BIOSYSTEMS ENGINEERING 140 (2015) 34–47

Available online at www.sciencedirect.com

ScienceDirect

journal homepage: www.elsevier.com/locate/issn/15375110

Research Paper

Calibration of a growth model for tomato seedlings (TOMSEED) based on heuristic optimisation

Nikolaos Katsoulas [a,*], Konstantinos Peponakis [a],
Konstantinos P. Ferentinos [b], Constantinos Kittas [a,b]

[a] University of Thessaly, Department of Agriculture Crop Production and Rural Environment, Fytokou str., 38446, Volos, Greece
[b] Centre for Research and Technology – Hellas, Institute for Research and Technology of Thessaly, 95 Dimitriados str., 38333, Volos, Greece

ARTICLE INFO

Article history:
Received 9 May 2015
Received in revised form
24 August 2015
Accepted 16 September 2015
Published online xxx

Keywords:
Growth model
Tomato seedlings
Model calibration
Genetic algorithms
Simulated annealing
Tabu search

A mechanistic growth model for tomato seedlings cultivated in unheated beds is developed, based on modifications of existing tomato growth models. Photosynthetically active radiation at crop level, air temperature and CO_2 concentration are taken into account, while simulated variables include dry weights of leaves, shoot and root, leaf area index (LAI) of the seedlings, number of leaves per plant and finally, shoot length and thickness. Model calibration is formed into an optimisation problem, taking into account model errors of the first five simulated variables, i.e., dry weight of leaves, shoot and root, LAI and number of leaves per plant. Three heuristic optimisation algorithms are explored during model calibration: genetic algorithms, simulated annealing and tabu search. Genetic algorithms proved to be the most successful approach, resulting in an overall average deviation between simulated and measured values of around 16%. The calibrated model is tested and validated on measurements not used for calibration, showing a satisfactory performance in modelling most seedling characteristics, like the number of leaves per plant, the shoot length and thickness and the dry weight distribution, while it is not so accurate in predicting of other features like leaf and shoot dry weight and LAI. The tomato seedling characteristics that are mainly related to seedling quality (shoot length and dry weight, and number of leaves) were satisfactorily modelled with an average deviation between measured and simulated values around 16% for most of the simulation period. Finally, possible improvement strategies for future research are also discussed.

1. Introduction

The overall quality of seedlings of high-yield crops plays an important role in the growth rate of the plants after transplanting and therefore to the quality and quantity of production (Pavlovic, Petrovic, & Stevanovic, 1998). Modern greenhouse cultivation and commercial trends call for uniform and standard quality seedlings, regardless of season. Seedling cultivation should therefore be scheduled and

---

* *Corresponding author*. Tel.: +30 24210 93249.
E-mail address: nkatsoul@uth.gr (N. Katsoulas).

http://dx.doi.org/10.1016/j.biosystemseng.2015.09.004

Nomenclature

$[CO_2]$	CO_2 concentration of the air, ppm
ASR_{leaf}	assimilate requirements of the leaf, g (CH_2O) g^{-1} (DM)
ASR_{root}	assimilate requirements of the root, g (CH_2O) g^{-1} (DM)
ASR_{shoot}	assimilate requirements of the shoot, g (CH_2O) g^{-1} (DM)
C_i	CO_2 intercellular concentration, ppm
D_{root}	increase demands of the roots, g CH_2O
D_{shoot}	increase demands of the shoot, g CH_2O
D_{total}	total increase demands of all organs, g CH_2O
DM	dry matter, g
DM_{leaf}	leaf dry matter, g
DM_{root}	root dry matter, g
DM_{shoot}	shoot dry matter, g
DWLS	dry weight per length unit of the shoot, g cm^{-1}
e_m	photochemical efficiency, μmol (CO_2) μmol^{-1} (photon)
e_O	potential photochemical efficiency, μmol (CO_2) μmol^{-1}(photon)
fasm	coefficient that incorporates the effect of sugars availability on the development of new leaves, –
fT	coefficient that incorporates the effect of lower than optimal temperatures on the development of new leaves, –
GA	genetic algorithm, –
I_O	radiation intensity above canopy level, μmol m^{-2} (ground) s^{-1}
I_{abs}	radiation absorbed by the canopy, μmol m^{-2} (ground) s^{-1}
I_{absSum}	accumulated amount of absorbed solar radiation during the entire cultivation period, MJ m^{-2}
k	light extinction coefficient, –
K_c	CO_2 Michaelis–Menten constant, ppm
K_o	O_2 Michaelis–Menten constant, ppm
L	canopy depth as a function of the LAI of the overlying canopy, –
LA	leaf area of a single plant, mm^2
LAI	leaf area index, –
Lsm	minimum specific leaf area, m^2 g^{-1}
Lss	maximum specific leaf area, m^2 g^{-1}
MAINT	demands in CH_2O for maintenance of 1 g of dry matter, g (CH_2O) g^{-1} (DM) day^{-1}
ME	model error, –
minr	percentage of available sugars that will be distributed to the root, –
N_{leaves}	number of leaves per plant, –
O_α	ambient O_2 concentration, ppm
Pl	possible increase rate of leaf area, as a percentage of the existing leaf area, –
Pool	quantity of available sugars in the “pool”, g m^{-2}
PoolLimit	capacity (limit) of the “pool”, g m^{-2}
Pr	possible increase rate of the root, as a percentage of existing root weight, –
pSLA	specific leaf area, m^2 g^{-1}
$Q_{10(\tau)}$	changing factor of τ for a 10 °C temperature change, –
$Q_{10,resp}$	changing factor of respiration for a 10 °C temperature change, –
r_m	maintenance respiration rate, $g(CH_2O)$ day^{-1}
rLA	increase rate of leaf area, as a percentage of the existing leaf area, –
rM	maximum rate of new leaves development, leaves $plant^{-1}$ day^{-1}
rN	rate of development of new leaves, day^{-1}
rRoot	increase rate of the root, as a percentage of existing root weight, –
SA	simulated annealing, –
SL	shoot length, cm
ST	shoot thickness, mm
T_{leaf}	leaf temperature, °C
Tl	coefficient for influence of temperature on the increase rate of leaf area, –
Tr	coefficient for influence of substrate temperature on root growth, –
TS	tabu search, –
v_c	CO_2 absorption rate, limited by Rubisco activity, μmol (CO_2) m^{-2} s^{-1}
v_q	CO_2 absorption rate, limited by light intensity, μmol (CO_2) m^{-2} s^{-1}
$V_{c,max}$	maximum CO_2 absorption rate from Rubisco activity, μmol (CO_2) m^{-2} s^{-1}
Γ^*	CO_2 equilibrium point, ppm
ΔDM_{root}	root dry matter accumulation, g m^{-2}
ΔDM_{shoot}	shoot dry matter accumulation, g m^{-2}
ΔDM_{leaf}	leaves dry matter accumulation, g m^{-2}
ΔLAI	increase of leaf area index, –
ρ	radiation reflectance coefficient –
τ	CO_2/O_2 specificity factor, –

optimised by means of controlling plant growth through greenhouse microclimate management. This is achieved by specialised control systems that manage the basic control actions of greenhouse seedling cultivation, like heating, ventilation and CO_2 enrichment. The operation of such systems can be improved by means of precise greenhouse and crop models, so that their control actions can be optimised by accurate environmental and plant growth predictions. Several greenhouse climate and crop models have been developed but only a few models concern seedlings.

Nowadays, specialised nurseries for the development of high quality seedlings (the plant characteristics relevant to quality are given later in this section) have started to operate and expand, trying to fulfil growers' demand for high quality transplants, deliverable on specific dates. The most important factor in the operation of these nurseries is the appropriate control of the microclimatic conditions that would lead to high quality seedlings. Optimal control practices should rely on Decision Support Systems (DSS), which in turn rely on greenhouse climate models, crop growth models, weather forecasts and cost estimation models. Despite the development of tomato crop models, such as TOMGRO (Jones, Dayan, Allen, van Keulen, & Challa, 1991), TOMSIM (Heuvelink, 1996) and others, only a few tomato seedling growth models have existed until recently. Among them are CROPGRO-TOMATO (Scholberg, Boote, Jones, & McNeal, 1997) and GreenLab (Yan, Kang, de Reffye, & Dingkuhn, 2004), which are specialisations of other, more general crop growth models. However, to the best of our knowledge, all tomato growth models focus on the reproductive stage of cultivation, in which there is concurrent development of shoot and reproductive organs so they cannot be used for seedling modelling, since several differences exist between seedlings and crops in the production stage [*e.g.*, a seedling does not have reproductive organs and all organs are young; seedling growth rate is exponential while that of a productive crop is linear (Goudriaan and van Laar, 1994), etc].

