Available online at www.sciencedirect.com

SCIENCE DIRECT®

Agricultural Systems 90 (2006) 5–31

AGRICULTURAL SYSTEMS

www.elsevier.com/locate/agsy

Review

Performance of CERES-Rice and CERES-Wheat models in rice–wheat systems: A review

J. Timsina *, E. Humphreys

CSIRO, Land and Water, Research Station Road, Hanwood, Griffith, NSW 2680, Australia

Received 1 December 2004; received in revised form 22 September 2005; accepted 23 November 2005

Abstract

Rice–wheat (RW) systems are critical to food security and livelihoods of rural and urban poor in south Asia and China, and to regional economies in southeast Australia. The sustainability of RW systems in south Asia is, however, threatened by yield stagnation or decline, and declining partial factor productivity, soil organic C and water availability. Crop models potentially offer a means to readily explore management options to increase yield, and to determine trade-off between yield, resource-use efficiency and environmental outcomes. This paper reviews the performance of CERES-Rice and CERES-Wheat in Asia and Australia in relation to their potential application towards increasing resource use efficiency and yield of RW systems.

The performance of the models was evaluated using simulated and observed data on anthesis and maturity dates, in-season LAI and growth, final grain yield and its components, and soil water and N balances from published studies across Asia and Australia, and then by computing the statistical parameters for the major characters. Over the four data sets examined for anthesis and six for maturity dates, CERES-Rice predicted those dates fairly well (normalised RMSE = 4–5%; *D*-index = 0.94–0.95), but over the 11 sets for grain and 4 for biomass yield, the predictions were more variable (normalised RMSE = 23% for both; *D*-index 0.90 and 0.76, for grain and biomass, respectively). Model performance was poorer under conditions of low N, water deficit, and low temperatures during the reproductive stages. Over the three data sets examined, CERES-Wheat predicted the anthesis and maturity dates quite well (normalised RMSE = 4–5%; *D*-index = 0.94–0.99), and over eight sets for grain and two sets for biomass

---

* Corresponding author. Tel.: +61 2 6960 1571; fax: +61 2 6960 1600.
E-mail address: jagadish.timsina@csiro.au (J. Timsina).

0308-521X/$ - see front matter
doi:10.1016/j.agsy.2005.11.007

yield the model predicted them also reasonably well (RMSE = 13–16%; *D*-index = 0.86–0.97). Only one study evaluated the DSSAT RW sequence model with fairly satisfactory predictions of rice and wheat yields over 20 years with adequate N, but not the long-term change in soil organic C and N. Predictions of in-season LAI and crop growth, and soil and water processes were quite limited to investigate the robustness of model processes.

Application of models to evaluate options to increase water and N use efficiency requires the ability to perform well at the margin where deficit stress begins. While both models generally perform satisfactorily under water and N non-limiting conditions, the little evidence available suggests that they do not perform well under resource-limiting situations. We recommend that the models' key processes under the water and N limiting conditions be further evaluated urgently. The DSSAT sequence model also needs to be further evaluated against observations for a range of locations and management using data from long-term experiments in RW systems.

Keywords: CERES-Rice; CERES-Wheat; Rice–wheat system; Model evaluation; Asia; Australia

---

Contents

---

1. Introduction

Rice and wheat are the two most important crops in Asia and Australia. While predominantly grown as monoculture, these crops are also grown in annual sequences in sub-tropical climates on about 13 Mha in the Indo-Gangetic Plains (IGP) of south Asia and on about 3.4 Mha in the Yangtse River Basin in China (Timsina and Connor, 2001; Dawe et al., 2004), and in less intensive, highly mechanised conditions in the temperate climate of south eastern Australia (Humphreys et al., 2004). Rice–wheat (RW) systems are critical to food security and fundamental to employment, income, and livelihood for hundreds of millions of rural and urban poor of south Asia and China, and also contribute to national revenue to varying degrees (Paroda et al., 1994; Timsina and Connor, 2001). However, the sustainability of RW system is threatened by stagnating or declining yields, increasing gaps between potential, experimental, and farmers' yields, and declining partial factor productivity, soil organic carbon, and water availability (Timsina and Connor, 2001; Ladha et al., 2003; Humphreys et al., 2004).

Rice–wheat systems are complex from both bio-physical and socio-economic viewpoints. Understanding the problems and developing solutions for RW systems require systematic research, the resources for which are shrinking globally. Moreover, field experimentation can only be used to investigate a limited number of variables, in a reductionist manner, under a few site-specific conditions. Crop models, on the other hand, can potentially be used to integrate knowledge of the bio-physical processes governing the plant–soil–atmosphere system, to evaluate the production uncertainties associated with various management options, and to extrapolate results to other sites and climates. Simulating RW sequences requires models that can describe the dynamic and, at times, extreme changes in soil hydrological conditions associated with growing continuously flooded rice in sequence with wheat grown in unsaturated soil, and their profound impacts on nutrient dynamics, especially N.

There are many rice and wheat simulation models but only a few that simulate crop sequences. Initiatives for modelling RW sequences began as early as 1994 through collaboration between the International Rice Research Institute (IRRI) and Wageningen University, and between IRRI and the International Fertilizer Development Centre (IFDC). The former collaboration through the Systems Analysis for Rice Production (SARP) project resulted in the development and validation of a RW prototype model for wheat after rice in a tropical environment in the Philippines (Bouman et al., 1994; Timsina et al., 1994), while the latter resulted in the validation and application of CERES-Rice and CERES-Wheat (Crop Estimation through Resource and Environment Synthesis, CERES) for RW sequences of northern Bangladesh and northwest India (Timsina et al., 1995, 1996, 1998).

The decision support system for agrotechnology transfer (DSSAT) developed by the International Benchmark Systems Network for Agrotechnology Transfer (IBSNAT) contains multiple crop models including CERES-Rice and CERES-Wheat, and provides a facility for simulating crop sequences. The DSSAT/CERES models simulate crop growth, development, and yield taking into account the effects of weather, management, genetics, and soil water, C and N. These models are continuously being refined and modified. The latest version (DSSAT ver. 4.0) differs from earlier versions with minor changes to growth and water balance processes, but large changes to integration procedures, modularity and the user interface (Porter et al., 2003; Jones et al., 2003; Hoogenboom et al., 2004). The model processes for CERES-Rice and CERES-Wheat are documented in a fragmented way in various publications (Table 1). Some researchers have also modified the CERES code to improve or add processes relevant to their applications, or to build new models, such

Table 1
Sources of documentation of model processes in CERES-Rice and CERES-Wheat

Crop	Process	Documentation
All cereals	Phenology	Ritchie and NeSmith (1991); Ritchie et al. (1998)
Rice and wheat	Phenology	Alociljha and Ritchie (1991); Hodges and Ritchie (1991); Singh (1994); Singh et al. (1999)
All crops	Water balance	Ritchie (1998)
All cereals	N and soil organic matter	Godwin and Jones (1991); Singh (1994); Godwin et al. (1990); Godwin and Singh (1991, 1998); Gijsman et al. (2002)
Rice	Cold injury	Godwin et al. (1994); Timsina et al. (2004)

as for methane generation and emission (MERES, Methane Emissions from Rice Ecosystems, Matthews et al., 2000a), tillage (Andales et al., 2000), salinity (Castrignano et al., 1998; Humphreys et al., 2003a,b), cold injury (Godwin et al., 1994; Timsina et al., 2004) and seedling growth and transplanting shock (Salam et al., 2001).

All crop models should be evaluated in the environment of interest if the results of applications are to be credible. To facilitate the model evaluation process, Hunt and Boote (1998) provided the data required for the calibration and validation of, while Tsuji et al. (1994) provided the descriptions of input and output file structures for, the DSSAT models. There are many studies on the calibration and validation of CERES-Rice and CERES-Wheat around the world, but there has been very little quantitative and systematic evaluation on their performance using robust statistical criteria, and no attempt to synthesise the results of those evaluations. The objectives of this review are: (i) to compile and analyse results of the performance of CERES-Rice and CERES-Wheat using various statistical criteria; (ii) to identify the capabilities and limitations of these models for application in RW systems; (iii) to identify future needs to enable the models to be applied to contribute to solving the resource and food security problems of these systems.

2. Procedure for model evaluation

Model evaluation requires well-defined criteria. Several statistical methods for analysing model performance are available (Jones and Kiniry, 1986; Willmott, 1982; Willmott et al., 1985). Jones and Kiniry (1986) used means and standard deviations (SD) for observed and simulated data, and linear regression parameters such as intercept (α), slope (β), and coefficient of determination (R^2). However, Thornton and Hansen (1996) concluded that use of the F-test (to test the null hypothesis that the regression line has unit slope and an intercept of zero) can be severely misleading. They showed that the probability of rejection of a valid model may increase with sample size, and pointed out that this is not an acceptable behaviour for a validation test of complex simulation models. Willmott (1982) and Willmott et al. (1985) recommended the use of RMSE or RMSD (root mean square error or deviation), RMSEs (root mean square error systematic), RMSEu (root mean square error unsystematic), and *D*-index (index of agreement), but suggested that RMSE is the "best" measure as it summarises the mean difference in the units of observed and predicted values. RMSEs indicates the bias (deviation of the actual slope from the 1:1 line) compared with the random variation (RMSEu) that may occur. The *D*-index is a descriptive (both relative and bounded) measure, and can be applied to make cross-comparisons between models (Willmott, 1982; Willmott et al., 1985). Kobayashi and Salam (2000) used mean squared deviation (MSD, the square of RMSE or RMSD), but concluded that RMSD is probably more intuitive than MSD because the deviation is expressed on a relative basis against observations.

To evaluate the performance of the models, we compiled the published data on simulated and observed results from several studies across Asia and Australia, and then computed a range of statistical parameters for each data set where these had not been originally presented by the authors (Tables 4 and 5). These included means, SDs, α, β, R^2, RMSE, RMSEs, RMSEu, relative RMSE, MSD, and *D*-index. A model reproduces experimental data perfectly when α is 0, β is 1, R^2 is 1, RMSE (and RMSEs) is 0, RMSEu approaches the value of RMSE, and *D*-index is 1. These parameters were calculated as follows:

$$\mathrm{RMSE} = \left[N^{-1} \sum_{i=1}^{n} (P_i - O_i)^2 \right]^{0.5}$$

$$\mathrm{RMSEs} = \left[N^{-1} \sum_{i=1}^{n} \left(\widehat{P}_i - O_i \right)^2 \right]^{0.5}$$

$$\mathrm{RMSEu} = \left[N^{-1} \sum_{i=1}^{n} \left(P_i - \widehat{P}_i \right)^2 \right]^{0.5}$$

$$\text{Normalised (relative) RMSE} = \mathrm{RMSE}/\bar{O}$$

$$\text{D-index} = 1 - \left[\sum_{i=1}^{n} (P_i - O_i)^2 \Big/ \sum_{i=1}^{n} \left[|P'_i| + |O'_i| \right]^2 \right]$$

$$|P_i| = a + bO_i$$

$$P_i = P'_i - \bar{O};\; O'_i = O_i - \bar{O}$$

where P_i and O_i are predicted and observed values, and $\bar{O}$ is the mean observed value over several replicates. $\bar{O}$ has an associated standard deviation which is often ignored in model evaluation, and any difference between simulations and observations is attributed solely to model inadequacies (e.g., Kobayashi and Salam, 2000; Gauch et al., 2003). In this review, model performance was judged on various qualitative and quantitative criteria, but with more weight on RMSE and *D*-index.

