Annals of Botany **107**: 1367–1376, 2011
doi:10.1093/aob/mcr071, available online at www.aob.oxfordjournals.org

ANNALS OF BOTANY
Founded 1887

Simulating the impact of genetic diversity of *Medicago truncatula* on germination and emergence using a crop emergence model for ideotype breeding

S. Brunel-Muguet[1], J.-N. Aubertot[2] and C. Dürr[3,*]

[1]*INRA, UMR 950 Ecophysiologie, Agronomie et Nutritions NCS, Esplanade de la Paix, F-14032 Caen, France,*
[2]*INRA, UMR1248 Agrosystèmes et développement territorial, chemin de Borde-Rouge F-31326 Castanet-Tolosan, France and*
[3]*INRA, UMR 1191 Physiologie Moléculaire des Semences, 16 bd Lavoisier, F-49045 Angers, France*
** For correspondence. E-mail Carolyne.Durr@angers.inra.fr*

Received: 27 September 2010 Returned for revision: 1 February 2011 Accepted: 14 February 2011 Published electronically: 18 April 2011

• *Background and Aims* Germination and heterotrophic growth are crucial steps for stand establishment. Numerical experiments based on the modelling of these early stages in relation to major environmental factors at sowing were used as a powerful tool to browse the effects of the genetic diversity of *Medicago truncatula*, one of the model legume species, under a range of agronomic scenarios, and to highlight the most important plant parameters for emergence. To this end, the emergence of several genotypes of *M. truncatula* was simulated under a range of sowing conditions with a germination and emergence simulation model.
• *Methods* After testing the predictive quality of the model by comparing simulations to field observations of several genotypes of *M. truncatula*, numerical experiments were performed under a wide range of environmental conditions (sowing dates × years × seedbed structure). Germination and emergence was simulated for a set of five genotypes previously parameterized and for two virtual genotypes engineered to maximize the potential effects of genetic diversity.
• *Key Results* The simulation results gave an average value of 5–10 % difference in final emergence between genotypes, which was low, but the analysis underlined considerable inter-annual variation. The effects of parameters describing germination and emergence processes were quantified and ranked according to their contribution to the variation in emergence. Seedling non-emergence was mainly related to mechanical obstacles (40–50 %). More generally, plant parameters that accelerated the emergence time course significantly contributed to limiting the risk of soil surface crusting occurring before seedling emergence.
• *Conclusions* The model-assisted analysis of the effects of genetic diversity demonstrated its usefulness in helping to identify the parameters which have most influence that could be improved by breeding programmes. These results should also enable a deeper analysis of the genetic determinism of the main plant parameters influencing emergence, using the genomic tools available for this model plant.

Key words: Germination, emergence, modeling, numerical experiments, ideotypes, *Medicago truncatula*.

INTRODUCTION

Germination and heterotrophic growth are crucial steps for crop stand establishment. Both are under the control of environmental factors that interact with plant genotype. Plant genetic diversity is a potential source of adaptive responses to explore for the improvement of the processes leading to emergence. Crop emergence models have been designed to predict germination and emergence under different sowing and climatic conditions (Bouaziz and Bruckler, 1989*a*–*c*; Carberry and Campbell, 1989; Mullins *et al.*, 1996; Finch-Savage *et al.*, 1998; Forcella *et al.*, 2000; Dürr *et al.*, 2001; Colbach *et al.*, 2006*a*, *b*). Field studies are time consuming, costly, and not always feasible, but as alternative tools, numerical experiments enable exploration of a wide range of environmental conditions. However, simulation studies are still rare and most were carried out to predict the effects of farming practices and climatic conditions (Colbach *et al.*, 2005). The SIMPLE model (SIMulation of PLant Emergence; Dürr *et al.*, 2001) was designed to predict the effects of the main physical factors within the seedbed, i.e. soil temperature and water potential as well as mechanical obstacles to germination and emergence. Previous studies on germination and emergence using SIMPLE evaluated the effects of sowing conditions, e.g. sowing date, sowing depth and seedbed preparation, or of seed lot characteristics (Dürr *et al.*, 2001; Dorsainvil *et al.*, 2005; Moreau-Valancogne *et al.*, 2008). Fewer studies have been carried out to evaluate the effects of genetic diversity on output variables of existing predictive models [peach fruit growth (Quilot *et al.*, 2005*a*, *b*); nitrogen nutrition in pea (Voisin *et al.*, 2007)]. Another possible use of ecophysiological models is to engineer virtual genotypes or ideotypes [Rasmusson (1987); nitrogen absorption in pea (Voisin *et al.*, 2007); control of leaf surface in barley (Yin *et al.*, 1999, 2003); adaptation to nitrogen deficiency in wheat (Laperche *et al.*, 2006); peach fruit quality (Quilot *et al.*, 2005*b*); drought tolerance in maize (Chenu *et al.*, 2009)], and to test their performance under different environmental conditions. In the present study, the SIMPLE model was used to analyse the extent of the effects of genetic diversity observed in *Medicago truncatula*, a model species for

TABLE 1. *Basic equations and input parameters used to model seed germination and seedling emergence in the SIMPLE model (adapted from Dürr* et al.*, 2001 and Dorsainvil* et al.*, 2005)*

Phase	Equations and output variables	Input variables and parameters
Sowing to germination	$STT_{id} = \sum_{d=1}^{n} \left[(T_d - T_{b,germ}) I_{\lvert\Psi_d\rvert - \lvert\Psi_{b,germ}\rvert} \right]$ (eqn 1)	T_d and Ψ_d: daily mean soil temperature and water potential at SD_i sowing depth of the seed i STT_{id}: calculation of the sum of thermal time cumulated at day d by seed i
	where $I_{\lvert\Psi_d\rvert - \lvert\Psi_b\rvert} = 1$ if $\lvert\Psi_d\rvert < \lvert\Psi_{b,germ}\rvert$	$T_{b,germ}$ and $\Psi_{b,germ}$ base temperature and water potential values for germination
	otherwise $I_{\lvert\Psi_d\rvert - \lvert\Psi_{b,germ}\rvert} = 0$	STT_i: cumulated thermal time required for germination of seed i, drawn at random from the $STTg$ distribution of the studied seedlot – this value allows the calculation of day Gi of seed i germination using eqn (1)
	Gi when $STT_{id} = STT_i$	
Germination to emergence	$l(t) = a[1 - \exp(-bt_d)^c$ (eqn 2)	$l(t)$: calculation of hypocotyl length at time t after germination. a, b, c: hypocotyl Weibull elongation function parameters
	where $t_d = TT_d = \sum_{d=STT_i}^{n} (T_d - T_{b,elon})$ (eqn 3)	TT_d: cumulative thermal time from $STTi$ to day d $T_{b,elon}$: base temperature value for elongation
	HL_i is a function of SD_i and soil structure	HL_i is the hypocotyl length to reach the soil surface; it depends on SD_i and clods circumvented by seedling i. HL_i is used to calculate the day Ei on which seed i reaches the soil surface according to eqns (2) and (3)
	$P_i = 0$ if $L < L_0$	Seedling i grows along the clod if the seedling does not remain trapped under the clod
	$P_i = 100\{1 - \exp[-\alpha(L - L_0)]\}$ if $L \geq L_0$	P_i: probability for the seedling i to be trapped under a clod of a given diameter L
	E_i	α, L_0 parameter values are given for buried clods and clods laid on the soil surface
Soil surface emergence	*Soil is crusted if cumulated rainfall since sowing (CR) or if daily rainfall (DR) over threshold values. A probability p for seedling emergence is associated each day after soil crusting to wet or dry crust.	$*CR = 12$ mm; $*DR = 5$ mm p_i: probability for the seedling i to emerge through a wet or a dry crust with $p_i = 1$ for a wet crust or $p_i =$ genotype-specific value, for a dry crust
	$TT_{survival}$: maximal below-ground survival time for a seedling after germination	$*TT_{survival} = 175$ °Cd

