Optimization of parameters of the 'Virtual Fruit' model to design peach genotype for sustainable production systems

Bénédicte Quilot-Turion, Mohamed Mahmoud Ould Sidi, Abdeslam Kadrani, Nadine Hilgert, Michel Génard, Francoise Lescourret

Optimization of parameters of the 'Virtual Fruit' model to design peach genotype for sustainable production systems

Quilot-Turion Bénédicte[a], Ould-Sidi Mohamed-Mahmoud[b], Kadrani Abdeslam[b], Hilgert Nadine[c], Génard Michel[b], Lescourret Françoise[b]

[a] *UR1052, Génétique et Amélioration des Fruits et Légumes, INRA, F-84000 Avignon, France*
[b] *UR1115, Plantes et Systèmes de Culture Horticoles, F-84000 Avignon, France*
[c] *UMR 729 Mathématiques, Informatique et STatistique pour l'Environnement et l'Agronomie, INRA SUPAGRO, F-34000 Montpellier, France*

Corresponding author: B. Quilot-Turion
INRA GAFL, BP 94 - 84143 MONTFAVET cedex (France)
Tél: +33 432722763 Fax: + 33 432722702
E-mail address: quilot@avignon.inra.fr

ABSTRACT

Designing sustainable production systems that are respectful of the environment and produce safe food of quality is a challenge for the future. A critical step is to find the best combinations of genetic resources and cultural practices adapted to target environments. The objective of this work is to illustrate how an ecophysiological process-based simulation model could be possibly used to design genotypes and to propose innovative production systems, by applying a methodology of optimization.

As example, we studied the peach-brown rot system and used the 'Virtual Fruit', a process-based model that has been extensively tested, to perform virtual experiments. The challenge was to optimize the trade-off between antagonistic criteria of major importance for both fruit quality (increasing fruit mass and sweetness) and sensitivity to brown rot (decreasing skin density of cracks) in four different cultural scenarios. A multiobjective evolutionary algorithm, namely NSGA-II, was applied to solve this multiobjective optimization problem based on the 'Virtual Fruit'. The optimized variables were six parameters of the 'Virtual Fruit', selected on the basis of a sensitivity analysis.

This optimization method provided a large diversity of solutions among which the decision-maker can choose the best suited trade-off between criteria according to a particular objective. Most of the optimized solutions were distributed along Pareto fronts suggesting a good convergence of the algorithm. Moreover, it also provided some solutions located in non-crowded zones which constitute some original alternatives for the final decision-maker.

The results confirmed the strong antagonism between the criteria considered. Large fruits had a weak sweetness and high crack density and for a given mass, those with improved sweetness had higher crack density. In a current breeding scheme, fruit mass would be the only criteria considered but alternative schemes could be considered for future, favoring organoleptic quality or environment friendly practices. In those cases, some interesting optimized solutions were identified.

The work described in this paper supports that multiobjective evolutionary algorithms should be used to optimize parameters of process-based models and help identifying trade-off in complex systems. The use of the 'Virtual Fruit' to design sustainable production systems combining genotypes and innovative practices is further discussed.

1. Introduction

Production systems have to adapt, in a changing climatic context, to face the growing social demand in terms of organoleptic, nutritional and environmental quality of food products. Indeed, improving quality is ever a matter of importance for fruit production, and controlling plant diseases is now a priority for plant health, product safety and environmental issues. Consequently, the critical question for the future is how to design best combinations of genetic resources and cultural practices adapted to, and respectful of specific environments, to meet the social demand for a multiobjective quality.

In this context, dealing with storage disease resistance is a priority to reduce fruit chemical residues and increase food safety. Among them, brown rot in peach fruits caused by *Monilinia* spp, can cause as much as 30 to 40% of crop losses. No other alternative to chemical treatment is available (biological control, genetic resistance…), hence fungicide applications are required till pre-harvest. Currently, all cultivated peaches are more or less sensitive to brown rot. The only cultivar known to have a high level of resistance is the Brazilian cultivar Bolinha (Feliciano et al., 1987), which is not available in Europe. Its resistance is thought to be based on complex mechanisms that are largely linked to fruit characteristics such as skin conductance and density of cracks on the skin surface and epidermis and cuticle biochemical compositions (Gradziel et al., 1998; Isaacson et al., 2009).

Considering *Monilinia* spp, conidial infection is largely occurring through wounds, so cuticular cracks are likely to play a major role for fungal infection (Gibert et al., 2009). Cuticular cracks can represent more than 10% of the fruit surface area. The cuticular crack density was shown to mainly vary with the intensity of the fruit growth, itself varying with management practices such as irrigation regimes and fruit thinning (Gibert et al., 2010). Sensitivity of cuticle to cracks may also be controlled by the genotype. Those cracks are not only large opportunities for fungal infection, but also contribute to fruit transpirating losses that influence fruit growth and quality build-up.