3. Calibration of models

Model calibration or parameterisation is the adjustment of parameters so that simulated values compare well with observed ones. The genetic coefficients used in CERES models characterise the growth and development of crop varieties differing in maturity. The coefficients for CERES-Rice and CERES-Wheat reported in the literaure are summarised in Tables 2 and 3. The coefficients for rice from Asia are mostly for short duration Indica varieties, while the two from Australia are for long duration Japonicas, one of which (Calrose), a tall variety, is no longer grown. The coefficients for wheat variety RR21 derived from separate studies in India and Nepal were very similar (Timsina et al., 1995, 1997), but there were differences for HD2329 derived by Hundal and Kaur (1997) and Pathak et al. (2003) using CERES-Wheat ver. 2 and ver. 3.5, respectively. The reasons for differences in the coefficients in the latter two studies for the same genotype are unknown, and neither report provided data or methodologies used to calculate the coefficients. The coefficients reported in Tables 2 and 3 were mostly derived from results of a single site, soil, season and sowing date, without explicitly reporting any methodology or even without the values of the coefficients. Many studies did not even report the model version used, nor provide cultivar names or phenological information, making it difficult for the reader to calculate or assign any coefficients. Mall and Aggarwal (2002), however, used data from more than one site or climate to derive the coefficients of seven rice cultivars, and Hundal and Kaur (1997) used more than one year of data to derive the coefficients for a wheat cultivar. Coefficients derived from observations for a range of sowing dates and seasonal conditions are likely to be more robust than those derived from a limited range of conditions.

4. Evaluation of CERES-Rice

CERES-Rice has been evaluated for many tropical and sub-tropical locations across Asia and in temperate climates in Japan and Australia (Table 4; Figs. 1 and 2). Most of these studies, however, provided only qualitative or graphical analysis, and many did not provide sufficient detail to calculate statistical parameters for quantitative evaluation of the models.

4.1. Phenology

Tropics. Alociljha and Ritchie (1991) reported good agreement between observed and predicted number of days to anthesis and maturity, with normalised RMSE of 4% and 3%, and *D*-index of 0.65 and 0.87, respectively for three upland rice cultivars (IR 43, UPLRi 5, UPLRi 7) in the Philippines. For three sites in Bangkok, with ver. 2.0, Tongyai (1994) concluded that the number of days to physiological maturity was overestimated by 9–12 d. In north and northeast Thailand, also with ver. 2.0, Jintrawat (1995), however, reported accurate predictions of phenology for both photoperiod sensitive and insensitive cultivars, but the heading dates were underestimated

Table 2
Genetic coefficients for rice cultivars, grown in Asia and Australia, determined for CERES-Rice

Cultivar	Type	Duration	Stature	Location	Soil	Coefficients							Source
						P1	P2O	P2R	P5	G1	G2	G3	
BR14	Indica	Short (130 d)	Semi-dwarf	Nashipur, Bangladesh	Sandy clay loam	560	11.5	200	500	45	0.026	1	Timsina et al. (1998)
BR11	Indica	Medium (150 d)	Semi-dwarf	Nashipur, Bangladesh	Sandy clay loam	825	11.5	300	390	52	0.024	1	Timsina et al. (1998)
Pant-4	Indica			Pantnagar, India	Silty clay loam	830	11.4	160	300	35	0.03	1	Timsina et al. (1995)
Masuli	Indica	Long	Semi-dwarf	Nepal		830	11.4	200	600	35	0.03	1	Timsina et al. (1997)
Chaite-2	Indica	Short	Dwarf	Nepal		560	11.5	200	500	45	0.026	1	Timsina et al. (1997)
IR52	Indica	Short	Semidwarf	Philippines		425	11.8	125	454	72	1	1	Hoogenboom et al. (1997)
Jaya	Indica			Kerala, India	Loam	830	15	50	277	72.8	0.028	1	Saseendran et al. (1998a,b)
IR36	Indica	Short	Semidwarf	IRRI, Philippines		450	11.7	149	350	68	0.023	1	Hoogenboom et al. (1997)
IR8	Indica			Pantnagar, India	Loam, silty clay loam	880	12.1	52	550	65	0.028	1	Timsina et al. (1995)
IR72	India	Short	Semidwarf	Philippines		400	12	100	580	76	0.023	1	Hoogenboom et al. (1997)
Amaroo	Japonica	Long	Semidwarf	Southern NSW, Australia	Mundiwa clay loam	370	14.5	750	85	80	0.026	1	Meyer et al. (1994)
Calrose	Japonica	Lonh	Tall	Southern NSW, Australia	Mundiwa clay loam	325	14	600	92	75	0.025	1	Meyer et al. (1994)
PR106	Indica			Several sites, India	–	500	11.5	150	300	60	0.024	1	Pathak et al. (2003)
China early	n/a			n/a		100	12.4	130	332	55	0.025	1	Hoogenboom et al. (1997)
China late	n/a					100	12	166	328	55	0.025	1	Hoogenboom et al. (1997)

Table 3
Genetic coefficients for spring wheat cultivars, grown in Asia and Australia, determined for CERES-Wheat

Cultivar	Stature	Duration	Location	Soil	Climate	Coefficients							Source
						P1V	P1D	P5	G1	G2	G3	PHINT	
Kanchan	Semi-dwarf	Medium (105 d)	Nashipur, Bangladesh	Sandy clay loam	Sub-tropical	1.0	1.3	4.3	5.0	4.0	2.7	99	Timsina et al. (1998)
Sowgat	Semi-dwarf	Medium (108 d)	Nashipur, Bangladesh	Sandy clay loam	Sub-tropical	1.0	3.0	4.0	5.5	4.5	2.7	97	Timsina et al. (1998)
RR21	Semi-dwarf	Short	Pantnagar, India	Loam, silty clay loam	Sub-tropical	1.0	1.5	5.0	4.0	2.9	2.4	95	Timsina et al. (1996)
RR21	Semi-dwarf	Short	Nepal	–	Sub-tropical	1.0	1.5	5.0	3.5	2.9	2.4	90	Timsina et al. (1997)
UP262	Semi-dwarf	Medium	Nepal	Loam, silty clay loam	Sub-tropical	1.0	3.0	5.0	4.0	2.9	2.4	90	Timsina et al. (1997)
HD2009	Semi-dwarf	Medium	Pantnagar, India	Loam, silty clay loam	Sub-tropical	1.0	2.3	4.0	4.0	2.9	2.4	95	Timsina et al. (1996)
HD2329	Semi-dwarf	Medium	Ludhiana, India	Sandy loam	Sub-tropical, semi-arid	0.5	2.5	3.5	2.5	2.9	4.0		Hundal and Kaur (1997)
HD2329	Semi-dwarf	Medium	Several sites, India	–	–	0.5	3.2	2.6	3.4	3.5	4.2	95	Pathak et al. (2003)
ANZA	Dwarf	Short	Los Baños, Philippines	Clay	Tropical	0.0	3.4	2.0	3.5	2.7	4.4	95	Timsina et al. (1996)
UPLW1	Dwarf	Short	Los Baños, Philippines	Clay	Tropical	0.0	4.0	7.0	5.0	4.0	3.0	95	Timsina et al. (1996)
UPLW3	Dwarf	Short	Los Baños, Philippines	Clay	Tropical	0.0	3.5	7.0	5.0	4.0	3.0	95	Timsina et al. (1996)
Janz	Semi-dwarf	Long	Southern NSW, Australia	Clay	Temperate, semi-arid	0.5	3.0	2.8	2.9	1.3	4.4	95	Humphreys et al. (2005)
Yecora	Dwarf	Short	Southern NSW	Clay	Temperate	0.5	2.8	2.0	3.2	3.0	4.4	95	Humphreys et al. (2005)
Bindawarra	Semi-dwarf	Long	Southern NSW	Clay	Temperate	0.5	3.2	2.5	2.6	2.5	4.4	95	Humphreys et al. (2005)

Table 4
Evaluation results for CERES-Rice predictions of phenology, grain and biomass yields for various data sets (simulated and observed data are from various studies reported under 'data sets')

Data sets	N	X_{obs} (SD)	X_{sim} (SD)	α	β	R^2	Absolute RMSE	RMSEs	RMSEu	RMSE normalised (%)	D-index
Grain yield (t/ha)											
Alociljha and Ritchie (1991)	18	4.9 (0.95)	4.3 (1.20)	2.91	0.46	0.34	1.14	0.64	0.95	23	0.70
Godwin et al. (1994)	12	4.6 (2.2)	4.6 (2.0)	0.46	0.88	0.64	1.30	0.59	1.16	29	0.96
Jintrawat (1995)	31	4.0 (1.2)	4.7 (1.3)	1.56	0.55	0.35	1.28	0.80	1	32	0.74
Timsina et al. (1995)	27	5.5 (1.5)	6.0 (1.9)	2.66	0.47	0.33	1.66	0.69	1.51	30	0.79
Timsina et al. (1998)	12	5.7 (2.2)	5.6 (2.8)	1.60	0.73	0.81	1.19	0.24	1.17	21	0.94
Amien et al. (1996)	26	5.2 (1.5)	4.4 (1.4)	0.95	0.96	0.83	0.98	0.81	0.56	19	0.89
Saseendran et al. (1998a,b)	8	5.2 (1.0)	5.3 (1.2)	0.52	0.89	0.99	0.18	0.14	0.10	3	0.99
Matthews et al. (2000a,b)	18	3.8 (1.2)	4.6 (1.5)	0.70	0.68	0.74	1.10	0.77	0.72	27	0.84
Rao et al. (2002) – 1993 data	14	5.0 (0.7)	4.9 (0.8)	0.75	0.87	0.98	0.17	0.13	0.11	3	0.99
Pathak et al. (2004)	58	5.3 (1.6)	5.3 (1.8)	1.11	0.79	0.84	0.72	0.98	0.72	18	0.95
Timsina et al. (2004)	25	6.4 (3.0)	6.2 (3.4)	1.25	0.82	0.86	1.26	0.22	1.25	20	0.96
All data sets	250	5.0 (1.8)	5.1 (1.9)	1.19	0.74	0.67	1.14	0.22	1.12	23	0.90
Biomass yield (t/ha)											
Alociljha and Ritchie (1991)	18	11.4 (2.2)	10.8 (2.7)	5.99	0.50	0.38	2.28	0.86	2.05	20	0.96
Saseendran et al. (1998a,b)	8	10.1 (1.5)	11.7 (1.1)	−1.01	0.95	0.45	1.89	1.74	0.74	19	0.56
Matthews et al. (2000b)	18	11.8 (3.2)	13.5 (3.1)	6.83	0.37	0.12	3.90	2.72	2.79	33	0.55
Pathak et al. (2004)	58	12.4 (2.9)	11.1 (3.5)	5.18	0.67	0.61	2.55	1.37	2.15	22	0.83
All data sets	102	12.0 (2.8)	11.5 (3.2)	5.89	0.53	0.37	2.74	0.96	2.57	23	0.76
Anthesis (d)											
Alociljha and Ritchie (1991)	15	94.9 (4.0)	94.9 (3.3)	47.18	0.50	0.17	3.86	2.55	2.91	4	0.65
Timsina et al. (1995)	13	95.3 (6.6)	91.7 (5.4)	26.37	0.75	0.39	6.24	4.73	4.08	7	0.72