Plant input parameters are in bold.
* Default values were obtained from previous characterization (Duval and Boiffin, 1994; Brunel *et al.*, 2009).

genomic studies, on emergence rates under a wide range of environmental conditions. Several genotypes of *M. truncatula* were parameterized in a previous study (Brunel *et al.*, 2009) that explored possible genetic diversity in plant parameters and highlighted genotypes with contrasting parameter values in the germination and pre-emergence processes as formalized in SIMPLE. SIMPLE was also run to simulate the behaviorsof virtual engineered genotypes to maximize the possible effects of genetic diversity on emergence and to identify the parameters that have the most impact under specific environmental conditions.

MATERIALS AND METHODS

Description of the SIMPLE model

SIMPLE is a stochastic model simulating germination and emergence as a function of seedbed characteristics (sowing depth distribution, and the proportion and spatial organization of aggregates), soil and climate conditions (soil temperature and water potential, rainfalls) and plant parameters. Its functioning is described in full in Dürr *et al.* (2001). Here, only the main principles and characteristics are recalled. Table 1 lists the main equations used for the prediction of germination and emergence with the required plant parameters, which are genotype specific.

A numerical 3-D representation of the seedbed is produced using input variables related to the number, size, proportion and spatial distribution of clods. Daily top-soil temperatures and water potentials are also input variables. Simulations are drawn for each seed (i). For the seed i, its sowing depth (SD_i) and the time required for its germination (STT_i) are drawn at random from the distributions of sowing depths and of seed germination thermal times (STT_g). The calculation of the required thermal time STT_i depends on the genotype base temperature ($T_{b,germ}$) and on mean daily soil temperature (T_d). Thermal time is cumulated from sowing to day d only when the mean daily soil water potential of day d (Ψ_d) in the soil layer corresponding to the sowing depth is greater than the base water potential ($\Psi_{b,germ}$) characterizing the genotype. If STT_i is reached, then the seed is declared germinated and seedling growth starts. Seedling length needed to reach the soil surface (HL_i) depends on its sowing depth (SD_i), and on the length needed to reach the soil surface among buried clods and clods on the surface of the soil. The thermal time required to grow until this length (HL_i) is reached depends

TABLE 2. *Values of the plant input parameters of SIMPLE for the five observed genotypes and the two engineered genotypes*

	Genotype						
	A17	F83005·5	DZA315·16	Paraggio	DZA045·5	*MtrFirst*	*MtrLast*
Germination							
Base temperature, $T_{b,germ}$ (°C)	1·8	1·7	0·9	1·9	1·9	1·9	3
Germination percentages per thermal time class STT_g							
0–10 °Cj (%)	19	2	0	3	5	61	17
11–20 °Cj (%)	67	76	55	81	81	39	74
21-25 °Cj (%)	6	11	24	10	9	0	5
>25 °Cj (%)	4	4	19	6	4	0	2
Residual percentage of non-germinated seeds	4	7	2	0	0	0	2
Base water potential $\Psi_{b,germ}$ (Mpa)	−0·7	−0·75	−0·55	−0·6	−1·21	−1·32	−0·57
Heterotrophic growth							
Base temperature for elongation $T_{b,elon}$ (°C)	7·1	6·6	3·1	5·6	7	5·4	7·4
Parameters of the Weibull elongation function							
a (mm^{-1})	65·9	68·9	64·5	86·5	66·1	72·4	62·7
b ($°Cd^{-1}$)	0·04	0·03	0·01	0·02	0·02	0·032	0·023
c	1·97	1·28	1·29	1·47	1·05	1·592	2·003
Mechanical obstacles							
Parameters of the probability function of death due to clod trapping							
(1) Buried clods							
α (mm^{-1})	0·03	0·04	0·02	0·02	0·03	0·02	0·04
L_0 (mm)	15	10	15	15	15	15	10
(2) Clods laid on the soil surface							
α mm^{-1}	0·02	0·03	0·02	0·02	0·02	0·02	0·03
L_0 mm	25·5	26·0	26·0	29·1	25·5	29·1	26·0
Distribution of seedling emergence forces per force value class (%)							
≤0·10 (N)	94	69	90	55	94	55	90
>0·10 (N)	6	31	10	45	6	45	10
Probability for a seedling to emerge through a dry crust							
p %	40	40	40	50	40	50	40

on T_d and on the base temperature for elongation ($T_{b,elon}$), which is genotype specific. It is calculated from the time of germination for seed i (STT_i) to time n using a Weibull elongation function whose parameters (a, b, c) are genotype specific. The seedling i may face clods along its path in the seedbed, or a crust at the soil surface that prevents its emergence. The proportion of seedlings trapped under a clod depends on the size and position of the clod in the soil, described by a probability function (P_i), the coefficients (α and L_0) of which are genotype specific. Rainfall data are used for prediction of crust formation (Table 1). When a seedling reaches the soil surface, its ability to emerge through a crust depends on its emerging force drawn at random from a distribution, which is genotype specific. This distribution is used to determine the probability (p_i) for a seedling to remain blocked under a crust. If it does not emerge, the seedling is considered as dead when time from germination is greater than a survival time value ($TT_{survival}$) below the soil surface. If this time is not yet reached, then the state of the crusted soil surface is analysed again the following day and the emergence prediction process is reiterated until seedling emergence or seedling death if the survival time is reached.