For the production of high quality tomato seedlings, the characteristics relevant to their quality and sale prices have to be initially determined. Some of those features that are closely related to seedling quality have been empirically determined in the literature and some relevant indicators have been defined. However, their optimal values have not been investigated, so there is no clear relation between indicator values and quality level of the seedlings. For the seedling state evaluation, both morphological and physiological indicators have been used. The first category includes height, shoot diameter and shoot-to-root weight ratio, while the second includes tissue electrolyte leakage, enzymatic activity, water potential and mineral nutrition. Radoglou and Raftoyannis (2002) investigated the potential of using several physiological indicators (like moisture content in shoot and root, shoot water capacity and tissue electrolyte leakage) to predict future quality and growth performance of the cultivated plants. Their results were quite promising. Hoek, Ten Cate, Keijzer, Schel, and Dons (1993) concluded that the development of the 5th leaf expressed as leaf area, can constitute a reliable indication of tomato growth at low temperatures. Pucheta, Schugurensky, Fullana, Patiño, and Kuchen (2006) tackled the problem of optimal control of a tomato nursery so that seedling production with specific desired characteristics could be achieved in the most economical way. The characteristics that were taken into account were dry weight and number of leaves, with desired values of 0.21 g and 3 leaves, respectively. Andersen (1986) used quality indicators for tomato seedlings based on dry weight to height and dry weight to fresh weight ratios, to assess supplemental lighting methodologies. Herrera, Castillo, Chica, & López Bellido (2008) used an extensive number of indicators to estimate tomato seedling quality, like height, shoot thickness, height to thickness ratio, dry weight of leaves, shoots and roots, number of leaves, leaf area, leaf area index (LAI) and leaf area to dry weight ratio. Morimoto and Hashimoto (1996) used the total leaf length to shoot diameter ratio as an index of balanced growth for tomato seedlings. That ratio has been recognised as a good indicator for growth prediction of reproductive organs and shoots, with higher values being associated with higher growth. Finally, Gupta, Chandra, Samuel, and Singh (2012) developed a mathematical model to correlate tomato seed germination with soilless medium temperature and used it to develop a decision support system to attain the desired dry weight of tomato seedlings at the time of transplanting (Gupta, Samuel, & Sirohi, 2010).

Thus, it can be concluded that a combination of indicators should be used to better estimate the overall quality of the seedlings and that the characteristics that have the strongest correlation with quality are the dry weight of roots, shoots and leaves and also the plant leaf area. Accordingly, the aim of this work is to present a tomato seedling growth model (TOMSEED) that has been developed based on appropriately modified features of the TOMGRO model (Jones et al., 1991), and to optimally calibrate it and evaluate its performance with experimental measurements.

2. The TOMSEED model

During the period from germination to transplanting, only vegetative growth takes place in seedlings. Although this facilitates their growth simulation, several modelling challenges remain, mainly because of the changing characteristics of the growing seedlings. Initially, there is no inter-leaf overlapping and growth rate is nearly proportional to the leaf area, resulting in an exponential growth rate. However, as leaf area increases and leaves start to overlap, growth rate converges to a constant value, thus growth becomes independent of the leaf area. Additionally, it has been reported that seedling growth depends on the air temperature, with the optimal night temperature during the first month being 4 °C lower than that for the rest of the growing period (Seginer & Raviv, 1984).

The proposed growth model, TOMSEED, is a mechanistic model that simulates the growth of tomato seedlings cultivated on unheated beds, as a function of photosynthetically active radiation (PAR) at crop level, air temperature and CO_2 concentration. Simulated variables include leaf, shoot and root dry weight (per unit ground area), LAI of the seedlings, number of leaves per plant and shoot thickness and length.

The model consists of two loops, following the TOMGRO approach. The first loop performs the time steps of the simulation, which are subdivisions of the day. Even though the number of time steps is user-defined, it is suggested that each step's duration is not less than one hour (*i.e.*, a maximum of 24 time steps). At each time step, photosynthesis and transpiration are estimated, while the products of pure photosynthesis are stored. The second loop is executed once at the end of each day. Potential growth rates for the environmental conditions of the specific day are estimated and photosynthetic products are distributed to the existing organs. When the produced metabolites can sufficiently cover total needs, their distribution is based purely on the demand of each organ, while, in the case where the production cannot cover the overall demand, the distribution is based on the priority of each organ.

The initial point of model application is one week after emergence. At that point, the first real leaf has appeared and in some plants even the second leaf has started to appear. Initial variable values are given to the model by an external file and they include LAI, dry weight of leaves, shoots and roots per m^2, number of leaves per plant and density of the seedlings. Cotyledons are considered to operate as normal leaves, thus they are included in the calculation of LAI, however they are not included in the count of the number of leaves per plant. Details of the two loops of the model are presented in Appendix A.

3. Materials and methods

3.1. *Greenhouse facilities and plant material*

The experiments were performed in an arch type greenhouse, N–S oriented (36° declination from north 0°), located at the University of Thessaly near Volos (latitude 39° 44′, longitude 22° 79′, altitude 85 m) on the continental area of eastern Greece. The greenhouse was covered by a double inflated polyethylene film for the roof and by glass for the sidewalls and gables. The geometrical characteristics of the greenhouse were as follows: eaves height, 3 m; ridge height, 4.6 m; total width, 10 m; total length, 30 m. The greenhouse was equipped with two side flap vents and a roof vent. Greenhouse heating was by means of a fan heater and plastic pipes located above the ground. The heating was controlled by an on-off controller following the daytime and night-time set point.

The experiments were conducted during three periods between November 2007 and May 2008. The starting date of each period was taken as the date when 75% of seedlings had emerged. The temperature regime differed from period to period, and the mean air temperature values for daytime and night-time were as follows:

- Period 1 (10 November 2007 to 15 December 2007): 20.5 ± 0.7 °C and 14.8 ± 0.4 °C
- Period 2 (21 December 2007 to 30 January 2008): 24.6 ± 0.9 °C and 19.8 ± 0.6 °C
- Period 3 (31 May 2008 to 5 July 2008): 26.9 ± 1.2 °C and 21.4 ± 0.9 °C

for daytime and night-time respectively. The mean temperature difference between day and night-time temperature was about 5 °C.

Three different light treatments were tested for each period of measurement in three different parts of the greenhouse: a) reduced lighting, b) natural lighting and c) supplemental lighting. Reduced lighting was achieved by extending a shading net above the plants during the total period of measurements resulting in about 10% shading, while supplemental lightning was implemented using HPS lamps (Master GreenPower 600W EL 400V Mogul 1SL, Philips). The lamps were: (a) turned on whenever the solar radiation intensity in the control treatment was lower than 200 W m^{-2} and (b) switched off if the solar radiation integral in the supplemental lighting was more than 21% greater than that in the control treatment. The first condition was necessary to ensure that supplemental lighting is provided when solar radiation intensity is a limiting factor for plant development while the second one was used to limit supplemental lighting to a certain extent. During the three periods, the total light energy integrals were:

- Period 1 (10 November 2007 to 15 December 2007): 230 MJ m^{-2}, 177 MJ m^{-2} and 160 MJ m^{-2},
- Period 2 (21 December 2007 to 30 January 2008): 220 MJ m^{-2}, 171 MJ m^{-2}, 155 MJ m^{-2} and
- Period 3 (31 May 2008 to 5 July 2008): 403 MJ m^{-2}, 400 MJ m^{-2} and 368 MJ m^{-2}, for supplemental lighting, natural lighting and reduced lighting treatments, respectively.

It should be noted that during experimental period 3, which took place during summer, photoinhibition problems seem to have been induced in the seedlings due to the high levels of solar radiation and thus the measurements were stopped nearly 3 weeks after sowing. The high levels of solar radiation occurring during the last experimental period resulted also in only small differences between supplemental lighting and the control treatment, due to limitations in supplemental lighting from condition (a) of lamps' control described above.

The tomato seedlings (*Lycopersicon esculentum*, cv. Belladona) were grown in 150 cell plates, 0.8 m by 0.6 m. The resulting density of the seedlings was 312 plants m^{-2}. Irrigation and fertilisation were applied according to the practice followed by the growers of the region. Water and fertilisers were supplied by a sprinkler irrigation system, which was automatically controlled by a fertigation computer.

3.2. *Measurements*

The following microclimate data were recorded in each treatment during the period of measurements: Air temperature and relative humidity by means of temperature and relative humidity sensors (model H08-032-08, HOBO instruments, Bourne, MA, U.S.A.), total solar radiation by means of solar pyranometers (model CM-6, Kipp and Zonen, Delft, The Netherlands), and CO_2 concentration by means of a CO_2 sensor (model WMA-4 CO_2 gas analyser, PP Systems International, Inc., MA, USA).

All the measurements were collected by a data logger system (model DL3000, Delta-T Devices, Cambridge, U.K.). Measurements took place every 30 s and 10-min average values were recorded.