Fruit quality itself is a complex issue, which results from many overlapping physiological processes, genetically and environmentally controlled. In this context, the challenge is to specifically lead to a reduction of fruit sensitivity to brown rot while keeping or improving fruit quality, and thus to enhance ecological, economical and health benefits, combining different processes at the fruit and plant levels. An integrative approach is necessary to handle the links and the antagonisms between the biological processes under the influence of genotype and cultural practices.

The strong genotype x environment x practices interactions in agricultural systems make it necessary to design genotypes that are adapted to specific agro-environmental conditions (e.g. Hammer et al., 2002). This issue is of increasing interest to the crop modeling community. Thus, many researchers have used the approach that consists of analyzing plant traits via a model based on physiological mechanisms, which describes the development of such traits through environment conditions and genetic control namely by means of genetic parameters (e.g. Yin et al., 2005). The next step is to optimize the combination of genetic parameters to design new genotypes, real or virtual plant cultivars with an ideal phenotype, adapted to target environments (Tardieu, 2003; Letort et al., 2008). To meet this challenge some modelers have first used techniques such as trial and error methods (Haverkort and Grashoff, 2004; Herndl et

al., 2007). These attempts were quickly confronted to the difficulty and the hardness of the task. Indeed, the design of innovative cultivars is based on antagonistic criteria with respect to strong constraints (economical or environmental constraints such as water saving…). This is a nonlinear multiobjective optimization problem known to be difficult to solve using classical methods. Nowadays, new computing technologies allow running millions of simulations very quickly and getting the entire landscape of the solutions. However, access to these technologies remains limited and some problems may lead to non solvable combinational explosion. Moreover, the enormous number of combinations to analyze to identify best-adapted genotypes highlights a major interest for predictive approaches. Therefore, thanks to collaborations between researchers from various disciplines, i.e. physiology, agronomy, and applied mathematics, effective methods have been recently proposed. In the domain of agriculture, several works studied multi-objective optimization problems (Raju and Kumar, 1999; Francisco and Ali, 2006; deVoil et al., 2006). However, these works mainly focused on agricultural systems and logistics, considering factors related to crop rotation and water planning only. Dealing with optimization of parameters that describe the inner physiological processes of plant growth has never been done before the two pioneer studies of Letort et al. (2008) and Qi et al. (2010). They took the lead in this particular issue and developed methods originating from the field of multiobjective optimization algorithms, e.g. genetic (Letort et al., 2008) and particle swarm (Qi et al., 2010) algorithms. However, the first example used a simple version of a genetic algorithm and dealt only with a mono-objective problem, namely to maximize the cob weight of maize. The second example used a particle swarm algorithm with its very good exploration (global search) abilities but with weak exploitation (local search) abilities. In addition this algorithm hardly maintains diversity within the swarm (population of solutions). The direct consequence is a poor choice for the decision-maker.

In this context, we propose to use the 'Virtual Fruit' (Lescourret and Génard, 2005, Génard et al. 2007 and Génard et al. 2010), a process-based model that has been extensively tested, to develop a methodology of optimization. We used the NSGA-II (Non-dominated Sorting Genetic Algorithm II) algorithm, well-known by the optimization community. Such multiobjective evolutionary optimization algorithms allow exploring highly dimensional solution spaces in a reasonable computation time. These methods provide the decision-maker with a set of diversified solutions with reduced, but sufficient, cardinality. The decision-maker will thus have the final choice of the best suited trade-off between criteria and will be provided with the corresponding optimal solutions.

The 'Virtual Fruit' was used to design peach genotypes with enhanced values for fruit quality and resistance aspects (brown rot sensitivity) and adapted to given cultural scenarios. We focused the work on six parameters of the model identified via a sensitivity analysis, to be combined to create the genotypes (a genotype is further considered as a set of the 6 parameters). Three traits (further called 'criteria') simulated by the model and of major importance for fruit quality and sensitivity to brown rot were taken into account to evaluate the individuals. No priority was given to any of those three traits compared to the two others. Simulations were performed for four cultural scenarios (two levels of crop load and two water regimes) in order to analyse the putative impact of cultural practices on the optimized solutions. Further development of such method and the limits to overcome were discussed.

2. Materials and methods

2.1. Description of the 'Virtual Fruit' model

The 'Virtual Fruit' represents the quality development of an 'average' peach fruit on a 1-year-old stem. . This is the basic production unit for peach growers which bears both fruit and leafy shoots.