Studied and engineered genotypes of Medicago truncatula

Genotypes of *M. truncatula* were sown in experiments carried out to compare model germination and emergence predictions with observations. Table 2 lists the set of parameter values measured in laboratory conditions required to run the simulations, established during a previous study whose aim was to analyse the genetic diversity of *M. truncatula* during germination and heterotrophic growth (Brunel *et al.*, 2009). Paraggio and Jemalong A17 are cultivars, the latter also being the reference for genomic studies. F83005·5 and DZA315·16 are genotypes derived from natural populations and belong to the nested core collections of *M. truncatula* (Ronfort *et al.*, 2006). In the seed lot used in the field experiment, DZA315·16 was shown to differ from the three others with slower germination and elongation rates. For numerical experiments, DZA045·5, another previously studied genotype in the nested core collections, was added because of its low sensitivity to water stress. Two virtual genotypes, *MtrFirst* and *MtrLast*, were engineered to maximize genotypic effects. These virtual genotypes were designed using the set of parameter values of the genotypes studied by Brunel *et al.* (2009), *MtrFirst* gathering all the favourable values for parameters, and *MtrLast* all the most unfavourable parameter values that were assumed to either improve or penalize their stand establishment performances.

Experiments for the evaluation of the model

Two experiments were carried out to test the model predictions. The aim of the first experiment was to test equations and parameter values for predictions of germination and emergence under conditions with no mechanical or water stresses.

Tanks ($40 \times 30 \times 10$ cm^3) were filled with soil (0·17 g g^{-1} clay, 0·75 g g^{-1} silt, 0·40 g g^{-1} sand) without any clods (<5 mm), with a volumetric mass of 1 kg dm^{-3}, maintained at 0·19 g g^{-1} water content (corresponding to −0·07 MPa for the texture of the soil used). Three replicates of the Paraggio cultivar were sown at a depth of 2 cm (seven lines of 30 seeds). The tanks were closed with aluminum foil and held at 15·1 ± 0·49 °C (recorded with temperature sensors; Testo 177-T3, calibrated by Testo and by internal procedures). At the seven sampling times, they were opened using the tank frontal door. At each time, a line of 30 seeds was carefully extracted, going from the front to the bottom of the tank. The opened soil layer was closed with aluminum foil, and the front door was closed so as to avoid soil disturbance and loss of water for the next remaining line sown (located 5 cm away). The number of the 30 extracted seeds that had germinated was recorded. Three other tanks were sown and placed in the same conditions and seedlings which emerged were observed regularly.

The aim of the second experiment was to test model predictions in realistic field conditions. It was carried out in the field at the National Seed Testing Station in Angers (western France, 47°28′N, 0°33′W) in a silt sandy soil (0·15 g g^{-1} clay, 0·40 g g^{-1} silt, 0·45 g g^{-1} sand). The seedbed was prepared with a harrow. The experiment used a randomized complete block design for observations of germination and emergence. Each of the three blocks comprised four sowing lines, i.e. one per genotype observed in the experiment: Jemalong A17, Parragio, F83005·5 and DZA315·16. On each line of each block, 100 seeds were sown on 16 October 2007, at an average depth of 1·6 (± 0·4) cm and 3-cm spacing. After sowing, rainfall events were recorded daily. Temperatures at the seed sowing depth were recorded at hourly intervals using temperature sensors (one per block; Testo 177-T3). Soil water contents were measured by sampling soil at different depths from the surface to a depth of 10 cm at 2-cm intervals. The soil water potential was then determined using a relationship between soil water content and soil water potential depending on soil texture (data given on request from Infosol, INRA Orléans; http://bdat.gissol.fr/geosol/index.php). Seedbed samples were taken and sieved to determine the number of clods per size class. Each block was divided into two areas: one dedicated to germination observations and another to emergence observations. Germinated seeds were observed in this specific area by destructive measurements on 30 seeds (10 per block) at six sampling dates. Seed sowing depths were measured at the same time. Seedling emergence was recorded twice a day until a plateau was reached. Then, causes of non-emergence were observed in the seedbed on non-emerged individuals identified by gaps along the sowing lines, i.e. a total of 35–40 non-emerged individuals observed according to the genotype. Non-germinated seeds, seedlings blocked under clods, seedlings trapped under a crust and glassy thick abnormal seedlings were distinguished. Observations of germination and emergence were compared with simulations by running the SIMPLE model using the genotype parameter values measured in the laboratory, and the environmental conditions measured during the field experiment.

Numerical experiments

The simulation was performed with the five studied genotypes and the two virtual genotypes, *MtrFirst* and *MtrLast* (Table 2). Meteorological data were from INRA station at Mons en Chaussée (northern France, 49°52'N, 3°0'E) over 9 years (1994–2002). Seven sowing dates were simulated from 1 August to 1 October. The simulated seedbed was a fine seedbed with no clods >20 mm in diameter. Soil surface crusting parameters were those for a silt clay soil type. The distribution of sowing depths ranged from 5 to 30 mm, with a mean value of 2 mm. The numerical experiment represented a total of 882 simulated sowings.

Statistical criteria for prediction analyses

The modelling efficiency statistic (EF) between n simulated (P_j) and measured (O_j) variables was used as an indicator of the goodness predictions (Smith *et al.*, 1996):

$$\mathrm{EF} = 1 - \frac{\sum_{j=1}^{n} (P_j - O_j)^2}{\sum_{j=1}^{n} (O_j - \overline{O})^2} \quad (1)$$

where j varies from 1 up to n ($n = 12$) for the successive times of observations, and $\overline{O}$ is the mean of the observed values. For a perfect model prediction EF = 1.

The root mean square error (RMSE) provides the mean difference between n predicted and observed values. The unit of the coefficient is the same as that of the analysed variables:

$$\mathrm{RMSE} = \sqrt{\sum_{j=1}^{n} \left[(P_j - O_j)^2 / n \right]} \quad (2)$$

The coefficient of residual mass (CRM) is a measure of the tendency of the model to under- or overestimate predicted values compared with observations. A negative value indicates that the majority of predicted values are less than the measured ones (i.e. simulated time courses are ahead of the observed ones).

$$\mathrm{CRM} = \frac{\left(\sum_{j=1}^{n} O_j - \sum_{j=1}^{n} P_j \right)}{\sum_{j=1}^{n} O_j} \quad (3)$$

Means and standard deviations of the analysed output variables were calculated by repeating the simulation process three times. Analyses of variance and mean comparison tests ($P < 0{\cdot}05$, Tukey's multiple comparisons procedure) for genotype ranking were carried out using STATGRAPHICS Plus 3·1 software.