Plant destructive measurements were carried out every week, starting from the first week after 75% of the seedlings had emerged, and finishing when most plants had five true leaves, which usually took 4–5 weeks. Each week, 25 plants were randomly chosen in each treatment. After measuring the leaf area, the number of leaves per plant and the shoot length and thickness, the different organs (stems, leaves and roots) were dried for 48 h at 85 °C, in order to measure their dry weight. Sampled plants where replaced by plants of the same age and growth stage, grown under the same conditions as the sampled ones.

3.3. *Model calibration and validation*

As analysed in Appendix A, where the details of the TOMSEED model operation are presented, the proposed model contains several constant parameters with unknown values that have to be calibrated. Thus, model calibration consisted of fine-tuning the following ten basic model parameters:

- *Lss* (maximum specific leaf area) and *Lsm* (minimum specific leaf area) parameters were selected for calibration because experimental measurements showed that the specific leaf area of the seedlings was around 0.1, which is much greater than the reported value for mature plants.
- The $Q_{10,resp}$ (factor for change in respiration for a 10 °C temperature change) parameter was selected because its reported values in the literature vary over quite a wide range, from 1.4 to 2.
- Pl_0 and *lgf* (empirical coefficients used for the computation of possible increase rate of leaf area), Pr_0 and *rgf* (empirical coefficients used for the computation of increase rate of the root weight), *minr* (percentage of available sugars that will be distributed to the root), *rM* (maximum rate of new leaf development) and *GENFAC* (coefficient of the effect of sugars availability to the development of new leaves), were also selected because there are no specific suggestions for their values in the literature, and even though some of them can be extracted from existing tomato growth models, their values do not seem to be adequate for the case of seedling cultivation.

It should be noted that (1) *Lss* and *Lsm* parameters both refer to specific leaf area (*pSLA*) computation, (2) $Q_{10,resp}$ refers to maintenance respiration (r_m) computation, (3) Pl_0 and *lgf* both refer to the computation of increase rate of leaf area (*rLA*), (4) Pr_0 and *rgf* refer to the computation of increase rate of the root weight (*rRoot*), (5) *minr* refers to the computation of root dry matter accumulation (ΔDM_{root}), and (6) *rM* and *GENFAC* refer to the computation of rate of development of new leaves (*rN*). Thus, although ten model parameters were selected to be estimated, only six plant functions were in fact calibrated.

Dry weight of leaves, shoots and roots, as well as LAI, number of leaves per plant and shoot length and thickness have been chosen for simulation, since these characteristics give satisfactory indication of the growth as well as the quality of tomato seedlings.

The calibration problem was formed into an optimisation task with a goal to minimise the model error when compared to real measurements. Because of the high number of exploration parameters and the non-continuous nature of the modelling error function to be optimised, the optimal calibration was tackled using heuristic and evolutionary optimisation algorithms, which have the advantage of having good performance in complicated, real-life applications with numerous parameters and complex search spaces. Several heuristics were developed for the calibration process, based on genetic algorithms (Holland, 1975; Goldberg, 1989), simulated annealing (Kirkpatrick, Gelatt, & Vecchi, 1983) and tabu search (Glover, 1989) methodologies.

3.3.1. *Heuristic algorithms*

Genetic algorithms (GAs) (Holland, 1975; Goldberg, 1989) belong to a class of algorithms known as evolutionary computation. They imitate the process of natural evolution by assigning fitness values to possible solutions of the problem and applying a mathematical model of the Darwinian principle of survival of the fittest. A number of possible solutions of the problem evolve during several generations, through the application of specific genetic operators. Each individual (or chromosome, in the GA terminology) consists of a string of values (genes) for the set of parameters that formulate the problem to be solved or the system to be optimised. Through the genetic operators, the individuals are combined and transformed, and through some selection mechanism, which is based on the Darwinian principle of the “survival of the fittest”, the best ones survive and reproduce to form even better solutions.

Simulated annealing (SA) (Kirkpatric et al., 1983) is based on ideas first presented by Metropolis, Rosenbluth, Rosenbluth, Teller, and Teller (1953). Metropolis's algorithm simulates the change in energy of a system when subjected to a cooling process, until it converges to a steady, “frozen” state. The SA algorithm that is inspired by the physical process of annealing is typically implemented by initialising the tuning features of the algorithm (i.e., initial random set of parameter values, initial temperature and cooling rate) and calculating the cost of the initial set of parameters. Then, a random new set of parameters is chosen and its cost is evaluated and if this new set is an improvement, it is accepted, otherwise it is accepted with probability $e^{-\Delta C/T}$, where ΔC is the cost difference and T is the current temperature. This process repeats and when the maximum number of “constant-temperature repetitions” is reached, the temperature is decreased according to some “cooling schedule” and the entire process is repeated until a sufficient solution is found or the maximum number of iterations is reached.

Tabu search (TS) (Glover, 1989) is a powerful heuristic optimisation methodology, which is basically a local search strategy with a flexible memory structure. Its name comes from its basic feature, which is a list of solution points that must be avoided, i.e., they are not allowed to be accepted as possible solutions, leading to the exploration of new areas in the search space. This list is referred to as the “tabu list” and it is updated based on some memory structure (short-term memory). At each iteration of the algorithm, the best of a set of solutions neighbouring the current solution is chosen, unless it belongs to the tabu list. Another important feature of the algorithm is that at each iteration, the best neighbouring solution that does not belong to the tabu list will necessarily be selected, even if it is worse than the current solution. Finally, some “aspiration criteria” exist, which allow for exceptions from the tabu list, if such moves lead to promising solutions. The algorithm has several tuning features that are application-specific and have to be explored, like the size of tabu tenure, the type of tabu restriction, the size of the neighbourhood, etc.

3.3.2. *Calibration approach*

Each possible solution to the calibration problem consisted of a set of values for each of the ten calibration parameters. In heuristic optimisation, each solution must have a quality metric, usually referred to as “fitness” of the solution, which is estimated by an appropriate *fitness function*. In all three algorithms, the same fitness function was used, which was based on the sum of the absolute errors of the main model's outputs, compared to actual, measured values. The comparisons for the estimation of the model error were made at several time points that corresponded to specific measurement points of leaf (DM_{leaf}), shoot (DM_{shoot}) and root (DM_{root}) dry weights, leaf area index (LAI) and number of leaves per plant (N_{leaves}). Thus, at each specific time point that morphological measurements existed, the model error (ME) was estimated by the following equation:

$$ME = \frac{\left|DM^{s}_{leaf} - DM^{m}_{leaf}\right|}{DM^{m}_{leaf}} + \frac{\left|DM^{s}_{shoot} - DM^{m}_{shoot}\right|}{DM^{m}_{shoot}} + \frac{\left|DM^{s}_{root} - DM^{m}_{root}\right|}{DM^{m}_{root}} + \frac{\left|LAI^{s} - LAI^{m}\right|}{LAI^{m}} + \frac{\left|N^{s}_{leaves} - N^{m}_{leaves}\right|}{N^{m}_{leaves}}$$

that is, the sum of all normalised absolute differences between simulated values (denoted by the "s" superscript) and measured values (denoted by the "m" superscript) of each variable. For a total of M available measurement points (during all available cultivation periods), the fitness f of a specific solution is given by:

$$f = \frac{1}{\sum_{i=1}^{M} ME_i}$$

The goal of each algorithm was to maximise the value of f, that is, to find the solution with the maximum possible fitness value, which would be the optimal solution leading to the smallest possible model error.

In all three optimisation algorithms used for the calibration of TOMSEED, the same fitness function was used to qualify each possible solution, that is, each combination of model parameter values. GAs work with a population of possible solutions that evolve through iterations (namely, generations), while the other two algorithms are iterative too, but not population-based. Another common characteristic of all three algorithms is that their performance depends on the randomly selected starting solutions, thus they have to be repeated several times from different starting positions. Finally, they all have several tuning features that have to be explored and defined through experimentation, as these features are application-specific and their optimal values cannot be defined *a priori*. In GAs, these features are the population size, the percentages of crossover and mutation occurrence and the specific types of crossover and mutation. In SA, the basic tuning features are the initial value of the temperature factor and its decrease scheme. In TS, the features that have to be tuned are the size of the tabu list, its tenure, the size of the neighbourhood of the solutions and the definition of the aspiration criteria that allow exceptions in the tabu list.

Finally, for all algorithms, a stopping criterion has to be defined, as there is no way of knowing whether a specific solution is the global optimum. The general criterion (that was not satisfied in practice) was an average overall model error of 5%. Because of the different processes involved in the three algorithms, a criterion based on a maximum number of iterations of the algorithms would not be fair for algorithm comparison, thus a stopping criterion based on a maximum number of fitness function evaluations was used instead.

Several experiments were performed for each algorithm, in order to fine-tune the corresponding optimisation features.