(*continued on next page*)

Table 4 (*continued*)

Data sets	N	X_{obs} (SD)	X_{sim} (SD)	α	β	R^2	Absolute RMSE	RMSEs	RMSEu	RMSE normalised (%)	D-index
Timsina et al. (1998)	11	84.0 (12.9)	81.8 (13.4)	8.19	0.93	0.93	4.09	2.18	3.46	5	0.98
Saseendran et al. (1998a,b)	8	76.5 (2.8)	78.5 (4.4)	56.88	0.25	0.15	4.42	2.25	2.91	6	0.65
All data sets	47	89.3 (10.5)	88.2 (9.9)	5.70	0.95	0.80	4.77	1.98	4.34	5	0.94
Maturity (d)											
Alociljha and Ritchie (1991)	15	122.2 (6.6)	122.2 (4.8)	−13.94	1.11	0.66	3.79	2.64	2.73	3	0.87
Seino (1995)	4	133 (14.1)	133.3 (11.9)	−22.97	1.17	0.98	2.38	0.77	1.86	2	1.00
Timsina et al. (1995)	13	114.9 (15.7)	116.8 (16.8)	13.23	0.87	0.87	6.23	1.93	5.93	5	0.96
Saseendran et al. (1998a,b)	8	115.5 (6.1)	118.8 (4.3)	−38.0	1.29	0.84	4.15	3.82	1.63	4	0.84
Salam et al. (personal. communication)	4	108 (28.1)	107 (29.4)	6.8	0.94	0.97	4.50	2.94	5.09	4	0.99
Timsina et al. (1998)	18	120.6 (14.0)	119.3 (12.4)	−5.52	1.06	0.87	5.00	3.69	4.46	4	0.96
All data sets	62	118.9 (12.1)	119.4 (11.2)	1.88	0.98	0.83	4.94	1.89	4.56	4	0.95

N, number of data pairs; X_{obs}, mean of observed values; X_{sim}, mean of simulated values; SD, standard deviation; α, slope of linear relation between simulated and observed values; β, intercept of linear relation between simulated and observed values; R^2, adjusted linear correlation coefficient between simulated and observed values; RMSE absolute, root mean square error.

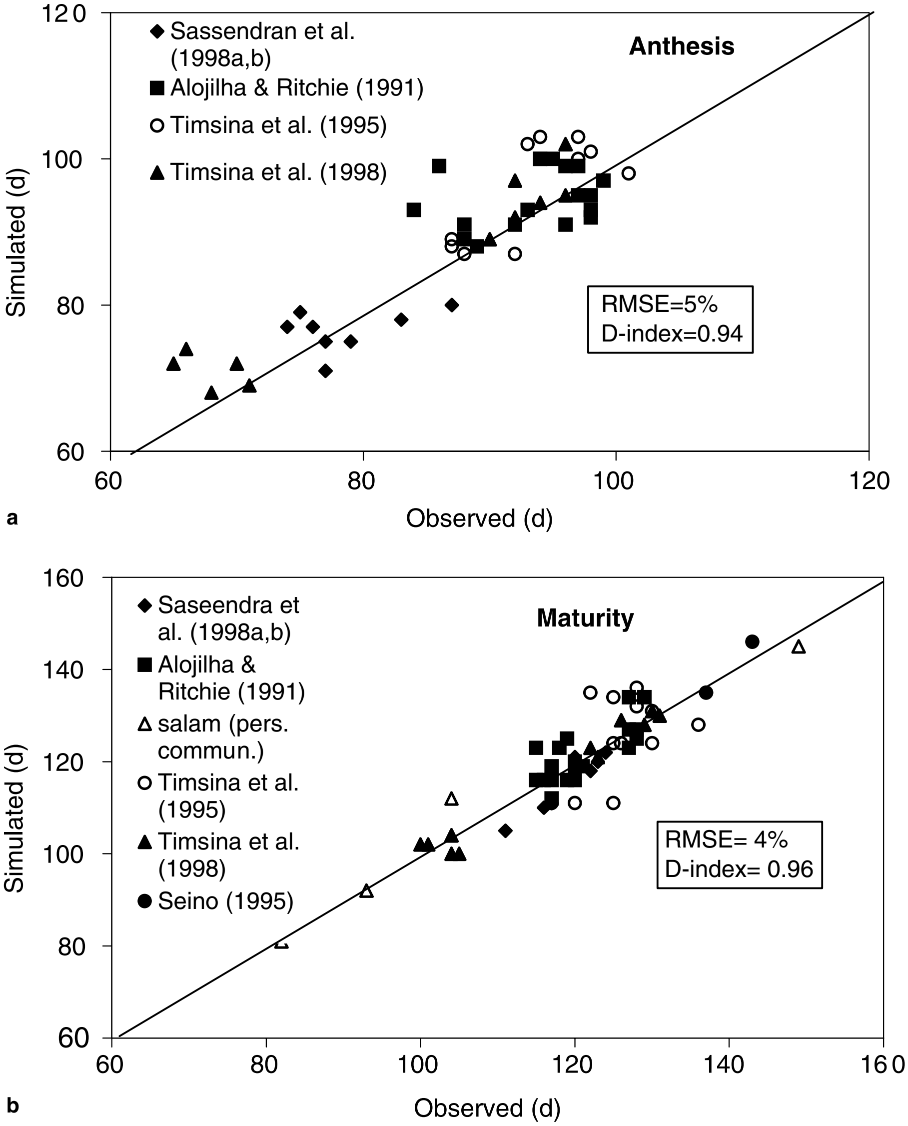


Fig. 1. Simulated and observed days to (a) anthesis and (b) maturity of rice across a range of experiments in Asia.

for a photo-sensitive cultivar, especially for early planting dates. The capability to simulate photoperiod effects has since been included in CERES-Rice. As a result, CERES-Rice accurately predicted the days to maturity of a photoperiod sensitive cultivar, KDML 105, at six sites in northeast Thailand, though the number of days to panicle initiation and anthesis were overpredicted (Boonjung, 2000).

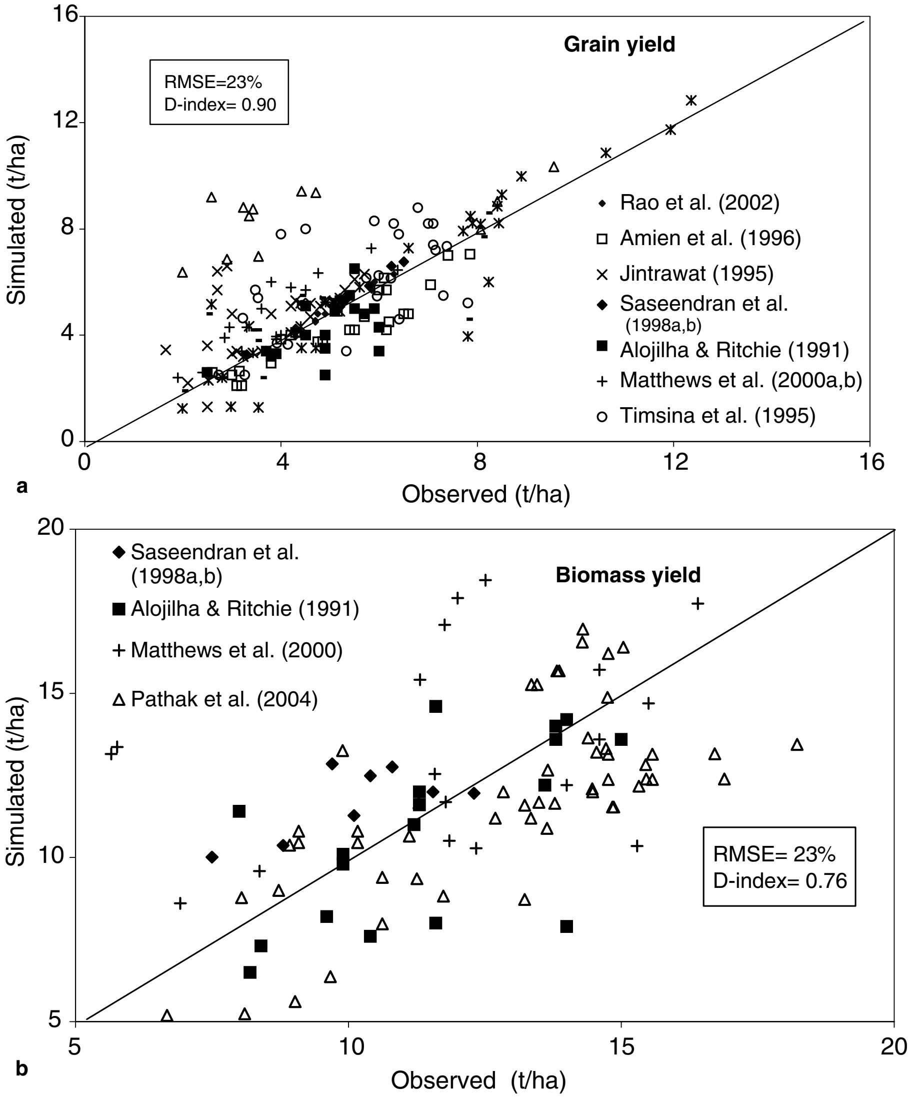


Fig. 2. Simulated and observed (a) grain and (b) biomass yields of rice across a range of experiments in Asia and Australia.