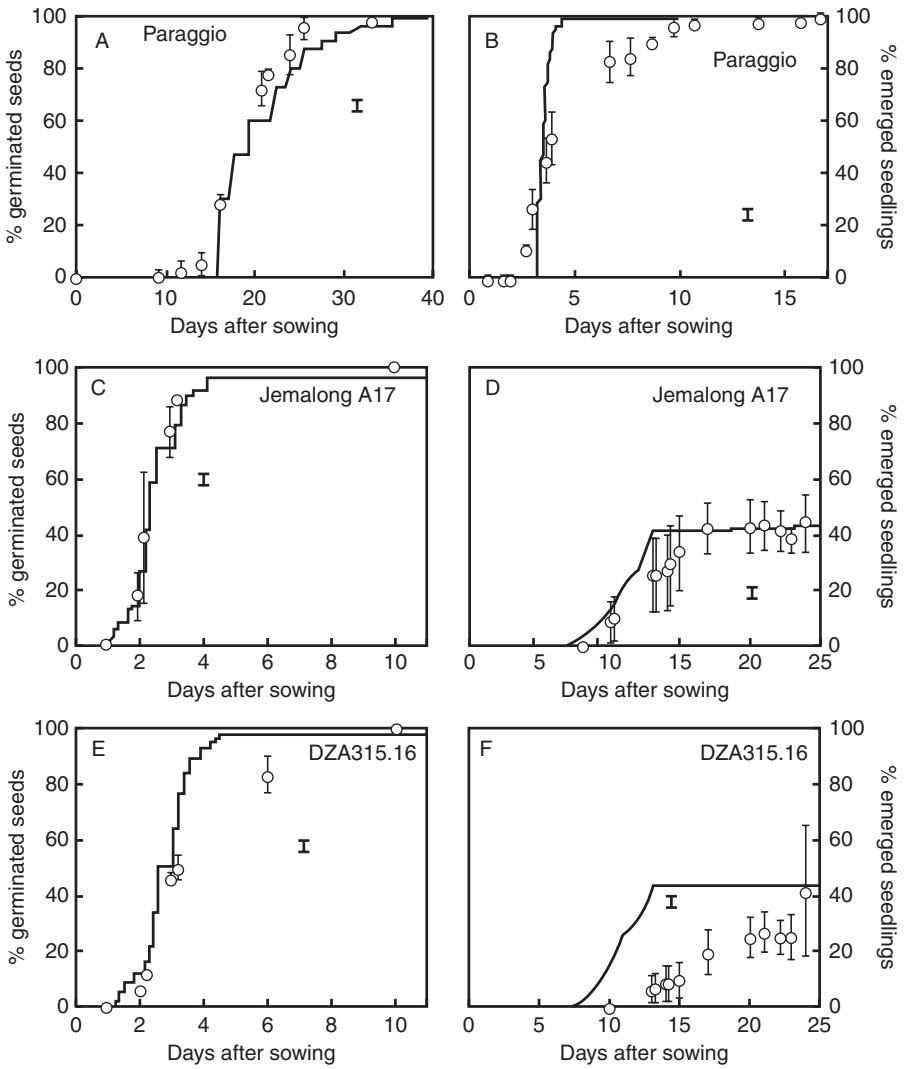

FIG. 1. Observed and simulated germination and emergence time courses in tanks for Paraggio (A,B) and (C–F) under field conditions for Jemalong A17 and DZA315·16. Symbols and lines represent observations and simulations, respectively. Error bars associated with symbols are s.d. for observations. The error bar below the line is the s.d. for three replicated simulations.

RESULTS

Comparison of simulation results with observed germination and emergence

Figure 1(A,B) compares observed and simulated germination and emergence for experiments in tanks. Germination and emergence rates were high, both close to 100 %. Simulated germination times and final rates, as well as emergence rates, were close to the observed results in these experimental conditions, with slight underestimation of germination rates (CRM $>$ 0; RMSE = 8·25 %) and overestimation of emergence rates especially between 5 and 10 d after sowing (CRM $<$ 0; RMSE 20·6 %). The model efficiency values (EF = 0·96 and 0·75 for germination and emergence, respectively) indicate good accuracy. Under conditions without mechanical or water stress, the model equations and the parameter values given for the studied genotypes enabled good predictions.

As regards the field experiment, soil temperatures were quite cold for *M. truncatula*, ranging from 8 °C to 12 °C and the soil water content in the first 2 cm of soil was low (0·06 g g^{-1}, i.e. below −1 MPa for this soil texture) for 24 h after sowing, until 10 mm of rain fell rewetting the soil but causing soil surface crusting. The soil surface remained dry thereafter as no rainfall occurred for 14 d, but the water content of the first 2 cm and below, where the seedling radicle elongated, remained over − 1MPa. Figure 1(C–F) illustrates observed and simulated germination and emergence times in the case of Jemalong A17 and DZA315·16, which presented the most extreme results in the field and, for comparison, between observed and predicted values. Observed final germination percentages were high, 95 to 100 % for the four genotypes. It took about 3 d for A17 to reach 80 % of germination (t80 %germ), i.e. 28 °Cd. Germination was slowed down because of the very low soil water content during the first day after sowing (e.g. t80 %germ for A17 was $<$20 °Cd in non-limiting

water conditions; Table 2). Germination of DZA315·16 was even slower than that of the other genotypes, requiring >50 °C to reach 80 % of germination. Seedling emergence was very slow and the final emergence percentages (FPemergence) were low, from 42 to 50 % depending on the genotype. The same as for germination, DZA315·16 emerged more slowly, the first seedlings emerging 3 d later than the other genotypes.

For all the genotypes, germination simulations were close to observations. The mean differences between predicted and observed germination percentages at a given time (RMSE) varied between 7 % and 28 % depending on the genotype, but differences in the prediction of germination times were less than half a day (Fig. 1C–F), with a tendency to a slight overestimation of the simulated time courses (positive values of CRM), except for DZA315·16.

Concerning emergence, simulated FPemergence was low and fitted that observed, i.e. 40–50 % (Fig. 1C–F). Field observations of non-emerging seedlings (data not shown) indicated that the main cause was seedlings trapped under a crusted soil surface (46–58 % of sown seeds) and the predictions of the percentage of seedlings trapped under a dry crust fitted the observations. The percentage of non-germinated seeds was nil in the field, and slightly overestimated in the simulations because of a slightly too-high percentage (0–7 %) given from the laboratory results (Table 2). The other two causes observed for non-emerging seedlings were (1) only a small percentage of seedlings were blocked under a clod (<5 %) as the seedbed had few clods, which was correctly predicted, and (2) abnormal seedlings (5–16 %, glassy thick seedlings), observed in the field, probably due to the cold conditions, but whose prediction was not included in the simulation model. Simulated emergence times were underestimated compared with field observations (Fig. 1C–F; CRM < 0). These discrepancies could be explained by the $T_{b,elon}$ parameter which was estimated from too few results for the seed lot sown in the field experiment, and also by the *b* coefficient of the elongation function obtained in laboratory experiments under low mechanical resistance, compared with more compacted soil conditions in the field experiment.

Finally, all the parameter values concerning germination allowed accurate predictions. Final emergence percentages were also correctly predicted. The largest discrepancies were the underestimation of emergence times of 2–5 d in poor emergence conditions lasting 20–25 d. These discrepancies could be partially explained by a lower quality of estimations of some of the model parameter values for the seed lot used in the field experiment. The other main discrepancy was abnormal seedlings not predicted by the model in cold conditions. The results of the simulation study should be considered keeping these main observations in mind.