3.3.3. *Validation approach*

The validation of the model was performed using climatic and morphological data from experimental periods that were not used during the calibration process of the model and were intentionally kept for the process of model validation. More specifically, those periods were:

- the natural lighting (control) treatment of experimental period 2 (21 December 2007 to 30 January 2008, average daytime temperature of 24.6 ± 0.9 °C, night-time temperature of 19.8 ± 0.6 °C, total solar radiation integral of 170.9 MJ m^{-2}) and,
- the reduced lighting (shading) treatment of experimental period 3 (31 May 2008 to 5 July 2008, average daytime temperature of 26.9 ± 1.2 °C, night-time temperature of 21.4 ± 0.9 °C, total solar radiation integral of 368 MJ m^{-2}).

The calibrated model was applied to these climatic data and its outputs (dry weight of leaves, shoot and root, LAI, number of leaves, shoot length and shoot thickness) were compared with the available measured values.

4. Results

4.1. *Calibration results*

Table 1 presents the final performance results of the three algorithms during model calibration. Genetic Algorithms (GAs) proved to be more successful in the model calibration process, followed by Simulated Annealing algorithm (SA), while Tabu Search algorithm (TS) was the least successful algorithm. Table 2 shows the optimal values of the calibration parameters of the model, which were found using the GA methodology. The average overall model error with these values was 16.1% and was achieved with a GA population of 10,000 chromosomes, probabilities of crossover and mutation equal to 0.8 and 0.1, respectively, and simple arithmetic crossover type. The evolution progress during model calibration for the best GA run is shown in Fig. 1, where both the fitness progress of the best solution found by the algorithm, as well as the average fitness of the entire GA population, are plotted. The total running time of the algorithm for the 100 iterations (1 million fitness function evaluations) was about 13 min on a single core of an i5 quad-core processor at 3.2 GHz.

Figure 2 shows the model predictions and the actual measurements with one standard deviation error bars, for all seven modelled variables during one of the experimental periods (Period 1, natural lighting) used in the calibration process. It can be seen that the model successfully simulates all variables and in most cases, even 35 days after germination,

Table 1 – Performance of the three algorithms during model calibration. Errors are represented as average percentage deviations from real measurements over all model outputs.

Algorithm	Average error (%)	Standard deviation of error	Minimum error (%)
GA	19.23	2.45	16.05
SA	26.37	3.65	20.06
TS	29.77	3.13	25.54

Table 2 – Optimal values of calibrated model constant parameters.

Model parameter	Optimal value	Unit
Lss	0.086	$m^2\ g^{-1}$
Lsm	0.0336	$m^2\ g^{-1}$
$Q_{10,resp}$	1.86	–
Pl_0	0.0074	–
lgf	0.7949	–
Pr_0	0.1543	–
rgf	0.3734	–
rM	0.16	Leaves plant^{-1} day^{-1}
minr	0.028	–
GENFAC	1.09	–

the predicted values are very similar to the measured ones. The increase trends are accurately followed in most cases, with the exception of root dry weight (Fig. 2c), which is overestimated between days 15–30. In general, there is very little deviation between measured and simulated values.

4.2. *Model validation*

As described in Section 3.3.3, the calibrated model was tested and validated using different data than those used for calibration. The performance of the GA-calibrated model in each of the output variables is presented hereafter.

Figure 3 shows the model predictions and the actual measurements for all seven modelled variables during experimental period 2 with natural lighting conditions (first validation period). It can be seen that, with the exception of the first two weeks, the model underestimates the values of shoot dry weight (Fig. 3b), and LAI (Fig. 3d), although the general trend is similar to the actual one, and also the variability of measurements drastically increases with time, thus the modelling error is expected to be higher. The maximum deviation between predicted and actual measurements reached 45%. Predicted values of leaf (Fig. 3a) and root (Fig. 3c) dry weight were quite satisfactory, as they deviated by around 35% during the first days, but this difference dropped to 2–8% for the rest of the testing period. In addition, measured and predicted values had very similar trends. The number of leaves per plant (Fig. 3e) is a characteristic which is also modelled quite well with slightly overestimated values in general, except for the final ones. The simulation of shoot length is very accurate (Fig. 3f), as the predicted values up to 21 days after germination are almost identical to the measured ones, and for the rest of the testing period, their deviation does not exceed 10%. Finally, the shoot thickness is simulated satisfactorily (Fig. 3g), as, with the exception of the first measurement, the deviation between predicted and measured values does not exceed 12%.

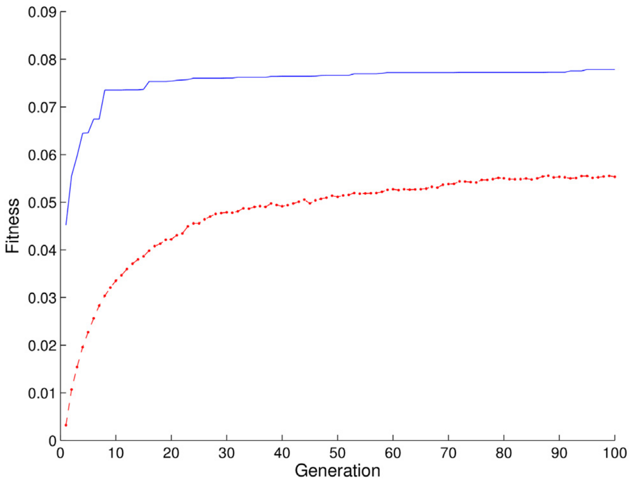

Fig. 1 – Evolution progress of the best solution, i.e. the best fitness value (—) and the entire population, i.e. the average fitness value (-·-) of the GA during the best run of the model calibration process.

Figure 4 shows the model predictions and the actual measurements for all seven modelled variables during experimental period 3 with reduced lighting (shading) conditions (second validation period).

It should be noted that although the set of measured data shown in Fig. 4 corresponds to shading conditions, due to the fact that period 3 was carried out during summer, the solar radiation integral of reduced lighting treatment of period 3 (368 MJ m^{-2}) was higher than the solar radiation integral of periods 1 and 2. In addition, the average air temperature values observed during period 3 were also higher than those observed during periods 1 and 2.

For leaf dry weight (Fig. 4a), the performance of the model is poorer than that during the first validation period (Fig. 3a), as there is significant overestimation which increases with time. The deviation between predicted values and actual measurements of shoot dry weight (Fig. 4b) is limited to a maximum of 15%, which is much better than that during the first validation period. In the case of root dry weight (Fig. 4c), model performance is very poor, as the deviation between predicted and actual values reaches a maximum of 92% and it only drops down to 10% at the end of the testing period. LAI, on the other hand, is better simulated (Fig. 4d), with a difference between predicted and measured values in the range of 20%–70%, with the exception of the second value, where the difference is just 1.7%. The simulation of the number of leaves is quite successful (Fig. 4e). The values are constantly slightly overestimated, but with predicted values very close to the measured ones. Shoot length is accurately predicted up to day 14 (Fig. 4f), but after that the deviation between predicted and measured values increases to 74% and finally drops to 28%. The model error increases after the second measurement due to the small actual increase of the shoot length and dry weight during that period. The problematic prediction of the shoot dry weight during these days (Fig. 4b) led to the relative large errors in shoot length simulation. It can be seen that during the entire period, shoot length is overestimated by the model. The same is true for the shoot thickness (Fig. 4g), where the model constantly overestimates the actual values. However, the increase rate of the predicted shoot thickness is the same as the measured one.

Thus, it can be seen that although the model performs quite well during the first validation period, it mainly overestimates the actual values during the second validation period where data from a summer experiment (with higher

