress, Nitra, Slovak Republic, 2:133-134. 1998.

Hoogenboom, G., Contribution of agro-meteorology to the simulation of crop production and its applications. *Agriculture and Forest Meteorology*, 103, 137-157. 2000.

ISCI – Crop Science, Via di Corticella, 133, 40128 Bologna, Italy.

Johnson, I.R., S.J. Riha, and D.S Wilks, Modelling daily net canopy photosynthesis and its adaptation to irradiance and atmospheric CO_2 concentration. *Agricultural Systems*, 50, 1-35. 1995.

Matthews, K.B., A.R. Sibbald, and S. Craw, Implementation of a spatial decision support system for rural land use planning: integrating GIS and environmental models with search and optimisation algorithms, *Computers and Electronics in Agriculture,* 23, 9-26. 1999.

Noreen, E.W., *Computer-intensive methods for testing hypotheses: an introduction*. John Wiley, 240pp, London. 1989.

Stöckle, C.O. and Nelson, R. *CropSyst Users Manual*. Washington State University, Pullman, WA. 1998.

Woodward, S.J.R., D.J. Barker, and R.F. Zyskowski, A practical model for predicting soil water deficit in New Zealand pastures. *New Zealand Journal of Agricultural Research*, 44, 91-109. 2001.