Numerical experiments with the genotypes of Medicago truncatula *in a range of environmental conditions*

Figure 2 presents the simulation results for the fine seedbed structure, for seven autumn sowing dates over 9 years for the five genotypes of *M. truncatula* studied and whose parameter values are summarized in Table 2, i.e. 315 sowing simulations. The output variables chosen for the illustration are FPemergence and the simulated causes for non-emerging seedlings, i.e. non-germination (%NG), seedlings blocked under clods (%clod) and seedlings trapped under a crust (%crust).

First, the simulated average FPemergence was as low as 40–70 % whatever the sowing date, close to those observed in the field experiment (Figs 1E,F and 2A). Interannual variations were very high as indicated by the large standard deviations. For instance, at the 20 September sowing date, extreme FPemergence varied from 20 % to 95 %, depending on the year, with t80 %germ ranging from 2 d to 10 d, and the time to reach 30 % emergence (t30 %emergence) from 3 d to 25 d (data not shown). The mean differences between genotypes were generally small, i.e. 10–15 % for FPemergence (Fig. 2), and 1–4 d for t80 %germ, 1–4 d for t30 %emergence depending on the sowing date (not shown). But when simulation results were analysed for each sowing date and year, the differences could be much larger. For instance, on 10 September 2002 and in the days that followed, conditions were very dry, and the time to reach 30 % of emergence varied from 3·5 d to 25 d after sowing with a FPemergence of 20 % for DZA315·16 whereas FPemergence reached 60 % for Paraggio and DZA045·5. Adverse conditions increased genotypic differences. The two genotypes F83005·5 and DZA315·16, exhibited the lowest FPemergence at each date, although the differences were not significant ($P < 0{\cdot}05$, Tukey test), because of the high interannual variation. These two genotypes were characterized by unfavourable parameter values related to the effects of water stress ($\Psi_{b,germ}$), temperature during germination ($T_{b,germ}$ and *STT*), the ability to elongate and the capacity to overcome mechanical obstacles (α, L_0 and *p*; Table 2).

The most important simulated cause of non-emerging seedlings was soil surface crusting, blocking on average as many as 30–55 % of the seedlings whatever the genotype and the sowing date (Fig. 2D). Between-year variability was very high whatever the sowing date. On average, DZA315·16 was the genotype that was the most affected by seedling trapping under soil crusts. It had a low *b* coefficient value, which increased its time to reach the soil surface and consequently the risk that rainfall would occur and lead to the formation of a crust at the soil surface before emergence. Paraggio, which had favourable general characteristics and the highest probability to emerge through a crust (*p*, Table 2), showed the lowest rate of seedlings trapped under a crust. Both parameters had a direct or indirect impact on the ability to emerge when the formation of a dry crust occurred. The second cause of non-emergence, although this had less impact on FPemergence, was non-germinated seeds, ranging from 0 % to 15 % on average but showing high between-year variability (Fig. 2B). Extreme differences between genotypes appeared under dry and cold conditions. Such conditions significantly affected F83005·5 and DZA315·16, which had the least favourable parameter values concerning cold and water stress, respectively (*STTg*, $T_{b,germ}$, $\Psi_{b,germ}$; Table 2). Moreover, the remaining non-germinated seeds were the highest for F83005·5 (Table 2). By contrast, DZA045·5, which was characterized by highest tolerance to water stress and, along with Paraggio, had the lowest %NG, had no non-germinating seeds under dry conditions such as those that occurred in the two last decades of September in 1996, 1997

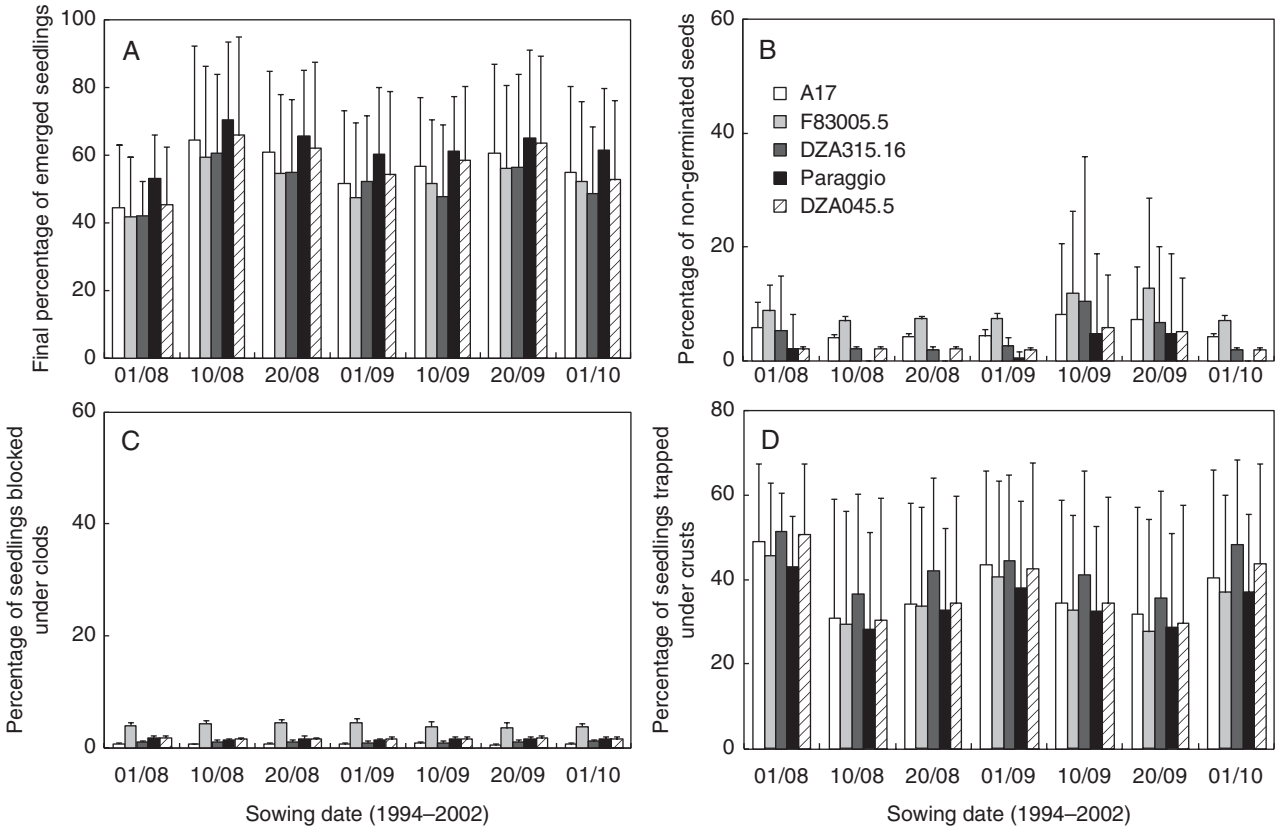

FIG. 2. Simulation results for seven sowing dates from 1994 to 2002 for the five genotypes: (A) final percentage of emerged seedlings; (B–D) percentages of the causes of non-emergence. Data are average values of the output variables; error bars are s.d. for the nine simulated years.

and 2002 (data not shown). Finally, the proportion of seedlings blocked under clods was very low as the simulated seedbed had very few clods (Fig. 2C). The most affected genotype was F83005·5 with >5 % of seedlings blocked under clods even in the fine seedbed. This genotype had the most unfavourable value of both parameters related to this effect, i.e. α and L_0 (Table 2). For numerical experiments performed by simulating a seedbed with more large clods but still within the range of those usually observed in seedbeds for this size of seeds under field conditions, the percentage of seedlings that remained blocked under clods reached 18–28 % on average, whatever the genotype and the sowing date (not shown).