The 'Virtual Fruit' integrates five sub-models (Fig. 1) and takes into account the effect of some climatic factors (radiation, air temperature and humidity), some practices (irrigation, thinning) and the genotype (Quilot et al, 2005b). A first step of integration was proposed by Lescourret and Génard (2005) which combines in a single model three existing process-based models describing dry mass, sugar and water accumulation in the flesh of fruit. This model was successfully confronted to data from factorial experiments from three different years under well-irrigated regimes. The relative root mean squared error of prediction was calculated using a cross-validation approach (Wallach et al., 2001): it was smaller than 20% in most cases (Lescourret and Génard, 2005). Moreover, a simplified version was tested on 87 different peach genotypes, and the error of prediction was also smaller than 20% (Quilot et al., 2005a). In water-stress conditions, simulated data from the sub-model describing the biophysical representation of water transport towards the fruit (Fishman and Génard, 1998) were comparable in terms of fresh mass reduction to the experimental data from Berman & DeJong (1996) and assessed the capacity of the water sub-model to simulate the combined effect of water stress and crop load on fruit growth. Additional sub-models describing skin conductance and microcracking (Gibert et al., 2010), and respiration (Génard and Gouble, 2005), were incorporated (Génard et al., 2010). Further details on the 'Virtual Fruit' and the submodels are available in the original papers and a synthesis of the complete 'Virtual Fruit' is described in Génard et al. (2010).

The inputs of the model are the photosynthetic active radiation, mean daily temperature, air relative humidity and leaf and stem water potentials (Fig. 1). Values for climatic data (see Fig. 1 in supplementary material) were obtained for 2009 from INRA weather stations close to the experimental orchards in Avignon (latitude 43°91′N, longitude 4°85′E and altitude 24 m.a.s.l.). Daily values of mean leaf and stem water potentials were defined by Gibert (2007) from pre-dawn and midday measurements in orchard-grown trees and in container-grown trees in 2004 and 2005.

The state of the fruit-bearing stem (number of leafy shoots and fruits on the stem, and initial values of the state variables) was also described. Parameter values (see Table 1 in supplementary material) were obtained from the literature or estimated for peach based on previous studies (Fishman and Génard, 1998; Lescourret et al., 1998; Génard et al., 2003; Lescourret and Génard, 2005).

The outputs used in this work are three fruit traits: fruit size, sugar concentration and skin density of cracks. The 'Virtual Fruit' is driven by its parameters, which are constant over time and supposed independent of the environment. The simulations were performed for the year 2009 from 87 days after bloom (DAB) to 150 DAB . The parameter values were obtained from the literature or estimated for peach based on previous studies (Fishman and Génard, 1998; Lescourret et al., 1998; Génard et al., 2003; Lescourret and Génard, 2005; Gibert et al., 2010).

2.2. Description of the scenarios

Four cultural scenarios were defined and used as inputs of the 'Virtual Fruit' for the optimization step. They were combinations of two contrasted levels of crop load and two contrasted irrigation regimes. The low (LC) and high (HC) crop loads counted respectively 4 and 20 fruits per fruit-bearing stem and are defined in the initial state of the model. These two crop loads were chosen to be highly contrasted, the former corresponding to an important commercial thinning and the latter to no fruit thinning. The two contrasted irrigation regimes are further called well-irrigated (WI) and water deficit (WD) regimes. The latter corresponds to a long and moderate stress during the final swelling, which presents the interest to limit cuticular crack apparition (Gibert et al., 2007). The water potentials which directly reflect the level of water stress experienced by the plant and are used by the growers as water status indicators were used as inputs of the model. Berman & DeJong (1996) and Naor et al. (1997) observed that for a same limiting irrigation regime, plants with high crop load are more stressed than plants with low crop load. Consequently, in case of the WD regime we modulated the water potentials according to the load (Fig. 2) to mimic the same level of restricted irrigation (data taken from Gibert, 2007). The four scenarios are further referred as WI_LC, WI_HC, WD_LC and WD_HC.

2.3. MultiObjective Evolutionary Algorithms

The multiobjective optimization aims at minimizing or maximizing simultaneously several criteria with respect to a set of constraints on the decision variables (here parameters of the 'Virtual Fruit' model). Constraints may be equalities or inequalities, bounds of decisions variables or more complicated. Two principal types of approaches are dedicated to solve multiobjective optimization problems. The first type consists of transforming the original multiobjective problem into a single objective problem using a weighted combination of the different objectives. The second type of resolution methods, which are used in this work, are said Pareto approaches and take into account all the objectives without aggregation (for more details see Coello et al. 2002, chapter 1). In order to compare the solutions, these approaches use the concept of dominance relation between solutions. Solution A is said to be better than solution B if A *dominates* B i.e. if and only if: (i) A is not worse than B in all objectives, and (ii) A is strictly better than B in at least one objective. Among a given set of solutions, the non-dominated subset of solutions are those that are not dominated by any member of the set. The non-dominated subset of the all solutions is the Pareto-optimal set. Each solution of the set represents a different compromise among design objectives. The Pareto optimal front (POF) is the set of Pareto-optimal solutions satisfying that no individual criterion can be improved without another criterion being made worse.