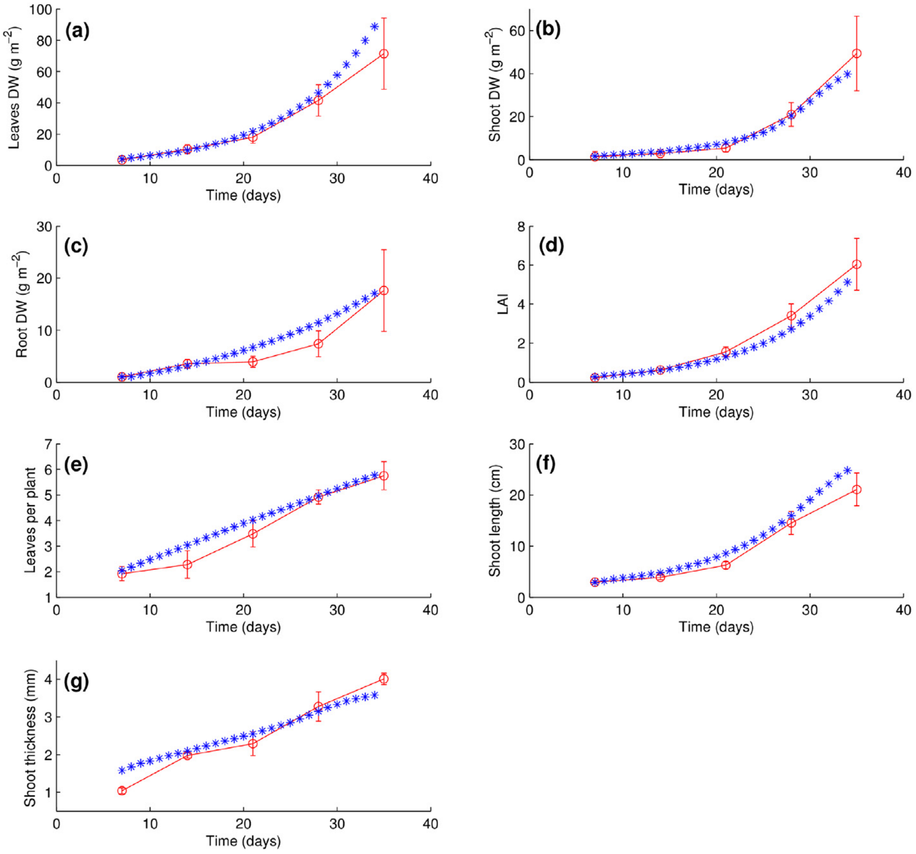

Fig. 2 – Calibrated model predictions (*) and actual measurements (—○—) during an experimental period used for calibration (period 1), with natural lighting. Time is in days after germination. (a) leaf dry weight, (b) shoot dry weight, (c) root dry weight, (d) LAI, (e) leaves per plant, (f) shoot length, (g) shoot thickness. The error bars represent the standard deviation of the measured values.

values of solar radiation integrals and average air temperature compared to the rest of the experimental periods available) were used.

5. Discussion

5.1. *Calibration process evaluation*

The proposed growth model for tomato seedlings incorporates the features that play an important role in seedling quality. The complex process of calibrating a number of model parameters was tackled using powerful heuristic optimisation algorithms. These algorithms are application-specific, thus their final performance and exact suitability cannot be known *a priori* (Michalewicz & Fogel, 2004). Furthermore, heuristic optimisation methods are application-specific and act as "black-boxes" in most cases, thus it is not possible to justify their relative performance based on the characteristics of the problem. The superiority of genetic algorithms in comparison to the other two algorithms (simulated annealing and tabu search) may be due to its population-based nature, where a large number of possible solutions evolve, instead of a single solution that is used by the other two algorithms. The structure of the specific optimisation problem seems to favour that approach.

The actual performance of the model was validated using two sets data, different from those used for calibration. The presented results of the validation data are promising. Not all model variables were accurately simulated, with leaf and shoot dry weight giving the worst prediction results. Similar problems in leaf and shoot dry weight prediction have also been observed by other authors working with tomato crop modelling (*e.g.* Dayan et al., 1993b; Dimokas et al., 2009).

As noted in Section 3.3.3, the model was validated using data from two experimental periods, a winter (21 December 2007 to 30 January 2008) and a summer (31 May 2008 to 5 July 2008) one. It could be seen that although the model performed quite well during the first validation period (Fig. 3), it mainly overestimated the simulated variables during the summer experimental period (Fig. 4), where relatively high values of solar radiation intensity and air temperature were observed. This may relate to the absence from the TOMSEED model of any function to take into account for the effect of such non-optimal climate conditions on seedling development and growth.

Vanthoor, De Visser, Stanghellini, and Van Henten (2011) reviewed the effects of temperature on a tomato crop, summarising that: (a) both instantaneous and mean temperatures

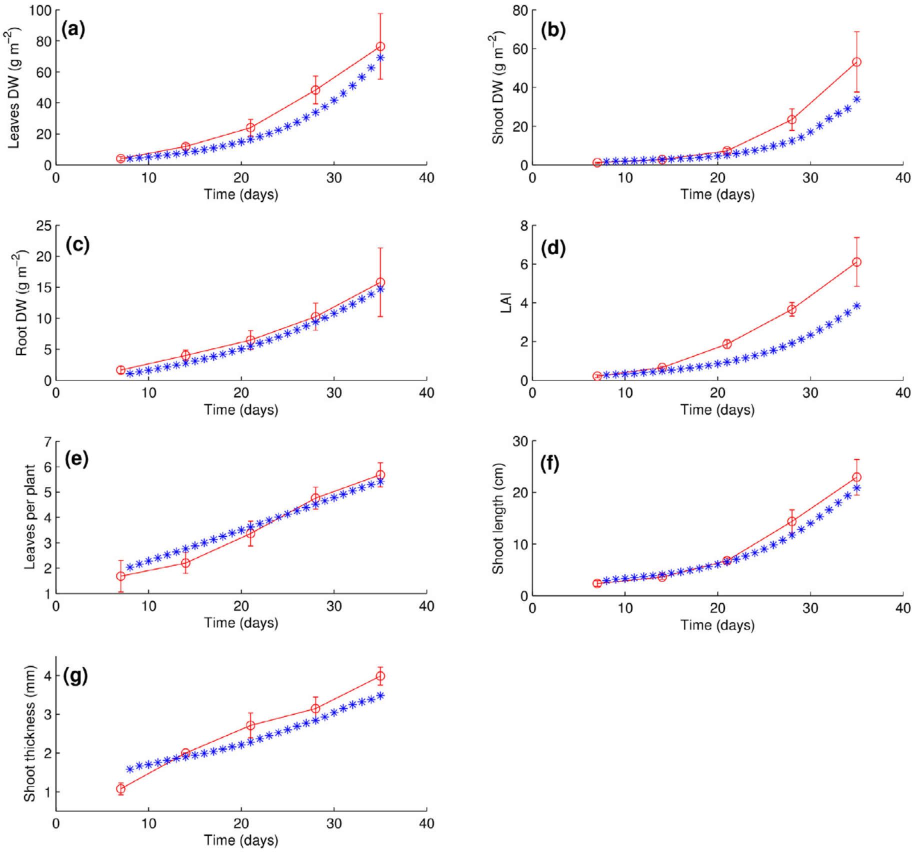

Fig. 3 – Model simulated (∗) and actual values (—○—) during validation in natural lighting conditions (period 2). (a) leaf dry weight, (b) shoot dry weight, (c) root dry weight, (d) LAI, (e) leaves per plant, (f) shoot length, (g) shoot thickness. The error bars represent the standard deviation of the measured values.

affect crop yield, (b) both sub-and supra-optimal temperatures affect several growth processes, resulting in lower yield, (c) it is difficult to identify one single growth process causing crop stress because growth processes influence each other, (d) stress sensitivity is cultivar-dependent. Thus, in the tomato yield model they developed, periods of unfavourable temperatures were taken into account in two ways: In the first place sub- and supra-optimal temperatures result in a decrement of actual photosynthesis in comparison to maximal photosynthesis under the given light and CO_2-conditions. In the second place temperature also affects the ability of organs to store and release assimilates. This is accounted for by functions for tomato growth inhibition at non-optimal temperatures, taking into account the difference between instantaneous and 24-h average temperatures. Thus, after an extensive review, they suggested that the optimal boundary for instantaneous values could be taken at 28 °C and for the 24 h mean could be taken at 22 °C.

Accordingly, taking into account the findings and suggestions given by Vanthoor et al. (2011), it seems that the air temperature values observed during the summer period in the present work (Period 3) were higher than the limits they indicated, inducing stress conditions that were not taken into account by our model. This could possibly explain the overestimation of some variables by TOMSEED model during the summer period.

5.2. *Model evaluation*

Considering that the proposed model is a first attempt to model tomato seedlings growth, its overall performance is regarded sufficient as a basis upon which, more accurate and complex models can be built. By incorporating additional processes and crop characteristics, the applicability of the TOMSEED model could be extended and the quality of the results improved.

Some assumption modifications and additions could improve the final performance of the model and will be investigated as a next step to TOMSEED development. The current model assumes that intercellular CO_2 concentration is 70% of the ambient concentration. However that value increases non-linearly when CO_2 enrichment takes place (Morison, 1987), thus an empirical equation or a special sub-model have to be used, something that was not feasible in this work, as there was no CO_2 enrichment during the experiments. Another drawback of the model was the use of not so accurate estimates of leaf temperature, especially during summer months (see Appendix Eq. (5)). This could be overcome in a future version of the model with the use of infrared sensors for the accurate measurement of leaf temperatures.