Numerical experiments with the two engineered genotypes with extreme parameter values

Figure 3 shows results for the two extreme virtual ideotypes along with those of Paraggio, which was used as a reference for a genotype cultivated in Australia with good emergence results. Like the previous simulations performed with the five genotypes studied, high variability of emergence results was observed for all sowing dates whatever the genotype i.e. FPemergence varied on average from 40 % to 75 % (Fig. 3A), t80 %germ from 3 d to 8 d, and t30 %emergence from 5 d to 14 d after sowing (not shown). Differences between *MtrFirst* and *MtrLast* varied from 15 % to 20 % on average for FPemergence, 1 d to 5 d for t80 %germ, 2 d to 6 d after sowing for t30 %emergence. Mean differences between Paraggio, the cultivar currently cultivated, and *MtrFirst*, including all the best values for emergence parameters, were very small, being <5 % for FPemergence whatever the sowing date. However, differences increased depending on climatic conditions as shown by the large inter-annual variability. For instance, under the dry conditions recorded on 10 September in 1996 and 2002, NG % could reach 85 % for *MtrLast* which greatly affected FP %emergence, whereas it remained as low as 5 % for *MtrFirst*. Under the cold and rainy conditions recorded on
10 September 1994, drastic differences were observed between the two virtual ideotypes, i.e. %crust was 50 % and 5 %, leading to 40 % and 95 % of FPemergence for *MtrLast* and *MtrFirst*, respectively. At this specific date, differences between Paraggio and *MtrFirst* were also greater with FPemergence 40 % lower for Paraggio. As these two genotypes mainly differed in their elongation parameter value (*b*, Table 2), the final difference in emergence resulted from more *MtrFirst* seedlings reaching the soil surface when no crust had yet been formed, unlike for Paraggio. Therefore, if the mean differences between the two extreme genotypes remained limited, contrasting behaviours were accentuated under unfavourable climatic conditions, especially dry conditions or cold associated with rainy conditions. Parameter values linked to water stress tolerance, and to the ability to break through a crust had the largest influence on emergence

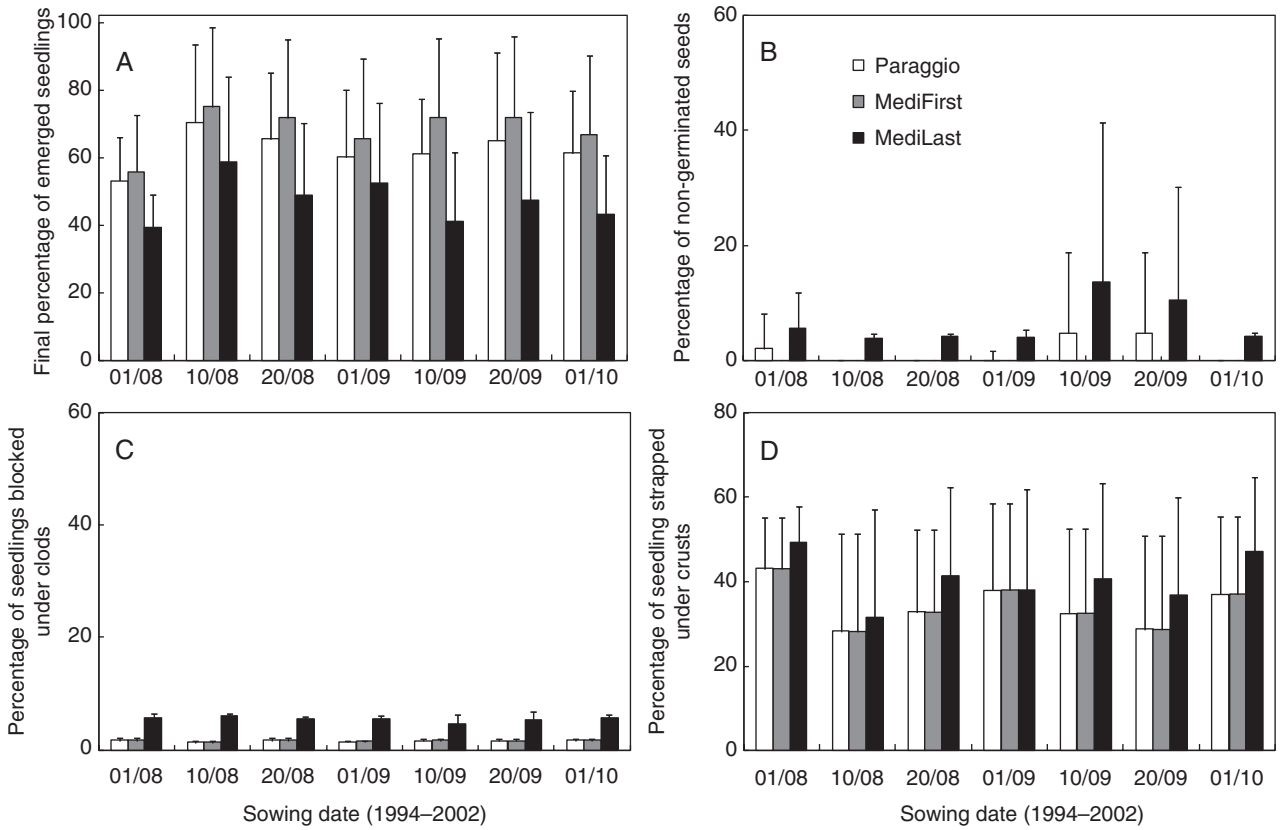

FIG. 3. Simulation results for seven sowing dates from 1994 to 2002 for the engineered genotypes and the cultivar Paraggio: (A) final percentage of emerged seedlings; (B–D) percentages of the causes of non-emergence. Data are average values of the output variables; errors are s.d. for nine simulated years.

results, along with those that reduced the time to reach the soil surface (high germination and elongation rates), which reduce the risk of coming up against a crusted soil surface.