Sub-tropics and temperate. At temperate and sub-tropical locations in Japan (Seino, 1995), the model predicted days to physiological maturity of unspecified cultivars quite well, with normalised RMSE = 2% and *D*-index = 1.0. In sub-tropical northwest India (Timsina et al., 1995), the absolute RMSE for both anthesis and maturity was 6 d, but *D*-index was 0.72 for anthesis and 0.96 for maturity, indicating less satisfactory performance of the model. In sub-tropical northern Bangladesh (Timsina et al., 1998), however, both graphical display and statistical analyses suggested very good agreement for anthesis and maturity dates of BR11 and BR14, with normalised RMSE of 5% and 4%, and *D*-index of 0.98 and 0.96, respectively. In south India (Saseendran et al., 1998a,b), prediction of anthesis was less accurate with RMSEs of Jaya and IR8 of 4.4 and 4.2 d, and D-index = 0.65 and 0.84, respectively.

Comparison with other models. Mall and Aggarwal (2002) compared the performance of CERES-Rice and ORYZA1N at 11 locations from north to south India, including four in northwest India. Predicted values were within 15% of observations for both models, with similar RMSE of 4.5 d (CERES-Rice) and 4.8 d (ORYZA1N). In another study, CERES-Rice, ORYZA1, ORYZA-European, RICAM, SIMRIW and TRYM were compared in four diverse rice-growing environments in Asia using mean squared deviation (MSD), which ranged from 87 to 238 d for growth duration. CERES-Rice, RICAM, SIMRIW and TRYM had smaller MSDs than ORYZA, and all models simulated growth duration well in China, but not in the Philippines and Japan (M. Salam, personal communication).

All environments. Combined analysis of all data from the 4–6 studies shows reasonable predictive capability of CERES-Rice for the number of days to anthesis and maturity (Fig. 1, Table 4). There is slightly greater precision for prediction of maturity date than anthesis, as reflected in the smaller scatter about the 1:1 line, intercept closer to 0 and slope closer to 1.0, and slightly higher R^2 and lower normalised RMSE. Whether any greater precision in predicting maturity than anthesis date actually reflects the model capability or relative accuracy of determination of actual anthesis and maturity dates in the field is an interesting question worthy of further investigation. Few studies report how they actually determined these stages in the field.

4.2. Grain and biomass yields, and yield components

Tropics. Amien et al. (1996) reported that the CERES-Rice (ver. 3) under-predicted grain yield by 10–20% at all locations in Indonesia, except at Sukamandi in west Java, due to underprediction of grain weight ($R^2 = 0.83$; RMSE = 0.98 t/ha). In different locations in Thailand, the model predicted grain and biomass yields quite well despite poor prediction of the dates of phenological events (Tongyai, 1994) (Table 4). Jintrawat (1995), however, reported that both grains/m^2 and single grain weight as well as phenological events were predicted well, leading to good predictions of grain yield in Thailand. Using MERES (with CERES-Rice growth routine), Matthews et al. (2000b) reported fairly good prediction of grain and above-ground biomass yield (RMSE = 1.1 and 3.9 t/ha, respectively) at Los Baños, Philippines and Hangzhou, China, except for three treatments with mid-season drainage in the dry season at Los Banos.

Sub-tropics. In Kerala, India, Saseendran et al. (1998a,b) predicted grain and straw yields of Jaya and IR8 within 3% and 27% of measured yields (RMSE = 0.2 and 1.9 t/ha) with *D*-index of 0.99 and 0.56, respectively (Table 4). Rao et al. (2002) also reported good yield prediction in Kerala (RMSE = 0.2 t/ha, D-index = 0.99) for all transplanting dates for three cultivars in one year. However, in another year in which heavy rains, and incidences of gallmidge, brownspot and rice bug during flowering stage resulted in high sterility, the data were not used for model evaluation. In northern Bangladesh, simulated yields of BR14 and BR11 were either over or underestimated (RMSE = 1.2 t/ha; *D*-index = 0.94), with large underpredictions for 0 N (Timsina et al., 1998). Some of the discrepancy at high N rates was due to insect damage and lodging, which the model is unable to predict. Mahmood et al. (2003) also reported satisfactory performance of the model, with observed yields from 2.9 to 6.7 t/ha and simulated from 2.6 to 7.3 t/ha, and RMSE of 1.3 t/ha, for central and northern Bangladesh.

In northwest India, RMSE for grain yield was 1.7 t/ha and *D*-index was 0.79, indicating large discrepancy between simulated and observed data (Timsina et al., 1995), largely due to inaccurate prediction of dates of phenological events. Pathak et al. (2004) evaluated CSM-CERES-Rice ver. 4.0 using data from a range of water regimes (saturated to frequent intermittent wetting and drying to prolonged mid-season drying) and N (0 to recommended to supra-optimal dose) management treatments for three RW growing environments in northwest India. There was good agreement for grain yield (RMSE = 0.72 t/ha; *D*-index = 0.95) and reasonable agreement for dry matter yield (RMSE = 2.6 t/ha; *D*-index = 0.83) in well-fertilized (N) treatments, but generally poor agreement for the 0 N treatments.

Temperate. In Japan, simulated grain yields at three sites were within 0–5% of observed yields (Seino, 1995). In one site in southern Australia, simulated grain and total biomass yields were within 10% and 3%, respectively, of the observed yields in 1 year, but across several years and sites, there were large discrepancies, especially in cold years, due to the model's inability to simulate sterility induced by cold damage (Godwin et al., 1994; Meyer et al., 1994).

All environments. Combined analysis of 11 data sets for grain and 4 sets for biomass yield revealed that the data were more scattered around 1:1 line, intercept and slope closer to 0 and 1.0, respectively, absolute RMSE lower, and *D*-index higher for grain yield than biomass yield, indicating better accuracy and precision for former than the latter. Normalised RMSE for both traits was 23%, while absolute RMSE and *D*-index were 1.14 and 2.74 t/ha, and 0.90 and 0.76, for grain and biomass yield, respectively. Accuracy and precision of their prediction was, however, not as good as anthesis and maturity dates (Fig. 2; Table 4). However, these data sets also include crops affected by lodging and insect damage at high N rate, heavy rain during anthesis and grain filling, and cold damage. When these crops are removed from the analysis, the precision of the model predictions is increased with normalised RMSE = 14% and *D*-index = 0.96, and intercept and slope became closer to 0 and 1.0, respectively, for the grain yield.

Comparison with other models. Bachelet et al. (1993) compared observed grain yields with yields predicted by CERES-Rice and MACROS (Modules of Annual Crop Growth Simulator) (Penning de Vries et al., 1989) under various climates. Using default coefficients, predictions by both models were fairly reliable, although the fit was closer for MACROS than CERES. With derived coefficients, however, the reverse was the case. Kropff et al. (1994) reported that ORYZA1, CERES-Rice, SIMRIW, and TRYM overestimated yields in the wet season at IRRI and, with the exception of ORYZA1, predicted LAI inaccurately in the dry and wet season at IRRI, and at Kyoto, Japan and Yanco, Australia. Mall and Aggarwal (2002), however, concluded that both CERES-Rice and ORYZA1N predicted grain yields satisfactorily (within ±15%), especially for yields above 4 t/ha, with RMSE of 0.7 and 0.6 t/ha, respectively. Both models predicted grain number fairly accurately over the range 15,000–32,000 grains/m^2, with ORYZA1N performing better than CERES-Rice at lower, but the latter performing better at higher, yield levels. Both models were unable to adequately simulate growth and yield under N and water deficit stresses.

In another study with six models, all models closely predicted yields for early sowing at Moroika, Japan and Nan Chang, China, but not for late sowing at Nan Chang and for the dry season at IRRI. The largest deviations were for SIMRIW and ORYZA1, followed in order by TRYM, CERES-Rice, ORYZA-European, and RICAM. For harvest index, RICAM and ORYZA1 had the largest MSD for Nan Chang early and late seasons, while CERES-Rice had the largest MSD at Moroika. CERES-Rice also had the highest deviations for biomass yield (M. Salam, personal communication).

4.3. Other variables

Growth. Only a few published reports are available on the performance of CERES-Rice in predicting in-season biomass and leaf area. Mall and Aggarwal (2002) reported good agreement for LAI using both CERES-Rice and ORYZA1N, with slightly better agreement by CERES. Both models underpredicted peak LAI due to limited calibration and inaccurate initialisation of soil mineral N and soil water. Meyer et al. (1994), in Australia, also reported underestimation of peak LAI and panicle density by CERES-Rice (ver. 2.1), although it predicted the time of complete canopy closure (LAI $\sim$ 3) reasonably accurately. Salam et al. (2001) reported that CERES-Rice overestimated root and leaf dry matter and underestimated culm dry matter during the 30-d period of seedling growth, and overestimated the duration of the transplanting shock period. With the inclusion of new nursery growth and transplanting shock routines, simulations of seedling growth and the transplanting shock period improved greatly, but not the yield response to transplanting shock.

Nitrogen. With ver. 2.0, Godwin et al. (1990) and Buresh et al. (1991) reported predicted N loss through urea of 6–58 vs. observed loss of 5–49 kg/ha from eight irrigated lowland rice experiments at three sites in the Philippines, with treatments ranging from continuous flooding to alternate wetting and drying (AWD). In most cases, especially in continuously-flooded soil with low percolation rates (<0.2 cm/d), the model predicted negligible denitrification loss, consistent with field measurements from ^{15}N balance (Buresh and de Datta, 1990). In the AWD treatments, however, the predicted losses exceeded those from measurements due to overprediction of soil drying during periods of water deficit, insufficient lag time for the onset of nitrification following soil drying, and overestimation of the denitrification upon reflooding. Buresh et al. (1991) also reported that CERES-Rice was sensitive to above- and below-ground residue inputs, initial soil ammonium–N concentration, and N mineralisation rate. Pathak et al. (2004) evaluated the model (ver. 4.0) for soil mineral N and loss processes from rice fields under RW systems for Delhi and Punjab in northwest India. Simulation of soil mineral N in the surface layer (0–15 cm) was generally poor, but the model predicted N leaching losses similar to those in field experiments reported in the literature. It also predicted higher denitrification and lower ammonia volatilisation losses, for urea applied at 120 kg N/ha, compared to observed losses.

Using ver. 2.0, Godwin et al. (1990) and Buresh et al. (1991) also reported predicted crop N uptake of 40–145 vs. observed uptake of 35–150 kg/ha from the above same experiments and treatments in the Philippines. With ver. 3.5, Timsina et al. (1998) reported that both observed and simulated total crop N uptake by two rice cultivars ranged from 48–175 kg/ha, with absolute RMSE of 17 kg/ha in northern Bangladesh, while with ver. 4.0, Pathak et al. (2004) reported good agreement, especially for grain N uptake, with RMSE of 21 kg/ha and *D*-index of 0.86, in a RW system in Delhi.