Over the last decades, a number of MultiObjective Evolutionary Algorithms (MOEAs) have been suggested (Deb 2001; Coello et al. 2002). The MOEAs are based on the concept of dominance relation between solutions. The main reason for the popularity of evolutionary algorithms for solving multiobjective optimization is their population-based nature and ability to find multiple optima simultaneously. These algorithms were used for solving the NP-hard problems which are considered as the most difficult optimization problems. The goals of MOEAs are to (i) identify the global Pareto-optimal set of solutions (ii) find a sufficient coverage of that set (iii) find an even distribution of solutions. Thus the decision-maker will be provided with a large choice of solutions and can select a design compromise that suits his/her preference structure (Bandyopadhyay et al. 2008). Once many Pareto-optimal solutions are obtained, POF can be represented by its own response surface (Goel et al.,

2007). The POF would allow the decision-maker to visualize and assess trade-offs among the objectives, to explore compromise solutions, and to take decisions based on realistic goals.

2.4. Optimization using NSGA-II

We used a Pareto approach to optimize genotypes, namely NSGA-II which is a MOEA developed by Deb et al. (2002). Fig. 3 presents a sketch of the NSGA-II algorithm inspired from the original one proposed by Deb et al. (2002). The pseudo-code of this algorithm is the following:

1. Initialization: randomly initialize parameters of the population of size npop (npop individuals). This population is called parent population and noted P.
2. Evaluation : calculate the values of criteria and constraints for each individual
3. Classification: rank the population using non-domination relation (best individuals in the first front).
4. Preserving the diversity within the population: calculate the crowding distance. The crowding distance computes the relative closeness of a given solution to other solutions on the same front (see Appendix A). The maximal the crowding distance the better the solution.
5. Child population: use genetic operators i.e. selection, crossover, mutation to create a child population from the parent population (npop children) (see Appendix A). Child population is noted Q.
6. Evaluation of child population
7. Recombined population: combine the parent and child population and calculate the crowding distance. The recombined population noted R is of size 2npop.
8. New parent population: select npop best individuals based on the rank and the crowding distance. This population will be used as parent population P in the next generation.
9. Go to the step 3 and repeat till stopping criterion is satisfied (number of generations in our case)

From two solutions selected randomly, the algorithm performs selection as follows: if the solutions belong to two different non-dominated ranks, the solution with lower rank is selected. If the solutions have the same rank, the algorithm prefers the solution located in a lesser crowded zone i.e. with higher crowding distance. Otherwise, one solution is picked randomly from the two solutions. Once the two parents selected in this way, the crossover operator produces two children. The NSGA-II algorithm uses the simulated binary crossover (SBX) operator (see Appendix A). The NSGA-II also uses polynomial mutation operator, this operator mutes solution x^1 into y^1 as described in Appendix A. More details on this algorithm are available in Deb et al. (2002) and Bandyopadhyay et al. (2008). The NSGA-II was applied to generate the combinations of the six parameters of the model which satisfy better our three criteria (fresh mass, sweetness, and crack density). The aim was to maximize fruit mass and sweetness and minimize skin crack density. The bounds of variations of the three criteria were set to 50 and 300g for the fruit mass, 4 and 20% for sweetness and 0 and 20% for skin crack density. The fitness of each solution measures how this solution satisfies fruit mass, sweetness, and crack density. We used the 'Virtual Fruit' to determine (calculate) the fitness of every individual in the population of solutions during the search process (over generations). For these simulations we used the mutation and crossover operators initially recommended by Deb et al. (2002). The population size and the number of generations were chosen after performing many simulations and taking into account the needs of our case study.

The following parameters were set for the simulations:

Population size (npop) 400

Generations 1500

Crossover probability (Pcross) 0.9

Distribution parameter (for crossover) 20

Mutation probability (Pmut) 0.10

Distribution parameter (for mutation) 20

2.5. Selection of six parameters of the 'Virtual Fruit'

A sensitivity analysis of the 'Virtual Fruit' model was performed to identify the parameters that mainly affect outputs of interest of the model and select them for the optimization step. The elementary effects screening method (Morris, 1991; Saltelli et al., 1999) was used to analyse all the parameters of the 'Virtual Fruit'. For a given model with k independent input parameters $X = (x_1, ..., x_k)$, suppose that each parameter x_i takes p selected values. Note Ω the k-dimensional unit cube defined by these selected values. Elementary effects of the parameter x_i can be defined as follows:

$$ee_i = \frac{\left[y\left(x_1, ..., x_{i-1}, x_i + \Delta, x_{i+1}, ..., x_k\right) - y\left(x_1, ..., x_{i-1}, x_i, x_{i+1}, ..., x_k\right) \right]}{\Delta}$$

where $\Delta \in \{\frac{1}{p-1}, ..., 1 - \frac{1}{p-1}\}$ and $X = (x_1, ..., x_k)$ is any point in Ω such that $(X + e_i\Delta) \in \Omega$ for $i \in \{1, ..., k\}$, with e_i a vector with 1 as i[th] component and zeros elsewhere (Saltelli et al., 2008).

For each parameter, two sensitivity measures are computed: the mean and the standard deviation of the distribution of the elementary effects associated with a given parameter. They respectively assess the overall influence of a given parameter on the output and the interactions of a parameter with other parameters. The sensitivity of each output variable to each parameter was quantified by considering an interval of variation of 10% around the reference value.