DISCUSSION

Simulations highlighted major genotypic-dependant traits for plant emergence under a wide range of environmental conditions

Results of model evaluation attested that its predictive quality was satisfactory for its further use, aiming at evaluating the impact of genotypic differences on germination and emergence under a wide range of environmental conditions. This study illustrates the possible use of an agro-ecophysiological crop model to describe genetic diversity in a model species mainly used for genomics and to predict emergence results for a wide range of environmental conditions. The formalism used to represent germination and heterotrophic growth processes and the parameterization of the species allowed genotypic differences to be simulated and the extent of their effects on emergence results to be quantitatively evaluated, even if the model should be further improved to take better account of some of the effects of cold conditions leading to abnormal seedling growth for which rates are not predicted.

Genotypic differences in mean final emergence rates did not exceed 5–10 % whatever the sowing date but were exacerbated under extreme conditions. The simulations also gave the main causes for non-emergence. These are labour-intensive data that are almost impossible to collect in field experiments. Analysis of these causes helps to focus on parameters that distinguish genotypes in real sowing conditions. These simulations required prior fine phenotyping to obtain parameters values that are genotype dependant. Although differences in parameter values were observed between the genotypes studied (Brunel *et al.*, 2009), it was not possible to evaluate the effects of such genetic diversity on emergence as these parameters interact with environmental conditions, which determine final emergence performance. This evaluation requires field experiments but these are time consuming and costly. Therefore, they are most often limited in time and space and combine and compare only a limited range of environmental conditions. Numerical experiments make it possible to dissect the effects of variations in plant parameters in a wider range of environmental contexts. They allow average genotypic differences to be revealed by multiplying sowing conditions and extreme differences to be highlighted under specific environmental conditions. Such conditions may be missing if the number of field experiments is limited.

This study demonstrated that modeling is a powerful tool to (*a*) analyse interactions between genotype characters and the environment, (*b*) determine a hierarchy of parameters according to their respective impact on emergence and (*c*) identify the most important environmental conditions that emphasize genotypic differences. The first numerical experiment with

the set of previously studied genotypes over a range of environmental conditions demonstrated the importance of parameter values limiting the effects of mechanical obstacles which accounted for the main limiting conditions in field conditions (crusts and clods). Among these parameters, the emergence force is important when a seedling is under mechanical constraints and has been shown to be correlated with seed mass (Sinha and Guildyal, 1979; Tamet *et al.*, 1996). Aside from this parameter, all the traits that enhance germination and elongation time courses are also favourable as they improve emergence speed, which limits the time to reach the soil surface and thus the risk of the occurrence of rainfall after sowing leading to crust formation. Water stress tolerance, i.e. low $\Psi_{b,germ}$, is another important trait for reducing the time to reach the soil surface. The other simulations emphasized the extent of genetic diversity by engineering genotypes that combined extreme parameters values, although still within a natural range of variations. These results once again showed the importance of specific environmental conditions in exacerbating genotypic differences since, on average, for a given sowing date, differences between *MtrFirst* and *MtrLast* remained limited. Interestingly, when compared with the most-efficient engineered genotype *MtrFirst*, the behaviour of the cultivar Paraggio was quite close. This indicates that a large part of genetic improvement has already been made, even though these favourable traits were not directly selected during breeding. Finally, the analyses resulting from simulations aimed at understanding processes and associated traits that would greatly improve crop stand establishment, i.e. adaptive mechanisms to extreme conditions (temperature and water stresses) and high potential for rapid heterotrophic elongation together with a high emergence force exerted by seedlings. Consequently, such indications should orientate towards the phenotyping of target traits on seeds and seedlings.

Perspectives for the use of the model

Pioneering studies integrate the genetic information of plant parameters in ecophysiological models. Their interest lies in the possibility of expressing the plant parameters, also referred as genotype-dependant parameters, since their values are independent of the environment, through the genetic effects of the QTLs (quantitative trait loci) involved in their variations (Quilot *et al.*, 2005*b*). This approach has already been used in other species and other stages of the plant cycle [bean (White and Hoogenboom, 1996); maize (Reymond *et al.*, 2003); barley (Yin *et al.*, 2003); rye (Chapman *et al.*, 2003; Quilot *et al.*, 2005*b*); soybean (Messina *et al.*, 2006; Dorlodot *et al.*, 2007); maize (Chenu *et al.*, 2009; Yin and Struick, 2010)]. The interest in using a parameter value that is predicted by a QTL model instead of a conventionally measured value is to engineer virtual ideotypes by judiciously combining certain alleles at the key QTLs controlling the parameter, in such a way that enhances its value. With the increasing development of molecular markers for crop breeding programmes, the selection of favourable alleles of the markers underlying the key QTLs controlling the parameter variation is conceivable. This modelling approach, referred to as 'Gene-to-Phenotype' (Cooper *et al.*, 2005, 2009; Chenu *et al.*, 2009; Yin and Struick, 2010) is challenging and would be useful for linking ecophysiological models designed to predict performances under varying environmental conditions to genetic models likely to capture the effects of molecular variability. With the aim of integrating genetic information in the agro-ecophysiological model SIMPLE, the choice of the model legume *M. truncatula* for this study, is of great interest, as a set of genetic and genomics tools are available making possible QTL mapping (e.g. Dias *et al.* 2011) and a deeper analysis of the genetic determinism of the main emergence parameters for a model plant.

The present work illustrates the necessity of the modeling approach which allows plant performances to be tested under a wide range of environmental conditions, not only for crop species as previously demonstrated (Dürr *et al.*, 2001; Dorsainvil *et al.*, 2005; Moreau-Valancogne *et al.*, 2008) but also for the model species. Owing to the formalisms used in the model, genetic variability can be investigated through values of the input variables and parameters. Such new approaches for further genetic studies on *Medicago truncatula* or other plants should contribute to a more effective dialogue between scientists working at different scales for effective agricultural research (Passioura, 2010).

ACKNOWLEDGEMENTS

We thank the National Seed Testing Station for helpful assistance in seed production and in field experiments. We thank L. Ledroit, E. Peltier and M.-H. Wagner, for their involvement in the experiments, and J. Léonard, for his help with model computation and parameterization. This research was funded by INRA, the Regional council of Pays de la Loire and GEVES (National Organization for Evaluation of Seeds and Cultivars).

LITERATURE CITED

Bouaziz A, Bruckler L. 1989*a*. Modelling of wheat imbibition and germination as influenced by soil physical properties. *Soil Science Society of America Journal* **53**: 219–227.

Bouaziz A, Bruckler L. 1989*b*. Modelling wheat seedling growth and emergence. I. Seedling growth affected by soil water potential. *Soil Science Society of America Journal* **53**: 1831–1838.

Bouaziz A, Bruckler L. 1989*c*. Modelling wheat seedling growth and emergence. II. Comparison with field experiments. *Soil Science Society of America Journal* **53**: 1838–1846.

Brunel S, Teulat-Merah B, Wagner M-H, Huguet T, Prosperi J-M, Dürr C. 2009. Using a model-based framework for analysing genetic diversity during germination and heterotrophic growth of *Medicago truncatula*. *Annals of Botany* **103**: 1103–1117.