In southern Australia too, with ver. 2.1, predicted total biomass and N uptake were within 3% of the observed values, but N partitioning between straw and grain was poor (Meyer et al., 1994).

Methane emission. With MERES, Matthews et al. (2000a,b) reported good agreement for seasonal patterns and quantities of methane emission from rice straw and green manure treatments for a dry season crop at IRRI (simulated 0–450 vs. observed 0–480 kg CH4-C/ha) and for seasonal methane emissions at Maligaya (simulated 150–450 vs. observed 125–570 kg), Philippines, and Hangzhou, China (simulated 30–200 vs. observed 70–360 kg), though with rice straw at IRRI the model slightly overpredicted the plume of methane at the second drainage just before harvest. They concluded that rice yield predicted by MERES was sensitive to root death coefficient, specific root exudation rate, root transmissivity, green manure addition, initial size of oxidised alternative electron acceptor pool, seasonal temperature, crop duration, floodwater depth, drainage period, type of organic amendments, phosphogypsum, and percolation rate. In a recent study, however, MERES greatly overestimated the emissions from a rice field in a RW system in Delhi (Pathak et al., 2004).

Evapotranspiration and deep drainage. Using ver. 2.1, Meyer et al. (1994) reported that, in temperate southern NSW, the pattern of observed ET matched predicted ET within 2% over the whole season. Using ver. 4.0, Pathak et al. (2004) also reported consistency between simulated ET and deep drainage from rice fields in Delhi and Punjab and typical values reported in the literature.

Cold damage. Cold damage due to low temperature during the early pollen microspore (EPM) stage is a common occurrence in Australian rice fields, with severest damage with high N fertility and shallow water (Williams and Angus, 1994; Gunawardena et al., 2003). Using ver. 2.1, Meyer et al. (1994) reported large discrepancies between observed and simulated yields, with most discrepancies when low temperatures occurred during EPM. Godwin et al. (1994) developed a chilling injury routine (ver. 2.1C) and obtained good agreement between simulated and observed yields for sprinkler irrigation treatments affected by cold damage. That version did not take into account the effects of floodwater temperature and assimilate availability on pollen cell survival throughout the chilling event (Doug Godwin, personal communication). Moreover, that version was not tested further against other data sets. Consequently, Timsina et al. (2004) developed new floodwater temperature and chilling injury routines and incorporated them into ver. 4.0 using the most recent understanding of the mechanisms of chilling injury in rice. The new version (ver. 4.0C) improved the ability to simulate rice response to low temperature compared with ver. 2.1C.

5. Evaluation of CERES-Wheat

CERES-Wheat has been evaluated in many places in north and south America, Europe, and Africa, but only in few parts of Asia and Australia. As for CERES-Rice, most of those studies provided only qualitative or graphical analysis, without calculating any statistical parameters.

5.1. Phenology

Prediction of wheat phenology was evaluated in three studies in tropical and sub-tropical RW regions of IGP in south Asia, one in temperate and sub-tropical region in Japan, and one in southeast temperate Australia (Fig. 3; Table 5). For the IGP, Timsina et al. (1995) reported that the model overestimated the time to anthesis (RMSE = 8.6 d; *D*-index = 0.87) and maturity (RMSE = 8.7 d; *D*-index = 0.99) of RR21 and HD2009 at Pantnagar, while Hundal and Kaur (1997) reported much better agreement (RMSE = 4.0 and 3.8 d; *D*-index = 0.84 and 0.93, respectively) for HD2329 at Ludhiana, both in NW India. In northern Bangladesh (Timsina et al., 1998) also, prediction of anthesis of Kanchan and Sowgat (RMSE = 2.0 d; *D*-index = 0.90) was good, while the date of maturity was generally overestimated (RMSE = 3.5 d; *D*-index = 0.52). Seino (1995) reported that CERES-Wheat predicted the days from sowing to physiological maturity within 1–2.5% of observed in northern (temperate) and central (sub-tropical) Japan. In southern NSW, Humphreys et al. (2005) reported that ver. 3.5 predicted anthesis and maturity of Janz and Yecora quite well in 2 out of 3 experiments. Taken all data sets together, predictions for both anthesis and maturity were quite similar, though there was a tendency for better prediction for maturity as indicated by slightly higher R^2 and *D*-index, lower normalised RMSE, and better spread of data around 1:1 line (Fig. 3; Table 5).

5.2. Growth and grain yield

Evaluation of CERES-Wheat and Wheat-W in a warm tropical environment at IRRI suggested that both models predict grain yield poorly due to high temperatures during the entire crop cycle, especially during emergence and grain-filling period (Timsina et al., 1995). Evaluation of CERES-Wheat for growth and grain yield of wheat in the sub-tropical environment of northwest India, however, showed

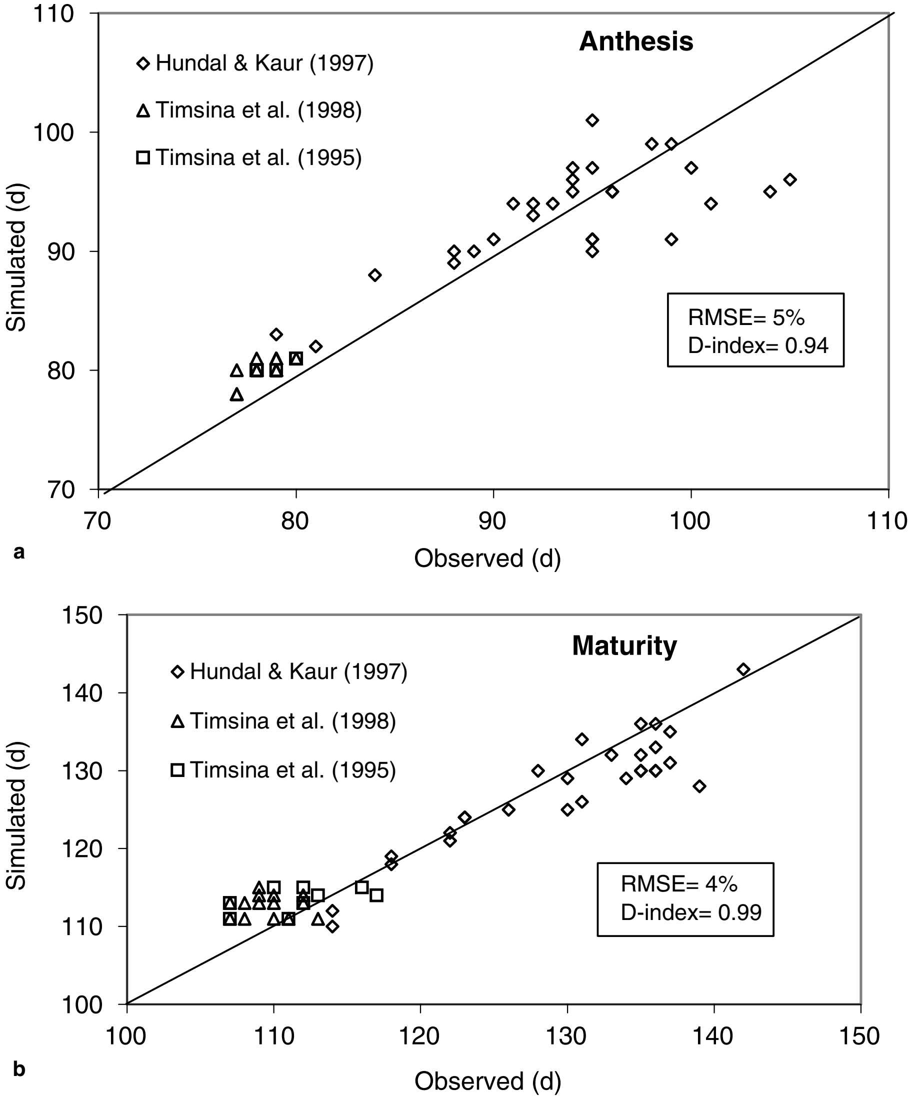


Fig. 3. Simulated and observed days to (a) anthesis and (b) maturity of wheat across a range of experiments in Asia.

reasonable predictive ability (Timsina et al., 1995; Hundal and Kaur, 1997). Hundal and Kaur (1997) concluded that the model can be used for yield prediction for the central irrigated plains of the Indian Punjab, but that there is also a need for closer examination of the quantitative relationships governing the partitioning of photosynthates into biomass and grain yield. Similarly, in sub-tropical northern Bangladesh, yield predictions for eight treatment combinations of N, water, and sowing

Table 5

Evaluation results for CERES-Wheat predictions of phenology, and grain and biomass yields for various data sets (simulated and observed data are from various studies reported under ‘data sets’)

Data sets	N	X_{obs} (SD)	X_{sim} (SD)	α	β	R^2	Absolute RMSE	RMSEs	RMSEu	RMSE (%)	D-index
Grain yield (t/ha)											
Seino (1995)	3	3.4 (0.7)	3.2 (0.6)	0.27	0.97	0.87	0.27	0.19	0.19	8	0.94
Timsina et al. (1995)	26	3.4 (1.3)	3.5 (1.1)	−0.26	1.10	0.85	0.51	0.27	0.44	15	0.96
Timsina et al. (1998)	24	3.1 (1.5)	3.3 (1.6)	0.14	0.90	0.97	0.35	0.21	0.28	11	0.99
Hundal and Kaur (1997)	29	4.2 (0.6)	3.9 (0.7)	1.50	0.70	0.63	0.53	0.31	0.42	12	0.84
Nain et al. (2002)	30	2.8 (0.9)	2.8 (0.9)	0.12	0.96	0.96	0.19	0.08	0.19	7	0.99
Godwin et al. (2002)	6	4.6 (0.9)	4.5 (0.9)	0.03	1.01	0.99	0.11	0.08	0.07	2	1.00
Heng et al. (2000)	19	4.6 (1.8)	4.8 (1.6)	−0.09	0.98	0.79	0.79	0.37	0.70	17	0.94
All data sets	137	3.6 (1.4)	3.6 (1.3)	0.15	0.96	0.88	0.48	0.10	0.22	13	0.97
Biomass yield (t/ha)											
Hundal and Kaur (1997)	29	10.6 (1.3)	11.6 (1.4)	3.93	0.57	0.40	1.54	1.11	1.06	15	0.66
Heng et al. (2000)	19	11.6 (3.9)	10.7 (3.1)	−0.48	1.14	0.80	2.02	1.50	1.34	17	0.91
All data sets	48	11.0 (2.7)	11.3 (2.2)	0.86	0.90	0.58	1.74	0.98	1.44	16	0.86
Anthesis (d)											
Timsina et al. (1995)	10	82.2 (12.0)	88.5 (12.6)	8.64	0.83	0.77	8.56	6.35	5.74	10	0.87
Timsina et al. (1998)	24	77.8 (3.4)	79.6 (2.7)	−18.8	1.21	0.96	2.01	1.93	0.55	3	0.90
Hundal and Kaur (1997)	29	93.7 (6.0)	93.2 (4.4)	−1.80	1.02	0.56	3.96	2.74	2.86	4	0.84
All data sets	64	85.7 (9.8)	87.2 (8.6)	−4.27	1.03	0.81	4.49	2.55	3.69	5	0.94
Maturity (d)											
Timsina et al. (1995)	13	141.8 (50.2)	141.1 (45.6)	−1.68	1.02	0.97	8.66	2.40	8.32	6	0.99
Timsina et al. (1998)	24	110.5 (2.7)	113.0 (1.5)	35.92	0.66	0.13	3.54	3.27	1.35	3	0.52
Hundal and Kaur (1997)	29	129.7 (7.7)	127.7 (7.1)	5.05	0.98	0.82	3.84	2.39	3.06	3	0.93
All data sets	66	125.1 (25.2)	125.0 (23.8)	−4.79	1.04	0.96	5.08	1.91	4.71	4	0.99