From about sixty parameters submitted to the sensitivity analysis, we selected six parameters (Table 1) with the highest significant impact on the model outputs (fruit fresh mass, sweetness, crack density) for each of the six different processes represented in the 'Virtual Fruit'. These processes are vegetative activity, fruit growth, sugar metabolism, stone-pulp partition, transpiration and water fluxes.

The optimization procedure explores the parameter values within a given domain of variation (Table 1). To define a reasonable domain of what could be possible to achieve in adequation with natural variation, we gathered experimental and literature data. For three parameters, the bounds were set to the extreme values observed within a highly variable interspecific peach population (Quilot et al., 2005a). For one parameter, the bounds were derived from values found in the literature for various species (Escobar-Gutiérrez and Gaudillére, 1997; Escobar-Gutiérrez et al., 1998; Lo Bianco et al., 2000). Unfortunately for the last two parameters, no data were available and the range was set to ±10% variation from the default value used in the 'Virtual Fruit' (supplementary data Table 1S).

2.6. Statistical analysis

Principal component analyses (PCA) were performed to analyse the diversity of the phenotypes within the Pareto-optimal set of 400 solutions, further called "optimized population". The goal of the PCA is to summarize a multivariate dataset (the solutions characterized by six parameters of the model and three fruit traits) as accurately as possible using a reduced number of uncorrelated components (Jolliffe, 2002).

All data analyses were carried out using R software (R Development Core Team, 2008). For the sensitivity analysis, we used the package available at http://cran.r-project.org/web/packages/sensitivity. The optimization procedure was performed using a connection between R and C source code. We adapted for that the R package "mco" developed by Heike Trautmann and available at http://cran.r-project.org/web/packages/mco.

3. Results

3.1. Determination of the genotypes optimizing three criteria

The NSGA-II algorithm was applied to generate the best combinations of the six parameters (Table 1) with respect to our three criteria (fresh mass, sweetness, and crack density). The optimization step was performed for each of the four cultural scenarios using 2009 climatic data. The optimized populations are four sets of 400 phenotypes, each combination of six parameters standing for their genotypes.

The influence of the parameters on the criteria is complex because a same parameter can have antagonist effects on the various criteria and different parameters can influence the same criterion. Table 2 presents the correlations between the parameters, between the criteria and between the parameters and the criteria for the 400 selected genotypes of WI_LC scenario. Many significant correlations with low or middle values, few correlations with high values and both positive and negative values well illustrate the complexity of the system studied.

The distributions of the optimized solutions of the six parameters of the model within the possible domain of variations are presented in Fig. 4. For the proportion of carbon as sucrose in phloemic sap, λ_{ph}, high values were favored. Despite an antagonist effect of this parameter on two criteria (high values increase sweetness but also crack density) (Table 2), all optimized solutions were concentrated towards high values. A small value of λ_{ph} is not an elected solution to reduce crack density. Similarly, optimized values of $share_1$ (a parameter of the relationship between stone dry mass and total fruit dry mass) were concentrated towards low values in a very small range of the possible domain of variation. Small values result in higher proportion of flesh in the fruit, which increases the entry of carbon in the fruit and thus the sweetness. Again, a large value of $share_1$ was not retained as solution to reduce crack density. Optimized values of parameter r_1 (part of the leaves in the structural part of the leafy shoots) were extremely concentrated towards high values but only in case of high crop load. High values of r_1 increase part of the leaves in the leafy shoot to the detriment of the stem, which increases the level of carbon assimilation. In case of low crop load, high values were also largely favored but the range was larger, the peak of distribution being smaller and expanded. The distribution of RGR_{ini} (initial relative fruit growth rate) values drew two peaks depending also on the crop load but not on the irrigation regimes. Their height and spreading were comparable but their position did not overlap. In addition, the range of selection was not as strict as for the previous parameters. Lastly, values of cut_2 and Lx (respectively, a parameter of the equation of the relative expansion rate of the cuticule surface area and the conductivity of the composite membrane for water transport) were distributed all along the possible

domain of variation, in a heterogeneous manner depending on the scenarios. In the same way as for RGR_{ini}, one can hardly prefigure best values of cut_2 and Lx.

3.2. A trade-off between antagonist criteria

The phenotypes of the optimized populations identified under four cultural scenarios were simulated by the 'Virtual-Fruit'. In Fig. 5 the optimized solutions were compared to a set of 400 random individuals (the 6 parameters were sampled at random 400 times) that represents part of the possible landscape. Best phenotypes outranged the random landscape (Fig. 5b), even though their parameters stood in the same plan as the one of the random individuals (Fig. 5a). This emphasized the benefit of the optimization step to design combinations of parameters resulting in good phenotypes. Indeed, best phenotypes stood in a zone hardly explored by random exploration suggesting highly refined combinations of parameters.