In addition, there is an obvious correlation between overestimation versus underestimation of the actual

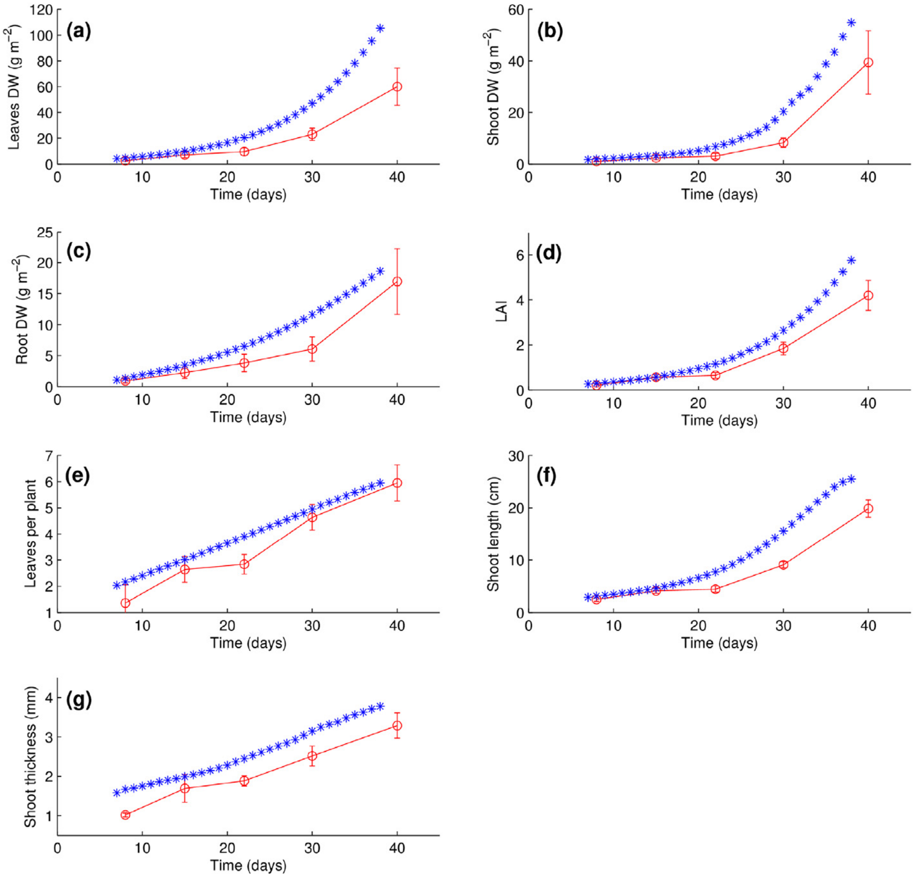

Fig. 4 – Model simulated (∗) and actual values (—○—) during validation in reduced lighting conditions (shading) (period 3). (a) leaf dry weight, (b) shoot dry weight, (c) root dry weight, (d) LAI, (e) leaves per plant, (f) shoot length, (g) shoot thickness. The error bars represent the standard deviation of the measured values.

measurements and the existing lighting and temperature conditions during seedling cultivation. Thus, the way that light intensity and air temperature are incorporated into the model has to be further studied and analysed, so that its outputs can be better adjusted to actual lighting conditions.

Furthermore, additions to the model, like the effects of high light intensity and air temperature, such as the one introduced by Vanthoor et al. (2011) to take into account supra-optimal conditions, would be useful and will be investigated for inclusion in the next version of TOMSEED.

Another area for development could be root growth which is treated in the simulation in a rather rudimentary fashion. A more detailed description is required that allows evaluation of the capacity of the root system to supply the required water and nutrients for optimum crop functioning.

Connecting the biological model with a climate model could allow the production process to be optimised. Although optimisation is rarely used for crop development (Garcia, 1999), it can be applied for greenhouse climate and crop production to provide the best strategy according to growers needs. Optimisation would allow the best management strategy to be determined for a range of possible future climates to which the greenhouse may be subjected. This would help growers in the strategic management of the crop and climate, in order to achieve the best benefit of the production according to their time schedule.

6. Conclusions

A mechanistic growth model for tomato seedlings (TOMSEED) was developed, based on modifications of existing tomato growth models. PAR at crop level, air temperature and CO_2 concentration were taken into account, while simulated variables included dry weights of leaves, shoot and root (per ground area unit), LAI, number of leaves per plant and finally, shoot length and thickness. Its calibration process was formed into an optimisation problem, taking into account model errors of five simulated characteristics: dry weight of leaves, shoots and roots, LAI and number of leaves per plant. Three heuristic optimisation algorithms were explored during model calibration: genetic algorithms, simulated annealing and tabu search. Genetic algorithms proved to be the most successful approach in the specific model calibration task, resulting on an overall average deviation between simulated and measured values of around 16%.

TOMSEED validation results showed that the proposed model was able to simulate satisfactorily most of the seedlings characteristics, like shoot length and thickness, number of leaves per plant and dry weight distribution, while other characteristics, like dry weight of leaves, shoot and root, as well as LAI, were modelled less successfully. Nevertheless, even with small differences between measured and simulated values, the model gives the opportunity to use it to develop and tune sophisticated climate control and management systems towards the development of high quality seedlings.

APPENDIX A. TOMSEED model loops

A.1. First loop or sub-daily loop

A.1.1. Photosynthesis

Photosynthesis can be restricted by several environmental factors (temperature, radiation) as well as biochemical and physiological factors (Rubisco activity, demand). In the proposed simulation model, photosynthesis is estimated as the minimum value affected by three limiting factors: light, Rubisco and demand (sink limitation).

A.1.1.1. Light limiting factor. The CO_2 absorption rate being limited by light intensity (υ_q) is calculated by the following equation (Teh, 2006):

$$\upsilon_q = e_m I_{abs} \frac{C_i - \Gamma*}{C_i + 2\Gamma*} \tag{1}$$

where e_m is the photosynthetic efficiency, I_{abs} is the radiation absorbed by the canopy, C_i is the CO_2 intercellular concentration, and Γ^* is the CO_2 equilibrium point.

The radiation that is absorbed by the canopy (I_{abs}) is calculated by the following equation:

$$I_{abs} = (1-\rho)\cdot I_0 \cdot \left(1 - e^{-k\cdot L}\right) \tag{2}$$

where ρ is the radiation reflectance coefficient, I_0 is radiation intensity above canopy level, k is the light extinction coefficient, and L is the canopy depth as a function of the LAI of the overlying canopy.

A typical value for ρ is 0.07 (Marcelis, Heuvelink, & Goudriaan, 1998), while a proper value of k for tomato seedlings is 0.57 (Peponakis, 2008).

Photosynthetic efficiency is estimated by the following equation (Bertin & Heuvelink, 1993):

$$e_m = e_0 \frac{[CO_2] - \Gamma*}{[CO_2] + 2\Gamma*} \tag{3}$$

where e_0 is the potential photochemical efficiency and $[CO_2]$ is the CO_2 concentration of the air.

The value of the potential photochemical efficiency for tomato plants is equal to 0.084. The CO_2 equilibrium point is estimated by the following equation (Teh, 2006):

$$\Gamma* = \frac{[CO_2]}{2\tau\cdot Q_{10(\tau)}^{((T_l-25)/10)}} \tag{4}$$

where τ is the CO_2/O_2 specificity factor, $Q_{10(\tau)}$ is the factor for the change in τ for a 10 °C temperature change, and T_{leaf} is the leaf temperature.

The O_2 concentration is 210,000 ppm, while $Q_{10(\tau)}$ is 0.57. For the estimation of leaf temperature (T_{leaf}), the empirical equation developed by Wang and Deltour (1999) is used, based on air temperature (T_{air}) and solar radiation intensity (I_i):

$$T_{leaf} = -2.05 + 1.01\cdot T_{air} + 0.00425\cdot I_i \tag{5}$$

The use of this equation was consider adequate in the case of seedlings that are not grown on heated beds. In addition, it was taken into account that seedling LAI, which affects leaf temperature, can reach or even surpass that of mature plants. The validity of the equation was not tested with regular measurements of leaf temperature, though it was validated with the sporadic use of infrared thermometers and proved to be quite accurate in most cases, with the exception of summer months, when the equation seemed to overestimate leaf temperatures.

A.1.1.2. Rubisco limiting factor. According to Teh (2006), the CO_2 absorption rate limited by Rubisco activity (υ_c) is calculated by the following equation:

$$\upsilon_c = \frac{V_{c,max}(C_i - \Gamma*)}{K_c(1 + O_a/K_o) + C_i} \tag{6}$$

where $V_{c,max}$ is the maximum CO_2 absorption rate from Rubisco activity, K_c and K_o are the CO_2 and O_2 Michaelis–Menten constants respectively, and O_a is the ambient O_2 concentration.