Carberry PS, Campbell LC. 1989. Temperature parameters useful for modeling the germination and emergence of pearl millet. *Crop Science* **29**: 220–223.

Chapman S, Cooper M, Podlich D, Hammer G. 2003. Evaluating plant breeding strategies by simulating gene action in dryland environment effects. *Agronomy Journal* **95**: 99–113.

Chenu K, Chapman SC, Tardieu F, McLean G, Welcker C, Hammer GL. 2009. Simulating the yield impacts of organ-level quantitative trait loci associated with drought response in maize: a "Gene-to-Phenotype"modeling approach. *Genetics* **183**: 1507–1523.

Colbach N, Dürr C, Roger-Estrade J, Caneill J. 2005. How to model the effects of farming practices on weed emergence. *Weed Research* **45**: 2–17.

Colbach N, Dürr C, Roger-Estrade J, Chauvel B. 2006*a*. ALOMYSYS: modelling black-grass (*Alopecurus myosuroides* Huds.) germination and

emergence in interaction with seed characteristics, tillage and soil climate. I. Construction. *European Journal of Agronomy* **24**: 95–112.

Colbach N, Busset H, Yamada O, Dürr C, Caneill J. 2006*b*. ALOMYSYS: modelling black-grass (*Alopecurus myosuroides* Huds.) germination and emergence in interaction with seed characteristics, tillage and soil climate. II. Evaluation. *European Journal of Agronomy* **24**: 113–128.

Cooper MD, Podlich DW, Smith OS. 2005. Gene-to-Phenotype models and complex trait genetics. *Australian Journal of Agricultural Research* **56**: 895–918.

Cooper M, Van Eeuwijk FA, Hammer GL, Podlich DW, Messina C. 2009. Modeling QTL for complex traits: detection and context for plant breeding. *Current Opinion in Plant Biology* **12**: 231–240.

Dias MBP, Brunel-Muguet S, Dürr C, *et al*. 2011. QTL analysis of seed germination and pre-emergence growth at extreme temperatures in *Medicago truncatula*. *Theoretical and Applied Genetics* **122**: 429–444.

Dorlodot S, Forster B, Pagès L, Prcie A, Tuberosa R, Draye X. 2007. Root system architecture: opportunities and constraints for genetic improvements of crops. *Trends in Plant Science* **12**: 474–481.

Dorsainvil F, Durr C, Justes E, Carrera A. 2005. Characterisation and modelling of white mustard (*Sinapis alba* L.) emergence under several sowing conditions. *European Journal of Agronomy* **23**: 146–158.

Dürr C, Aubertot JN, Richard G, Dubrulle P, Duval Y, Boiffin J. 2001. SIMPLE: a model for SIMulation of PLant Emergence predicting the effects of soil tillage and sowing operations. *Soil Science Society of America Journal* **65**: 414–42.

Finch-Savage WE, Steckel JRA, Phelps K. 1998. Germination and post-germination growth to carrot seedling emergence: predictive threshold models and sources of variation between sowing occasions. *New Phytologist* **139**: 505–516.

Forcella F, Benech Arnold RL, Sanchez R, Ghersa CM. 2000. Modeling seedling emergence. *Field Crop Research* **67**: 123–139.

Laperche A, Devienne-Barret F, Maury O, Le Gouis J, Ney B. 2006. A simplified conceptual model of carbon/nitrogen functioning for QTL analysis of winter wheat adaptation to nitrogen. *Theoretical and Applied Genetics* **113**: 1131–1146.

Messina CD, Jones JW, Boote KJ, Vallejos CE. 2006. A gene based model to simulate soybean development and yield responses to environment. *Crop Science* **46**: 456–466.

Moreau-Valancogne P, Coste F, Crozat Y, Dürr C. 2008. Field emergence of bean seed lots: characterisation, modelling and simulation. *European Journal of Agronomy* **28**: 309–320.

Mullins CE, Townend J, Mtakwa PW, *et al*. 1996. *EMERGE user guide: a model to predict crop emergence in the semi–arid tropics*. Department of Plant and Soil Science, University of Aberdeen, UK.

Passioura JB. 2010. Scaling up: the essence of effective agricultural research. *Functional Plant Biology* **37**: 585–591.

Quilot B, Génard M, Kervella J, Lescourret F. 2005*a*. Simulating genotypic variation of fruit quality in advanced peach × *Prunus davidiana* cross. *Journal of Experimental Botany* **56**: 3071–3081.

Quilot B, Kervella J, Génard M, Lescourret F. 2005*b*. Analysing the genetic control of peach fruit quality through an ecophysiological model combined with a QTL approach. *Journal of Experimental Botany* **422**: 3083–3092.

Rasmusson DC. 1987. An evaluation of ideotype breeding. *Crop Science* **27**: 1140–1146.

Reymond M, Muller B, Leonardi A, Charcosset A, Tardieu F. 2003. Combining quantitative trait loci analysis and an ecophysiological model to analyze the genetic variability of the responses of maize leaf growth to temperature and water deficit. *Plant Physiology* **131**: 664–675.

Ronfort J, Bataillon T, Santoni S, Delalande M, David JL, Prosperi JM. 2006. Microsatellite diversity and broad scale geographic structure in a model legume: building a set of nested core collection for studying naturally occurring variation in *Medicago truncatula*. *BMC Plant Biology* **6**: 28. doi:10.1186/1471-2229-6-28.

Sinha AK, Guildyal BP. 1979. Emergence force of crop seedlings. *Plant and Soil* **51**: 153–156.

Tamet V, Boiffin J, Dürr C, Souty N. 1996. Emergence and early growth of an epigeal seedling (*Daucus carota* L.): influence of soil temperature, sowing depth, soil crusting and seed weight. *Soil and Tillage Research* **40**: 25–38.

Smith J, Smith P, Addiscott T. 1996. Quantitative methods to evaluate and compare soil organic matter models. In: Powlson DS, Smith J, Smith P, eds. *Evaluation of soil organic matter models*. Berlin: Springer-Verlag, 181–199.

Voisin AS, Bourion V, Duc G, Salon C. 2007. Using an ecophysiological analysis to dissect genetic variability and to propose an ideotype for nitrogen nutrition in pea. *Annals of Botany* **100**: 1525–1536.

White JW, Hoogenboom G. 1996. Simulating effects of genes for physiological traits in a process-oriented crop model. *Agronomy Journal* **88**: 416–422.

Yin X, Struik PC. 2010. Modelling the crop: from system dynamics to systems biology. *Journal of Experimental Botany* **61**: 2171–2183.

Yin X, Kropff MJ, Stam P. 1999. The role of ecophysiological models in QTL analysis: the example of specific leaf area in barley. *Heredity* **82**: 415–421.

Yin X, Stam P, Kropff PS, Schapendonk AHCM. 2003. Crop modelling, QTL mapping and their complementary role in plant breeding. *Agronomy Journal* **95**: 90–98.