N, number of data pairs; X_{obs}, mean of observed values; X_{sim}, mean of simulated values; SD, standard deviation; α, slope of linear relation between simulated and observed values; β, intercept of linear relation between simulated and observed values; R^2, adjusted linear correlation coefficient between simulated and observed values; RMSE, root mean square error.

dates were quite good (RMSE = 0.4 t/ha; *D*-index = 0.99) (Timsina et al., 1998). In another study in India, Bangladesh and China (Heng et al., 2000), there were reasonably good predictions of grain and biomass yields as judged by several statistical parameters (low intercepts: slope = 0.98, $R^2 = 0.79$, normalised RMSE = 17%, and *D*-index = 0.94). All these studies in the sub-tropical environments were carried out in the IGP where RW systems predominate.

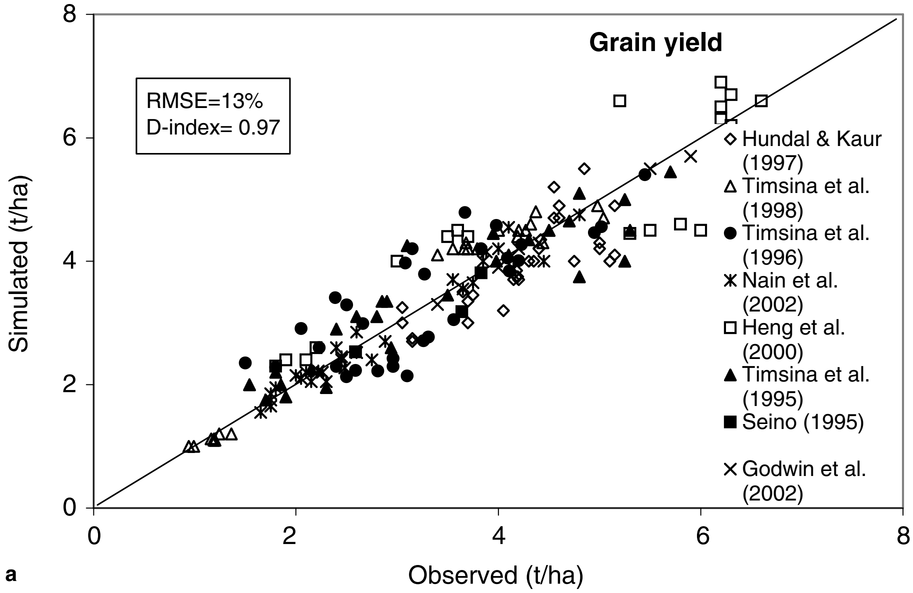
Seino (1995) reported predicted grain yields within 1–2% of observed (RMSE = 0.27 t/ha; *D*-index = 0.94) in several locations in Japan. Otter-Nacke et al. (1986) validated an early version of the model for a range of cultivars in many countries, including temperate rainfed environments across Australia. They reported that the model explained about 60% of the variation in grain yield, but the evaluation for time course of LAI and above ground biomass, N uptake, and root weight responses to N rate showed variable results. Humphreys et al. (2005) evaluated ver. 3.0 with three cultivars in field and lysimeter experiments in the rice-growing region of southern Australia where the model performed well for grain yield. Taken all data sets together, accuracy and precision of predictions of grain yield was better than biomass yield as reflected by better spread of data around 1:1 line, intercept and slope closer to 0 and 1.0, respectively, lower RMSE, and higher R^2 and *D*-index for grain than biomass yield (Fig. 4; Table 5).

5.3. Other variables

Humphreys et al. (2005) reported that the ver. 3.0 predicted well the time course of biomass production and LAI, volumetric soil water content and root length density, and daily ET of three cultivars in rice-growing areas of southern Australia. Recently, with CSM-CERES-Wheat (ver. 4.0), Timsina et al. (2005) reported simulated deep drainage losses beyond 0.9 m depth of 0–165 mm and ET of 338–662 mm for a range of sowing dates and irrigation methods. Deep drainage was higher for flood irrigation compared with sprinkler irrigation and rainfed but ET was higher for sprinkler. Both deep drainage and ET were within the ranges of observed values reported by Humphreys et al. (2003a,b). Comprehensive evaluations of this nature are important to demonstrate robustness of the model processes to simulate the effect of water stress on growth and yield for sub-tropical and temperate environments.

5.4. Model comparisons

Nain et al. (2002) used combined analyses from the technology-trend model and CERES-Wheat for several locations in India, and reported a mean grain yield of 2.82 t/ha, RMSE of 0.2 t/ha, and *D*-index of 0.99, indicating good performance of the model under "high" input conditions. Godwin et al. (2002) compared the predictions of CERES-Wheat (ver. 3.5) and SWAGMAN-Destiny against data from irrigated wheat in southern Australia, and concluded that both models predict in-season LAI, total biomass, root length density, and soil water content, and final grain yield quite satisfactorily in temperate Australia. In con-



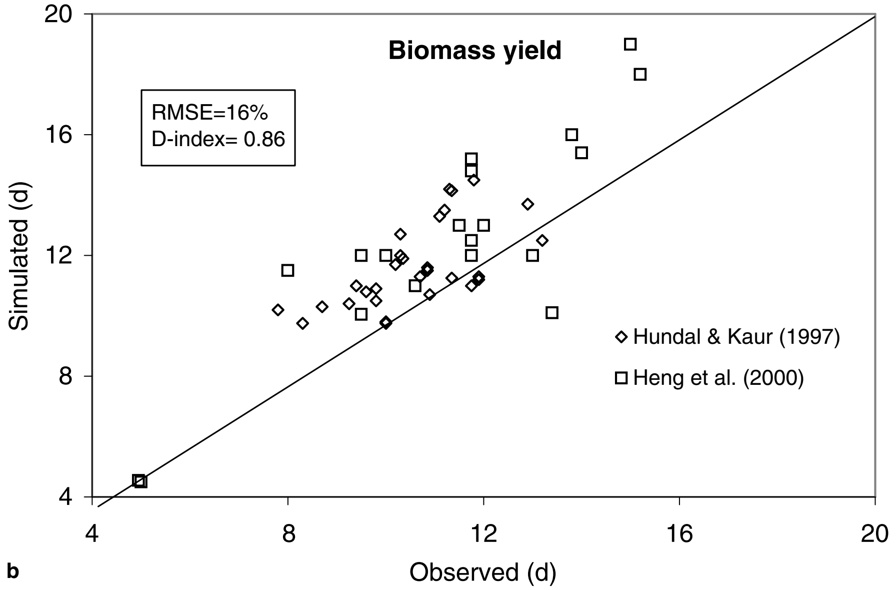


Fig. 4. Simulated and observed grain and biomass yields of wheat across a range of experiments in Asia and Australia.

trast, after comparing the accuracy of eight simulation models for several low yield potential temperate environments, including Australia, O'Leary (2000) concluded that yield underestimation under stressed conditions was a common problem of all models, and suggested the need of better models, especially for very low yield levels (<1 t/ha).

6. Evaluation of the CERES RW sequence model

Timsina et al. (1996), the only reported model validation study for RW system, used 20 years' experimental and historical weather data from Pantnagar, to validate the DSSAT sequence model (ver. 3.5). In the experiment, planting dates, cultivars, and initial conditions varied across years, so the model was manually reset prior to each season using observed initial soil mineral N and soil water, planting dates, and cultivars. The model satisfactorily predicted the seasonal fluctuations in rice and wheat yields, and the long-term trends for each crop, over 20 years, though the predictions were more accurate with N fertilizer application than with no added N. The model also greatly overestimated the long-term decline in soil organic carbon (SOC) and nitrogen (SON). The discrepancies in yield were partly due to inadequate model input data, especially initial soil mineral N and soil water prior to each crop, but poor predictions of SOC and SON were also due to the presence of only one pool of SOM in that version of DSSAT (ver. 3.5). DSSAT ver. 4.0 includes the option of using a CENTURY-based SOM routine which has three SOM pools (passive, slow, and active microbial) (Gijsman et al., 2002). However, this is yet to be tested for RW system, and the CENTURY module does not simulate flooded soils.

7. General discussion and conclusions

There are many reports of the evaluation of CERES-Rice in the RW areas of tropical and sub-tropical Asia, but only a few for CERES-Wheat. Most evaluations have been limited to dates of anthesis and maturity, and grain and biomass yields. Both models generally predicted anthesis and maturity within a few days of observed dates. CERES-Rice predicted grain and biomass yields quite variably, across a range of tropical to temperate locations, with generally good predictions under optimal N and water conditions (predicted values usually within 10–15% of observations.), but not under stress (N deficit, water deficit, low temperature). CERES-Wheat, however, generally predicted the grain and biomass yields quite satisfactorily, except at very low yield levels (<1 t/ha) or in high temperature environments.

The variable performance of the models is probably due to a combination of deficiencies in model inputs, experimental observations, inclusion of non-modelled factors (such as disease, lodging, pests, storms) in model validation, and insufficient capture of model processes. Possible input deficiencies include insufficient data for derivation of robust genetic coefficients, lack of data on initial soil mineral N and water, and lack of proper soil characterisation (especially hydraulic properties). Few of the above reports of model evaluation state how these inputs were derived, or how genetic coefficients were determined.

The variable performance of the models, and in particular of CERES-Rice, highlights the importance of proper calibration and evaluation in the environment of interest before applying them to evaluate management options. This is especially important in the absence of reports on evaluation of model processes as reflected in the models' relative inabilities to predict a range of crop, soil and water parameters.