The value of $V_{c,max}$ for a temperature of 25 °C ($V_{c,max(25)}$) is equal to 200 μmol m^{-2} s^{-1} (Bernacchi, Singsaas, Pimentel, Portis, & Long, 2001). For various temperature values, it can be estimated by the following equation, which also takes into account that Rubisco activity ceases for temperatures above 40 °C:

$$V_{c,max} = \frac{V_{c,max(25)} Q_{10(vc)}^{(T_{leaf}-25)/10}}{1 + e^{0.128(T_{leaf}-40)}} \tag{7}$$

A.1.1.3. Demand limiting factor (sink limitation). The CO_2 absorption rate limited by demand (υ_s) is calculated by the following equation (Teh, 2006):

$$\upsilon_s = \frac{V_{c,max}}{2} \tag{8}$$

A.1.2. Respiration

Similarly to TOMSIM and TOMGRO models, respiration is divided into maintenance respiration and growth respiration. Maintenance respiration depends on temperature and some factor that expresses the CH_2O demands for the maintenance of 1 g of existing dry matter. The value of that factor is different for each plant organ. Thus, maintenance respiration is estimated by the following equation:

$$r_m = (DM_{root}\cdot MAINT_{root} + DM_{leaf}\cdot MAINT_{leaf} + DM_{shoot}\cdot MAINT_{shoot}) \cdot Q_{10,resp}^{(T-25)/10} \tag{9}$$

where DM_{root}, DM_{leaf} and DM_{shoot} are the root, leaf and shoot dry matter, respectively, and $MAINT_{root}$, $MAINT_{leaf}$ and $MAINT_{shoot}$ are the root, leaf and shoot demand in CH_2O for maintenance of 1 g of dry matter, respectively. According to Heuvelink (1995), *MAINT* coefficients for a temperature of 25 °C have the following values: $MAINT_{root}$ = 0.01 g (CH_2O) g^{-1} (DM) day^{-1}, $MAINT_l$ = 0.03 g (CH_2O) g^{-1} (DM) day^{-1} and $MAINT_{st}$ = 0.015 g (CH_2O) g^{-1} (DM) day^{-1}. It should be noted that these values are general and have not been proposed specifically for tomato plants at any growth stage.

A.2. Second loop or daily loop

At the end of each sub-daily cycle, the photosynthesis products minus the maintenance respiration are added to a “pool”. The pool's content is available for plant growth during the execution of the daily loop. The pool is not reset at the end of each daily loop, thus any remaining content is available for use during the following day. The capacity of the pool was set to 4.2 g (CH_2O) m^{-2} (leaf area), based on the relevant suggestion by Sitheswary, Putman, and Janes (1990).

A.2.1. Leaf growth demand

Initially, the estimation of the possible growth rate of the plants' organs is performed. The increase rate of leaf area depends only on the temperature, as it does not seem to be influenced by radiation (Peponakis, 2008). Thus, the following formula was used for the estimation of possible increase rate of leaf area (*rLA*):

$$rLA = Pl \cdot Tl \tag{10}$$

where *Pl* is the possible increase rate of leaf area as a percentage of the existing leaf area, and *Tl* is the coefficient for the influence of temperature on the increase rate of leaf area. The values of *Tl* have been calculated at specific temperature values between 0 and 50 °C by Dayan et al. (1993a) and are then estimated by interpolation. For the estimation of *Pl*, the following exponential equation was derived, as a function of the number of leaves per plant (N_{leaves}), which is an indication of plant growth:

$$Pl = Pl_0 \cdot e^{N_{leaves} \cdot lgf} \tag{11}$$

where Pl_0 and *lgf* are empirical coefficients, and N_{leaves} is the number of leaves per plant.

In order to estimate the total demand of the leaves, the specific leaf area (*pSLA*) has also to be taken into account. *pSLA* depends on the availability of photosynthetic products (sugars) and the increase rate of leaves. High increase rate of leaves and low sugars availability would lead to high *pSLA* values. This happens when solar radiation intensity is low, limiting photosynthesis, while temperature is high, and keeping the plant's metabolism high. For the estimation of *pSLA*, the following approach was developed and used in the model:

- If the ratio *Pool* to *PoolLimit* is greater than the ratio *rLA* to *Pl* (where, *Pool* is the quantity of available sugars in the “pool” and *PoolLimit* is the capacity of the “pool”), then:

$$pSLA = Lsm + (Lss - Lsm) \cdot (rLA/Pl) \tag{12}$$

otherwise:

$$pSLA = Lsm + (Lss - Lsm) \cdot [1\text{-}(Pool/PoolLimit)] \tag{13}$$

where *Lsm* is the minimum specific leaf area, and *Lss* is the maximum specific leaf area. It should be noted here that *pSLA* does not express the specific leaf area of the crop, but it rather constitutes a coefficient that determines the increase of leaf area per g of dry matter that is added to the leaves. For that reason, it is presented as *pSLA* and not SLA, and it is subject to calibration.

A.2.2. Root growth demands

The rate of increase of root weight depends on the substrate temperature, which is assumed to be equal to the air temperature, for seedling cultivation in unheated beds. Its estimation is similar to that for leaf area. Another crucial factor that influences the root growth is water availability, but there was assumed to be no water limitation, as is usually the case for seedlings. Thus, the following equation was used, assuming that there is sufficient amount of sugars:

$$rRoot = Pr \cdot Tr \tag{14}$$

where *rRoot* is the rate of increase of root weight as a percentage of existing root weight, *Pr* is the possible increase rate of the root as a percentage of existing root weight, and *Tr* is the influence of substrate temperature. The *Tr* coefficient has been calculated in the same way as the *Tl* coefficient, by Dayan et al. (1993a). The root demand (D_{root}) is calculated based on the rate of root growth (*rRoot*), and the demands for sugars for the development of root dry matter (ASR_{root}):

$$D_{root} = rRoot \cdot ASR_{root} \tag{15}$$

The *Pr* coefficient, like *Pl*, is given by an exponential equation in relation to N_{leaves} (Pr_0 and *rgf* are empirical coefficients):

$$Pr = Pr_0 \cdot e^{N_{leaves} \cdot rgf} \tag{16}$$

A.2.3. Shoot growth demand

By making the assumption that the shoot is a cylinder and that its thickness increases linearly with shoot dry matter, the following empirical equation was used for its estimation, based on 2422 values from six experimental periods ($R^2 = 0.89$):

$$ST = 0.6317 \cdot \log(DM) + 5.2358 \tag{17}$$

where *ST* is the shoot thickness and *DM* is the dry matter. The length of the shoot (*SL*) is estimated using the following empirical equation:

$$SL = 10^{-7} \cdot LA^2 + 0.0031 \cdot LA + 1.6024 \tag{18}$$

where *LA* is the leaf area of a single plant. This equation was derived from data from three different experimental periods ($R^2 = 0.88$).

Dry weight per unit shoot length (*DWLS*) is a linear function of the accumulated radiation (I_{absSum}). The following empirical equation estimates the shoot dry weight per unit length ($R^2 = 0.95$):

$$DWLS = 5 \cdot 10^{-5} \cdot I_{absSum} + 1.1 \cdot 10^{-3} \quad (19)$$

Thus, the dry weight of the shoot can be estimated by simply multiplying the shoot length by the dry weight per unit length.

A.2.4. Dry matter distribution

It was considered that the plant distributes the products of photosynthesis with priority to the organs that have to be developed in order to compensate for some specific deficiency. Thus, as light becomes a limiting factor, metabolite distribution towards the upper part will increase, so that the plant can increase in height (increase in shoot length) and increase its leaf area. Similarly, if water is the limiting factor, then the products of photosynthesis are distributed mainly towards the root, so that the root is developed to make possible water uptake from a larger volume of soil. When both light and water constitute limiting factors, root demands are initially covered, so that water needs are satisfied, as water is considered a more vital factor than light. When there is not any limiting factor present, the distribution is performed according to each organ's needs.

The criterion for light sufficiency was the ratio of CH_2O supply to total CH_2O demand, while the criterion for water sufficiency was the ratio of the evapotranspiration of the crop to the available water in the substrate. In the case where CH_2O quantities were not sufficient, the root assimilate requirements were partially satisfied. Thus, the root dry matter accumulation (ΔDM_{root}) is given by the following equation ($minr$ is the percentage of available sugars that will be distributed to the root):

$$\Delta DM_{root} = \frac{minr \cdot Pool}{ASR_{root}} \quad (20)$$

The ASR_{root} coefficient (assimilate requirements of the roots) is used so that the “transformation cost” of non-structural dry matter to structural dry matter is taken into account. Similarly, the accumulation of shoot dry matter (ΔDM_{shoot}) is given by the following equation:

$$\Delta DM_{shoot} = \left(\frac{(1 - minr)Pool}{D_{total}} D_{shoot}\right) ASR_{shoot} \quad (21)$$

where D_{shoot} and D_{total} are the demand of the shoot and the total demand of all organs, respectively, and ASR_{shoot} is the assimilate requirements of the shoot.