Fig. 6 presents the correlations between criteria for the WI_LC scenario (see Fig. 2 in supplementary material for 3D representation). Similar results were obtained for the other three scenarios (see online supplementary data Fig. 3). The results confirmed the strong antagonism between the three criteria considered at maturity (fruit mass, sweetness and skin density of cracks). Most of the optimized solutions were distributed along a Pareto front revealed on the plan linking fruit mass and sweetness (Fig. 6a) suggesting a good convergence of the algorithm. Phenotypes located along this front were the best compromises between fruit mass and sweetness and there was no possible way to further improve jointly these two criteria. However, those solutions were not the best regarding the third criteria and some alternative solutions in non-crowded zones, inside the front, displayed lower crack density. In Fig. 6b, two clouds of solutions can be observed. The lower cloud corresponded to smaller crack density values but also to decreased sweetness values compared to the upper cloud.

Some particular solutions displaying contrasted trade-offs between the three criteria are presented in Table 3 and reported in Fig 6. Solutions 1, 2 and 3, respectively favored high fruit mass, low crack density and high sweetness, whereas solution 4 optimized both sweetness and crack density to the detriment of fruit mass. Solutions 5, 6 and 7 proposed real compromises between the three criteria. As a result, the decision-maker disposes of solutions with a large range of fruit mass and for a given mass he can choose those with high sweetness but high crack density or on the contrary with low crack density but lower sweetness.

3.3. Which parameters drive the selected phenotypes?

In order to explore the phenotypes resulting from the optimization, a PCA analysis was performed on the three criteria of the optimized population simulated in the WI_LC scenario (Fig. 7) and in the three other scenarios (see Fig. 4 in supplementary material). To further identify the structure of the population, and to analyse the impact of the parameters on the resulting phenotypes, data points were coloured according to the criteria and the parameter values, respectively. The cloud of the WI_LC scenario data points displayed an amazing form with three opposite heads. Data points were not evenly distributed in the cloud, some zones being crowded and others sparsely populated. In addition, a gap with no solution split the cloud into two parts that did not correspond to two contrasted phenotypes of fruit. The three criteria, declined into levels (Fig. 7a, b, c), divided the cloud into slices each with a particular orientation. Slices were vertical for fruit mass, were oblique from top-right to bottom-left for sweetness, and oblique from top-left to bottom-right for crack density. Consequently, the two superior heads corresponded to phenotypes with maximal values of sweetness and crack

density, respectively and the bottom head was defined by low sweetness and crack density and intermediate fruit size. Finally, only the few phenotypes in the middle of the cloud corresponded to a real compromise between the three criteria: criteria had neither best values nor bad ones.

Regarding the parameters studied, three of them (r_1, λ_{ph} and $share_1$) displayed values randomly distributed in the cloud (Fig. 7d, f, g). They do not drive the phenotypes towards contrasting types and their values may be linked to the presence of local optima. This is in accordance with the low correlations observed between these parameters and the three criteria (Table 2). For the other parameters, clear grouping of levels were observed (Fig. 7e, h, i). The separation of the cloud into two parts resulted from cut_2 values which divided the cloud into horizontal slices. The pattern of distribution of RGR_{ini} values did not show any common point with other patterns whereas the distribution of Lx values was somehow opposed to that of sweetness. Indeed, sweetness seemed to be largely driven by Lx values, as suggested by the correlation table (Table 3). Low fruit mass seemed linked to both low RGR_{ini} values and low Lx values whereas large fruit mass was not linked to the highest values of RGR_{ini}. High values of Lx explained part of large fruit mass but some genotypes with high Lx values displayed low fruit mass. As for crack density it might be largely driven by the combination of Lx and cut_2 values.

3.4. Comparisons of the optimized populations between scenarios

The optimization step was performed under four scenarios contrasting for cultural practices. The phenotypes of the resulting populations are presented in Fig. 8. The global impact of crop load appeared huge on fruit mass and crack density and moderate on sweetness. Irrigation regime displayed a small impact on the three criteria in case of low crop load but gave rise to contrasting phenotypes in case of high crop load. Fruits were largely smaller, sweeter and less cracking when produced in WD_HC compared to WD_LC scenario. Beyond affecting peaks of distributions, constraint scenarios (WD and HC) also resulted in a shift of the distribution towards small values of fruit mass and crack density. Their impact on sweetness was not constant and depended on each particular practice. The distribution was shifted by high crop load towards poor sweetness only in case of well-irrigated regime. In case of water deficit regime and high crop load (WD_HC) small values of sweetness were not observed.

Data points of the optimized populations simulated in the three constraint scenarios (WI_HC, WD_LC and WD_HC) were added as supplementary variables on the first two PC plan of the PCA analysis (Fig. 9b, c, d) carried out on the criteria of the optimized population in the WI_LC scenario presented in section 3.3 (Fig. 9a). The dispersion of solutions on the plan was reduced for the three constraint scenarios compared to WI_LC. The shape and the position of the cloud observed for WD_LC was similar to that for WI_LC, but clouds for the two high crop load scenarios (WI_HC and WD_HC) were largely smaller and condensed towards the bottom-left of the plan. The impact of water deficit was small when crop load was low but much more pronounced in case of high crop load. Similarly, the impact of crop load level, reduced in case of well-irrigated regime, turned out very large when there was a water stress. These results demonstrate a strong interaction between the two types of cultural practices tested.