Better evidence of the ability of the models to simulate a range of important parameters other than yield, such as time course of biomass production, leaf area development, N uptake, soil water and mineral N dynamics, and components of the water and N balances, is also highly desirable to demonstrate the robustness of model processes and to increase the confidence in the use of model. There are no reports of rigorous evaluations of both models for variables related to increasing resource-use efficiency such as water and N management. The results of the few studies where some of these components have been determined are, however, generally encouraging.

Further model testing and improvements are thus needed under water and N limiting conditions before they can be used with confidence to explore management options to increase resource use efficiency, such as stretching irrigation intervals and placement of N fertilizers, for RW systems in both Asia and Australia. This is particularly important because of the strong dependence of N transport and transformations on the hydrology of the system, the risk of soil water deficit stress during non-ponded periods, and the problem of cold damage during early pollen microspore in the Australian environment. There is also a need to improve the ability of CERES-Wheat to simulate the effect of high temperatures, especially during grain filling, to examine trade-offs between yield and water use requirement by changing sowing date in both the tropical and sub-tropical RW regions of Asia and temperate Australia. In Australia, the model also needs to be able to examine the trade-off between sowing late to avoid frost damage vs. shortened grain filling during warmer weather.

Only one study evaluated the performance of the DSSAT RW sequence model, but with initial mineral N and soil water content based on observations inputted for each crop. This work demonstrated that the model predicts the long-term grain yield of component crops fairly satisfactorily, but not the long-term changes in SOC and SON, suggesting that the sequence model needs further evaluation against data from a range of locations and management. The results of long-term experiments could be useful for this purpose, however, availability of adequate and good quality data availability is always a problem.

In conclusion, while it seems that CERES-Rice and CERES-Wheat have performed reasonably well in RW regions of Asia and Australia, evaluations and applications addressing resource use efficiency and sustainability issues are lacking. Better data from field experiments designed to address these issues, and further model evaluations, improvements and applications, are needed to address the issues of yield stagnation or decline and increasing yield gaps, and finally to contribute to solving the resource and food security problems in RW systems of Asia and to regional economies of Australia.

Acknowledgements

We thank F.W.T. Penning de Vries (IWMI) and James W. Jones (University of Florida) for reviewing an earlier draft, and Ken Boote (University of Florida) for reviewing a later draft, of the manuscript. We are also grateful to Doug Godwin (Consultant to CSIRO Land and Water) for various discussions on the global performance of CERES models. We are grateful to the Australian Centre for International Agricultural Research for funding support for our research on RW system.

References

Alociljha, E.C., Ritchie, J.T., 1991. A model for the phenology of rice. In: Hodges, T. (Ed.), Predicting Crop Phenology. CRC Press, Boca Raton, FL, USA, pp. 181–189.

Amien, I., Rejekingrum, P., Pramudia, A., Susanti, E., 1996. Effects of interannual climate variability and climate change on rice yield in Java, Indonesia. Water Air Soil Pollut. 92, 29–39.

Andales, A.A., Batchelor, W.D., Anderson, C.E., Farnham, D.E., Whigham, D.K., 2000. Incorporating tillage effects into a soybean model. Agri. Syst. 66, 69–98.

Bachelet, D., Van Sickle, J., Gay, C.A., 1993. The impacts of climate change on rice yield: evaluation of the efficacy of different modelling approaches. In: Penning de Vries, F.W.T. et al. (Eds.), Systems Approaches for Agricultural Development. Kluwer Academic Publishers, The Netherlands, pp. 145–174.

Boonjung, H., 2000. Climate Variability and Rice Production in Rainfed Rice Area in Northeast Thailand: Risk Analysis and Management Applications, pp. 202–205. Available from: <http://www.earthscape.org/rmain/rsites/clch.html>.

Bouman, B.A.M., Penning de Vries, F.W.T., Wan Harun, W.S., Timsina, J., Hartkamp, A.D., 1994. Optimising RW rotations under water-limited conditions. Paper presented at the symposium "Scenario studies for the rural environment", 12–15 September 1994, Wageningen, Netherlands, and the "Third Congress of the European Society for Agronomy. 18–22 September 1994, Podowa, Italy, 4 p.

Buresh, R.J., de Datta, S.K., 1990. Denitrification losses from puddled rice soils in the tropics. Biol. Fert. Soils 9, 1–13.

Buresh, R.J., Singh, U., Godwin, D.C., Ritchie, J.T., de Datta, S.K., 1991. Simulating soil nitrogen transformations with CERES-Rice. Agrotechnol. Transf. 13, 7–10.

Castrignano, A., Katerji, N., Karam, F., Mastrorilli, M., Hamdy, A., 1998. A modified version of CERES-Maize model for predicting crop response to salinity stress. Ecol. Model. 111, 107–120.

Dawe, D., Frolking, S., Li, C., 2004. Trends in rice–wheat area in China. Field Crops Res. 87, 89–95.

Gauch, H.G., Hwang, J.T.G., Fick, G.W., 2003. Model evaluation by comparison of model-based predictions and measured values. Agron. J. 95, 1442–1446.

Gijsman, A.J., Hoogenboom, G., Parton, W.J., Kerridge, P.C., 2002. Modifying DSSAT for low-input agricultural systems, using a SOM/residue module from CENTURY. Agron. J. 94, 462–474.

Godwin, D., Jones, C.A., 1991. Nitrogen dynamics in soil-crop systems. In: Hanks, R.J., Ritchie, J.T. (Eds.), Modeling Plant and Soil systems, Agronomy Monograph 31. American Society of Agronomy, Madison, WI, USA, pp. 287–321.

Godwin, D., Singh, U., 1991. Modelling nitrogen dynamics in rice cropping systems. In: Deturck, P., Ponnamperuma, F.N. (Eds.), Rice Production on Acid Soils of the Tropics. Institute of Fundamental Studies, Kandy, Sri Lanka, pp. 287–294.

Godwin, D., Singh, U., 1998. Nitrogen balance and crop response to nitrogen in upland and lowland cropping systems. In: Tsuji, G.Y., Hoogenboom, G., Thornton, P.K. (Eds.), Understanding Options for Agricultural Production. Kluwer Academic Publishers, Dordrecht, The Netherlands, pp. 55–78.

Godwin, D.C., Singh, U., Buresh, R.J., De Datta, S.K., 1990. Modelling N dynamics in relation to rice growth and yield. In: Transactions of the 14th Int. Cong. Soil Sci., Kyoto, Japan. Int. Soc. Soil Sci., Japan, pp. 320–325.

Godwin, D.C., Meyer, W.S., Singh, U., 1994. Simulation of the effect of chilling injury and nitrogen supply on floret fertility and yield in rice. Aust. J. Exp. Agri. 34, 921–926.

Godwin, D., Humphreys, E., Smith, D.J., Timsina, J., Xevi, E., Meyer, W.S., 2002. Application of SWAGMAN destiny to rice–wheat cropping systems. In: Humphreys, E., Timsina, J. (Eds.), Modelling Irrigated Cropping Systems, with Special Attention to Rice–Wheat Sequences and Raised Bed Planting. Proceedings of a Workshop, CSIRO Land and Water, Griffith, NSW, Australia, 25–28 February 2002. CSIRO Land and Water Technical Report 25/02, pp. 69–82.

Gunawardena, T.A., Fukai, S., Blamey, F.P.C., 2003. Low temperature induced spikelet sterility in rice. I. Nitrogen fertilization and sensitive reproductive period. Aust. J. Agri. Res. 54, 937–946.

Heng, L.K., Baethgen, W.E., Moutoonnet, P., 2000. The collection of a minimum dataset and the application of DSSAT for optimizing wheat yield in irrigated cropping systems, pp. 7–17. In: Optimizing Nitrogen Fertilizer Application to Irrigated Wheat, IAEA TECDOC-1164, 2000, p. 245.

Hodges, T., Ritchie, J.T., 1991. The CERES-Wheat phenology model. In: Hodges, T. (Ed.), Predicting Crop Phenology. CRC Press, Boca Raton, FL, USA, pp. 133–143.

Hoogenboom, G., Georgiev, G., Singh, U., 1997. Performance of the CERES-Rice model for three locations in the Asian Region. Report of the Workshop "Climate Variability, Agricultural Productivity and Food Security in the Asian Monsoon Region". A Joint START/WCRP/ IGBP-GCTE Workshop, 19–22 February 1997, Bogor, Indonesia. START Report 1997/2, pp. 35–42.

Hoogenboom, G., Jones, J.W., Porter, C.H., Wilkens, P.W., Boote, K.J., Batchelor, W.D., Hunt, L.A., Tsuji, G.Y., 2004. DSSAT 4.0. vol. 1, Overview. ICASA, University of Hawaii, Honolulu, USA.

Humphreys, E., Edraki, M., Bethune, M., 2003. Deep drainage and crop water use for irrigated annual crops and pastures in Australia – a review of determinations in fields and lysimeters. CSIRO Land and water Technical Report 14/03, 33 pp.

Humphreys, E., Smith, D.J., White, R.J.G., Xevi, E., Godwin, D.C., 2003. MaizeMan – a decision support system for water and nitrogen management. In: Fifth Australian Maize Conference Proceedings, 18th–20th February, 2003, Toowoomba, Queensland, Australia, pp. 127–134.

Humphreys, E., Meisner, C., Gupta, R., Timsina, J., Beecher, H.G., Yong Lu, T., Yadvinder-Singh, Gill, M.A., Masih, I., Jia Guo, Z., Thompson, J.A., 2004. Water saving in rice–wheat systems. Plant Production Science 8, 242–258.

Humphreys, E., Smith, D.J., Fattore, A., Bhuiyan, A.M., Fawcett, B.J., 2005. Quantifying and maximising the benefits of crops after rice. CRC for Sustainable Rice Production Final Report, Project 1205, Yanco, NSW, Australia.

Hundal, S.S., Kaur, P., 1997. Application of the CERES-Wheat model to yield predictions in the irrigated plains of the Indian Punjab. J. Agric. Sci., Cambridge 129, 13–18.

Hunt, L.A., Boote, K.J., 1998. Data for model operation, calibration, and validation. In: Tsuji, G.Y., Hoogenboom, G., Thornton, P.K. (Eds.), Understanding Options for Agricultural Production. Systems Approaches for Sustainable Agricultural Development. Kluwer Academic Publishers, Great Britain, pp. 9–39.

Jintrawat, A., 1995. A decision support system for rapid assessment of lowland rice-based cropping alternatives in Thailand. Agri. Syst. 47, 245–258.

Jones, C.A., Kiniry, J.R. (Eds.), 1986. CERES-Maize: A Simulation Model of Maize Growth and Development. Texas A&M University Press, College Station, TX, p. 194.

Jones, J.W., Hoogenboom, G., Porter, C.H., Boote, K.J., Batchelor, W.D., Hunt, L.A., Wilkens, P.W., Singh, U., Gijsman, A.J., Ritchie, J.T., 2003. The DSSAT cropping system model. Eur. J. Agron. 18, 235–265.