The rest of the “pool” is distributed to the leaves, divided again by the ASR_{leaf} coefficient (the assimilate requirements of the leaves). The values proposed by Heuvelink (1995) for the ASR coefficients are between 1.39 and 1.45, depending on the organ. However, these values are not tomato-specific. Gijzen (1994) and Gary, Bot, Frossard, and Andriolo (1998) have proposed significantly lower values (1.02–1.16, depending on the organ). In the proposed seedlings model, the values given by Gary et al. (1998) were used, as they relate to young tomato plants. Thus, $ASR_{leaf} = 1.05$ and $ASR_{shoot} = 1.02$, while ASR_{root} was arbitrarily set equal to 1.05, as no specific value has been proposed in the literature.

The increase of leaf area index (ΔLAI) is calculated by the following straightforward equation:

$$\Delta LAI = \Delta DMleaf \cdot pSLA \quad (22)$$

where ΔDM_{leaf} is the leaves dry matter accumulation.

A.2.5. Number of leaves

TOMGRO provides a function to estimate the appearance of new leaves that seemed to be suitable for seedlings too. Based on that, the rate of the development of new leaves (rN) is estimated by the following equation:

$$rN = rM \cdot fT \cdot fasm \quad (23)$$

where, rM is the maximum rate of development of new leaves, fT is a coefficient that incorporates the effect of lower than optimal temperatures (estimated similarly to the Tl and Tr coefficients, as explained earlier), while the $fasm$ coefficient is given by:

$$fasm = \frac{Pool}{D_{total}} GENFAC \quad (24)$$

where $GENFAC$ is the coefficient of the effect of sugars availability on the development of new leaves.

REFERENCES

Andersen, A. (1986). Comparison of fluorescent lamps as an energy source for production of tomato plants in a controlled environment. *Scientia Horticulturae, 28*, 11–18.

Bernacchi, C. J., Singsaas, E. L., Pimentel, C., Portis, A. R., Jr., & Long, S. P. (2001). Improved temperature response functions for models of Rubisco-limited photosynthesis. *Plant, Cell & Environment, 24*, 253–259.

Bertin, N., & Heuvelink, E. (1993). Dry-matter production in a tomato crop: comparison of two simulation models. *Journal of Horticultural Science, 68*, 995–1011.

Dayan, E., van Keulen, H., Jones, J. W., Zipori, I., Shmuel, D., & Challa, H. (1993a). Development, calibration and validation of a greenhouse tomato growth model: I. Description of the model. *Agricultural Systems, 43*, 145–163.

Dayan, E., van Keulen, H., Jones, J. W., Zipori, I., Shmuel, D., & Challa, H. (1993b). Development, field calibration and validation of a tomato greenhouse model. II. Field calibration and validation. *Agriculture Systems, 43*, 165–183.

Dimokas, G., Tchamitchian, M., & Kittas, C. (2009). Calibration and validation of a biological model to simulate the development and production of tomatoes in Mediterranean greenhouses during winter period. *Biosystems Engineering, 103*, 217–227.

Garcia, F. (1999). Use of reinforcement learning and simulation to optimize wheat crop technical management. In *Proceedings of the International Congress on Modelling and Simulation ModSim'99* (pp. 801–806).

Gary, C., Bot, J. L., Frossard, J. S., & Andriolo, J. L. (1998). Ontogenic changes in the construction cost of leaves, stems, fruits and roots of tomato plants. *Journal of Experimental Botany, 49*, 59–68.

Gijzen, H. (1994). *Development of a simulation model for transpiration and water uptake and an integral crop model*. Wageningen: Plant Research International. Report 18.

Glover, F. (1989). Tabu search – Part I. *ORSA Journal on Computing, 1*, 190–206.

Goldberg, D. E. (1989). *Genetic algorithms in search, optimization and machine learning*. Reading, MA: Addison-Wesley.

Goudriaan, J., & van Laar, H. H. (1994). *Modelling potential crop growth processes: Textbook with exercises*. Kluwer Academic Publishers.

Gupta, M. K., Chandra, P., Samuel, D. V. K., & Singh, B. (2012). Modeling of tomato seedling growth in greenhouse. *Agricultural Research, 1.*, 362–369.

Gupta, M. K., Samuel, D. V. K., & Sirohi, N. P. S. (2010). Decision support system for greenhouse seedling production. *Computers & Electronics in Agriculture, 73*, 133–145.

Herrera, F., Castillo, J. E., Chica, A. F., & López Bellido, L. (2008). Use of municipal solid waste compost (MSWC) as a growing medium in the nursery production of tomato plants. *Bioresearch Technology, 99*, 287–296.

Heuvelink, E. (1995). Dry matter production in a tomato crop: measurements and simulation. *Annals of Botany, 75*, 369–379.

Heuvelink, E. (1996). Dry matter partitioning in tomato: Validation of a dynamic simulation model. *Annals of Botany, 77*, 71–80.

Hoek, I. H. S., Ten Cate, C. H. H., Keijzer, C. J., Schel, J. H., & Dons, H. J. M. (1993). Development of the fifth leaf is indicative for whole plant performance at low temperature in tomato. *Annals of Botany, 72*, 367–374.

Holland, J. H. (1975). *Adaptation in natural and artificial systems*. Ann Arbor: University of Michigan Press.

Jones, J. W., Dayan, E., Allen, L. H., van Keulen, H., & Challa, H. (1991). A dynamic tomato growth and yield model (TOMGRO). *Transactions of the ASABE, 34*, 663–672.

Kirkpatrick, S., Gelatt, C. D., & Vecchi, M. P. (1983). Optimization by simulated annealing. *Science, 220*, 671–680.

Marcelis, L. F. M., Heuvelink, E., & Goudriaan, J. (1998). Modelling biomass production and yield of horticultural crops: a review. *Scientia Horticulturae, 74*, 83–111.

Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H., & Teller, E. (1953). Equation of state calculations by fast computing machines. *Journal of Chemical Physics, 21*, 1087–1092.

Michalewicz, Z., & Fogel, D. B. (2004). *How to solve it: Modern heuristics*. Springer Science & Business Media.

Morimoto, T., & Hashimoto, Y. (1996). Optimal control of plant growth in hydroponics using neural networks and genetic algorithms. *Acta Horticulturae, 406*, 433–440.

Morison, J. I. L. (1987). Intercellular CO_2 concentration and stomatal response to CO_2. In E. Zeiger, G. D. Farquhar, & I. R. Cowan (Eds.), *Stomatal function* (pp. 229–252).

Pavlovic, R., Petrovic, S., & Stevanovic, D. (1998). The influence of transplants quality on the yield of tomato grown in plastic house. *Acta Horticulturae, 456*, 81–86.

Peponakis, K. (2008). *Effect of light and temperature on the growth and development of tomato seedlings for optimal control*. Master Thesis. School of Agriculture, University of Thessaly, 217 pp. (in Greek).

Pucheta, A. J., Schugurensky, C., Fullana, R., Patiño, H., & Kuchen, B. (2006). Optimal greenhouse control of tomato-seedling crops. *Computers & Electronics in Agriculture, 50*, 70–82.

Radoglou, K., & Raftoyannis, Y. (2002). The impact of storage, desiccation and planting date on seedling quality and survival of woody plant species. *Forestry, 75*, 179–190.

Scholberg, J. M. S., Boote, K. J., Jones, J. W., & McNeal, B. L. (1997). Adaptation of the CROPGRO model to simulate the growth of field-grown tomato. In M. J. Kropff, P. S. Teng, P. K. Aggarwal, J. Bouma, B. A. M. Bouman, J. W. Jones, et al. (Eds.), *Systems approaches for sustainable agricultural development: Applications of systems approaches at the field level*. Dordrecht, The Netherlands: Kluwer Academic Publishers.

Seginer, I., & Raviv, M. (1984). Optimal night temperatures for greenhouse seedlings. *Scientia Horticulturae, 23*, 203–216.

Sitheswary, L., Putman, H. F., & Janes, H. W. (1990). The influence of light period on carbon partitioning, translocation and growth in tomato. *Scientia Horticulturae, 42*, 75–83.

Teh, C. (2006). *Introduction to mathematical modeling of crop growth: How the equations are derived and assembled into a computer model*. Brown Walker Press.

Vanthoor, B. H. E., De Visser, P. H. B., Stanghellini, C., & Van Henten, E. J. (2011). A methodology for model-based greenhouse design: Part 2, description and validation of a tomato yield model. *Biosystems Engineering, 110*, 378–395.

Wang, S., & Deltour, J. (1999). Lee-side ventilation-induced air movement in a large-scale multi-span greenhouse. *Journal of Agricultural Engineering Research, 74*, 103–110.

Yan, H. P., Kang, M. Z., de Reffye, P., & Dingkuhn, M. (2004). A dynamic, architectural plant model simulating resource-dependent growth. *Annals of Botany, 93*, 591–602.