4. Discussion

In this paper we proposed a general framework associating a process-based model, the 'Virtual Fruit' and a multiobjective evolutionary algorithm in order to identify individuals with a good trade-off between three fruit traits and adapted to specific cultural practices.

4.1. A promising optimization method

The NSGA-II algorithm was chosen to generate combinations of six parameters of the 'Virtual Fruit' model that maximise fruit mass and sweetness and minimize skin crack density. Thus, if one single parameter has combined effects on the phenotype, with positive influence on some traits and negative on others, the algorithm can help to find the best compromise. Moreover, the algorithm provides a large diversity of solutions which represent a different compromise among design objectives.

The algorithm proved to be well adapted to solve a multiobjective problem with highly contradictory criteria. Indeed, the optimized solutions largely overcame non optimized solutions obtained from random sampling of parameters (Fig. 5). Results showed that most of the optimized solutions were distributed along Pareto fronts suggesting a good convergence of the algorithm towards best zones. In addition, some solutions were located in non-crowded zones, which constituted some singular alternatives displaying different compromise between the criteria. Although NSGA-II is one of the most efficient and reliable current multiobjective optimization algorithms, some weaknesses of this algorithm have been reported in the literature (Wang et al., 2010). They pointed out that the crowding distance does not accurately reflect the actual crowding degree for a given solution. This measure could lead to preserve a lot of solutions in some region of the search space while having only few solutions in other regions.

In the near future, we should improve the stopping criterion of NSGA-II which may reduce computational time. We also plan to introduce the suggestions proposed by Wang et al. (2010) to improve the crowding measure and the selection strategy of NSGA-II. In addition, we are currently working with alternative multiobjective algorithms offering complementary advantages in the successive phases of the optimization procedure. The final aim is to build a hybridization method by combining different algorithms.

4.2. Characteristics of the optimized solutions proposed to the decision-maker

The optimized solutions represented a real improvement in comparison to non-optimized solutions. The set of solutions proved to be highly contrasted with different types of solutions depending if one of the three criteria was favored or if a balance between the three criteria was reached. Therefore, within the set of solutions, the decision-maker can choose the best suited trade-off adapted to a particular objective. Indeed, in a current breeding scheme, fruit mass would be the only criteria considered. However, alternative schema could be considered for future, favoring organoleptic quality or environment friendly practices. In those cases, some interesting optimized solutions were identified with high level of sweetness and low crack density.

4.3. Best combinations of parameters to improve peach fruit

The solutions were obtained through simulations of the 'Virtual Fruit' model with the optimized combination of six parameters. Among the parameters with significant impact on the model outputs, these six parameters were selected to be involved in each of the six different processes. This option was followed in order to emphasize the complex and intricate

functioning of the model and the influence of each process on the fruit elaboration. However, three parameters (λ_{ph}, $share_1$ and r_1) proved to be of low interest. Indeed, they do not drive the phenotypes towards a particular type of solution and their optimization may be linked to local improvement of the solutions. Consequently, one can suggest fixing these three parameters and no longer including them in the optimization procedure for further analyses. However, it would be worth checking previously the accuracy of their values as well as their genetic variability in the species. Indeed, it might be an opportunity for the breeders to target the best values identified or even more extreme than the ones considered in this study.

4.4. Genotypes adapted to a target environment

The optimization step was performed under four scenarios contrasting for cultural practices. The distribution of the genotypes (set of the six parameters) in the optimized populations was almost comparable between scenarios (see section 3.1, Fig. 4). The only major difference was noticed for the initial relative fruit growth rate (RGR_{ini}) which displayed lower values in low than in high crop loads. Despite similar parameter values inside the four populations, the resulting phenotypes simulated by the 'Virtual Fruit' were different, which emphasizes the impact of conditions of growth on the phenotypes. This suggested that goals in terms of criteria might not be similar whatever the conditions of growth. Some excellent solutions cannot be reached in some particular scenario and one might adapt his goal to the type of cultural practices that can be carried out. The same way, it might be necessary to adapt the genotypes to the management of the orchard. To go further in this way, a more precise comparison of the genotypes is needed to identify those adapted to different scenarios or specifically adapted to a given one. Indeed, distributions of parameter values were similar between scenarios but the combinations of the six parameters might be slightly different. The analysis of interactions between genotype and cultural scenario is currently in progress by simulating the phenotypes of each population optimized in a given scenario in the conditions of the three other scenarios. This will give further insight into the behaviour of a given genotype in different cultural conditions Besides the cultural practices, it is important to also consider the impact of the climatic environment on the expression of the phenotypes. The work in this study has been performed under a constant environment (year 2009 in Avignon, France), as the ones carried out by Letort et al. (2008) and Qi et al. (2010). However, it appears essential to extend the work and consider various climatic environments as done by Chenu et al. (2009) for example. Indeed, the challenge for the future is to improve performance of crops across different environments. This can be achieved by defining target populations of environment as outlined in detail by Messina et al. (2009). Considering our work, further methodological developments are needed in the optimization side to reduce computational time in order to be able to consider multi-environments and large climatic series.