Kobayashi, K., Salam, M.U., 2000. Comparing simulated and measured values using mean squared deviation and its components. Agron. J. 92, 345–352.

Kropff, M.J., Williams, R.L., Horie, T., Angus, J.F., Singh, U., Centeno, H.G., Cassman, K.G., 1994. Predicting yield potential of rice in different environments. In: Humphreys et al. (Eds.), Temperate Rice

– Achievements and Potential. Proceedings of Temperate Rice Conference, Yanco, NSW, Australia, pp. 657–663.

Ladha, J.K., Pathak, H., Tirol-Padre, A., Dawe, D., Gupta, R.K., 2003. Productivity trends in intensive rice–wheat cropping systems in Asia. In: Ladha, J.K., Hill, J.E., Duxbury, J.M., Gupta, R.K., Buresh, R.J. (Eds.), Improving the Productivity and Sustainability of Rice–Wheat Systems: Issues and Impacts'. ASA Special Publication 65 (ASA Inc., CSSA Inc., SSSA Inc., Madison, USA), pp. 45–76.

Mahmood, Meo, M., Legates, D.R., Morrissey, M.L.. The CERES-Rice model-based estimates of potential monsoon season rainfed rice productivity in Bangladesh. The Professional Geographer 55, 259–273.

Mall, R.K., Aggarwal, P.K., 2002. Climate change and rice yields in diverse agro-environments of India. I. Evaluation of impact assessment models. Climate Change 52, 315–330.

Matthews, R.B., Wassmann, R., Arah, J., 2000a. Using a crop/soil simulation model and GIS techniques to assess methane emissions from rice fields in Asia. I. Model development. Nutr. Cycl. Agroecosyst. 58, 141–159.

Matthews, R.B., Wassmann, R., Buendia, L.V., Knox, J.W., 2000b. Using a crop/soil simulation model and GIS techniques to assess methane emissions from rice fields in Asia. II. Model validation and sensitivity analysis. Nutr. Cycl. Agroecosyst. 58, 161–177.

Meyer, W.S., White, R.J.G., Smith, D.J., Baer, B.D., 1994. Monitoring a rice crop to validate the CERES-Rice model. CSIRO Technical Memorandum 94/13. September 1994, 26 p.

Nain, A.S., Dadhwal, V.K., Singh, T.P., 2002. Real time wheat yield assessment using technology trend and crop simulation model with minimal data set. Curr. Sci. 82, 1255–1258.

O'Leary, G.J., 2000. Can contemporary wheat models simulate grain yield accurately in low-potential environments? In: White, J.W., Grace, P.R. (Eds.), Modelling Extremes of Wheat and maize Crop Performance in the Tropics. Proceedings of a Workshop, CIMMYT, El Batan, Mexico, 19–22 April 1999. NRG-GIS Series 00-01. Mexico, DF: CIMMYT, pp. 5–10.

Otter-Nacke, S., Godwin, D., Ritchie, J.T., 1986. Testing and Validating the CERES-Wheat Model in Diverse Environments. IFDC, Muscle Shoals, AL, USA, p. 147.

Paroda, R.S., Woodhead, T., Singh, R.B., 1994. Sustainability of RW Production Systems in Asia. RAPA Pub. 1994/11. FAO, Bangkok, Thailand, 209 p.

Pathak, H., Ladha, J.K., Aggarwal, P.K., Peng, S., Das, S., Yadvinder-Singh, Bijay-Singh, Kamra, S.K., Mishra, B., Sastri, A.S.R.A.S., Aggarwal, H.P., Das, D.K., Gupta, R.K., 2003. Trends of climatic potential and on-farm yields of rice and wheat in the Indo-Gangetic plains. Field Crops Res. 80, 223–234.

Pathak, H., Timsina, J., Humphreys, E., Godwin, D.C., Bijay-Singh, Shukla, A.K., Singh, U., Matthews, R.B., 2004. Simulation of rice crop performance and water and N dynamics, and methane emissions for rice in northwest India using CERES-Rice model. CSIRO Land and Water Technical Report 23/04. CSIRO Land and Water, Griffith, NSW 2680, Australia, 111 p. Available from: <www.clw.csiro.au/publications/technical2004>.

Penning de Vries, F.W.T., Jansen, D.M., ten Berge, H.F.M., Bakema, A., 1989. Simulation of Ecophysiological Processes of Growth in Several Annual Crops. PUDOC, Wageningen, The Netherlands, p. 271.

Porter, C.H., Hoogenboom, G., Batchelor, W.D., Jones, J.W., Gijsman, A.J., 2003. DSSAT v4.0 Crop Models: Overview of changes relative to DSSAT v3.5. University of Florida, Gainesville, USA.

Rao, G.S.L.H.V., Sebastian, S., Subash, N., 2002. Crop growth simulation models of rice under humid climates. Available from: <http:///www.commonorthwestealthknowledge.northeastt/MetCD/Chapter7/C7P06.htm>.

Ritchie, J.T., 1998. Soil water balance and plant stress. In: Tsuji, G.Y., Hoogenboom, G., Thornton, P.K. (Eds.), Understanding Options for Agricultural Production. Kluwer Academic Publishers, Dordrecht, The Netherlands, pp. 41–54.

Ritchie, J.T., NeSmith, D.S., 1991. Temperature and crop development. In: Hanks, R.J., Ritchie, J.T. (Eds.), Modeling Plant and Soil Systems, Agronomy Monograph 31. American Society of Agronomy, Madison, WI, USA, pp. 5–29.

Ritchie, J.T., Singh, U., Godwin, D., Bowen, W.T., 1998. Cereal growth, development, and yield. In: Tsuji, G.Y., Hoogenboom, G., Thornton, P.K. (Eds.), Understanding Options for Agricultural Production. Kluwer Academic Publishers, Dordrecht, The Netherlands, pp. 79–98.

Salam, M.U., Jones, J.W., Kobayashi, K., 2001. Predicting nursery growth and transplanting shock in rice. Exp. Agri. 37, 65–81.

Saseendran, S.A., Hubbard, K.G., Singh, K.K., Mendiratta, N., Rathore, L.S., Singh, S.V., 1998a. Optimum transplanting dates for rice in Kerala, India, determined using both CERES v3.0 and ClimProb. Agron. J. 90, 185–190.

Saseendran, S.A., Singh, K.K., Rathore, L.S., Rao, G.S.L.H.V.P., Mendiratta, N., Lakshmi Narayan, K., Singh, S.V., 1998b. Evaluation of the CERES-Rice v.3.0 model for the climate conditions of the state of Kerala, India. Meterol. Appl. 5, 385–392.

Seino, H., 1995. Implications of climate change for crop production in Japan. In: Climate Change and Agriculture: Analysis of Potential International Impacts. ASA Special Publication Number 59. ASA, Madison, WI, pp. 293–306.

Singh, U., 1994. Nitrogen management strategies for lowland rice cropping systems. In: Proceedings of the International Conference on Fertilizer Usage in the Tropics (FERTROP). Malaysian Society of Soil Science, Kuala Lumpur, Malaysia, pp. 110–130.

Singh, U., Wilkens, P.W., Chude, V., Oikeh, S., 1999. Predicting the effect of nitrogen deficiency on crop growth duration and yield. In: Proceedings of the Fourth International Conference on Precision Agriculture. ASA-CSSA-SSSA, Madison, WI, USA, pp. 1379–1393.

Thornton, P.K., Hansen, J.W., 1996. A note on regressing real-world data on model output. Agri. Syst., 411–414.

Timsina, J., Connor, D.J., 2001. Productivity and management of RW systems: issues and challenges. Field Crops Res. 69, 93–132.

Timsina, J., Bouman, B.A.M., Penning de Vries, F.W.T., Van Kraalingen, D.W.G., Wan Sulaiman Wan Harun, 1994. Modelling tropical RW systems. In: Proceedings of SARP "Application Workshop" on the Application Program 'Agroecological Zonation and Characterisation' and 'Crop rotation optimisation', IRRI, Philippines, pp. 58–70.

Timsina, J., Singh, U., Singh, Y., Lansigan, F.P., 1995. Addressing sustainability of RW systems: testing and applications of CERES and SUCROS models. In: Proceedings of the International Rice Research Conference, 13–17 February 1995. IRRI, Los Banos, Philippines, pp. 633–656.

Timsina, J., Singh, U., Singh, Y., 1996. Addressing sustainability of RW systems: analysis of long-term experimentation and simulation. In: Kropff et al. (Eds.), Applications of Systems Approaches at the Field Level. Kluwer Academic Publishers, Dordrecht, The Netherlands, pp. 383–397.

Timsina, J., Adhikari, B., Ganesh-KC, 1997. Modelling and simulation of rice, wheat, and maize crops for selected sites and the potential effects of climate change on their productivity in Nepal. Consultancy Report submitted to Ministry of Agric., Harihar Bhawan, Kathmandu, Nepal. 55 p.

Timsina, J., Singh, U., Badaruddin, M., Meisner, C., 1998. Cultivar, nitrogen, and moisture effects on a RW sequence: experimentation and simulation. Agron. J. 90, 119–130.

Timsina, J., Singh, U., Godwin, D., Humphreys, E., 2004. Modelling chilling injury in rice using CERES-Rice ver. 4.1. In: CD-Rom Proceedings of the International Rice Cold Tolerance Workshop, 22–23 July 2004, Canberra, ACT, Australia.

Timsina, J., Humphreys, E., Godwin, D., Matthews, S., 2005. Evaluation of options for increasing water productivity of irrigated wheat using CSM-Wheat V4.0. In: MODSIM05 Conference, Melbourne, Australia.

Tongyai, C., 1994. Impact of climate change on simulated rice production in Thailand. In: Rosenzweig, C, Iglesias, A (Eds.), Implications of Climate Change for International Agriculture: Crop Modelling Study. US Environmental Protection Agency. EPA 230-B-94-003, Washington, DC.

Tsuji, G.Y., Uehara, G., Balas, S., 1994. DSSAT v3, vol. 2. University of Hawaii, Honolulu, Hawaii, p. 284.

Williams, R.L., Angus, J.F., 1994. Deep water protects high-nitrogen crops from low-temperature damage. Aust. J. Exp. Agri. 34, 927–932.

Willmott, C.J., 1982. Some comments on the evaluation of model performance. Bull. Am. Meterol. Soc. 63, 1309–1313.

Willmott, C.J., Ackleson, S.G., Davis, R.E., Feddema, J.J., Klink, K.M., Legates, D.R., O'Connell, J., Rowe, C.M., 1985. Statistics for the evaluation and comparison of models. J. Geophys. Res. 90 (C5), 8995–9005.