4.5. From model parameters to genetic parameters

Ecophysiological models are increasingly expected to include genetic information (Boote et al., 2001). To fulfil this demand, one way is to include genotype-dependent parameters that may be stable across environments (Yin et al. 2000; Tardieu 2003; Hammer et al. 2006). In this case, the set of genetic parameters is a fingerprint of the genotype. Many studies have now proved the interest of this approach but they also enhanced the necessity to build up more sophisticated model versions, in order to fully unravel the genetic and environmental effects. In this paper, we developed a methodology to nest the 'Virtual Fruit'

model with an efficient optimization algorithm. We studied six parameters of the model in order to illustrate the interest of the approach to develop new breeding strategies based on the guidance from optimization techniques. The results are rather preliminary, since the parameters studied were not chosen on the basis of their genetic features. However, the previous studies of Quilot et al. (2005a and b) led us to consider a possible genotypic determination of some parameters of the 'Virtual Fruit'.

4.6. Towards model-based breeding

Our results provide a picture of the optimized space of solutions considering the functioning of the system driven by biophysical constraints only. Presently an optimized population represents ideal genotypes the breeder is not sure to be able to create. Thus priorities in our perspectives of work are given to the integration of genetic constraints to the scheme. Two options can be considered. First, known genetic constraints may be included in the definition of the space of variation of the parameters to be explored during the optimization step. Non continuous space and links between parameters could be added as constraints of optimization at this step. Then, from the parameter values associated allelic combinations may be inferred. This option was tested by Letort et al. (2008). The second option is based on the direct optimization of the allelic combinations. This option requires that genetic information is combined to the process-based model and may allow considering complex genetic models. Some steps have already been made towards this integration and the resulting models are referred as gene-to-phenotype or QTL-based models. We started using a simplified version of the 'Virtual Fruit' to describe genetic variations within a mapping population (Quilot et al., 2005a) and replaced parameters by the effects of quantitative trait loci (QTL) (Quilot et al., 2005b). In the model classification proposed by White and Hoogenboom (2003) the resulting QTL-based model belongs to the level 4 class: 'genetic differences represented by specific alleles, with gene action represented through linear effects on model parameters'. Several studies have followed the same approach with different systems and models (Yin et al., 1999; Buck-Sorlin and Bachmann, 2000; Reymond et al., 2003; Hammer et al., 2005) and this level 4 class has now concrete applications. Such a gene-to-phenotype model was used by Chenu et al. (2009) to explore the genotype x environment interactions and the complexity of the results highlighted the importance of genetic architecture in the generation of phenotypes.

However, to move from theoretical estimation to practical result is far from straightforward since many QTLs are likely to have limited utility in applied breeding because of their dependency on genetic background or their sensitivity to the environment (Chapman et al., 2003; Hammer et al., 2004; Podlich et al., 2004). Finally, work has to be done now to take into account more complex genetic models and to evaluate their predictive capacity against empirical data. Consequently, we can consider that the use of process-based models to help breeding by optimisation has never been stepped over since the only work published so far was based on virtual genes and virtual chromosomes (Letort et al., 2008). This study together with those of Letort et al. (2008) and Qi et al. (2010) represent a proof of concept of the approach linking process-based models and optimization methods and will hopefully give rise to more complete studies.

5. Conclusion

We linked a process-based model to a multiobjective optimization method to design genotypes adapted to cultural scenarios. The use of an advanced optimization method, the NSGA-II algorithm, proved to be highly efficient for this purpose. This study provides a general framework for further works which would highly benefit from the integration of genetic knowledge. Besides, the results obtained with contrasting cultural scenarios encourage extending the approach to the consideration of contrasted climatic conditions. Similarly, the optimization procedure could be expanded to the optimization of the cultural practices. All together, these forthcoming works would contribute to the design of sustainable production systems combining genotypes and specific practices adapted to a targeted environment.

Appendix A
The crossover operator used by NSGA-II:

First, NSGA-II generates a random u_i number in [0,1] interval. Then, it uses a specified probability distribution to calculate βq_i so that the cumulative probability from 0 to βq_i is equal to u_i. The probability distribution used is the following:

$$P(\beta) = \begin{cases} 0.5(\eta_c + 1)\beta_i^{\eta_c} & \text{if } \beta_i \leq 1 \\ 0.5(\eta_c + 1)\dfrac{1}{\beta^{\eta_c + 1}} & \text{otherwise} \end{cases}$$

So βq_i is given as follows:

$$\beta q_i = \begin{cases} (2u_i)^{\frac{1}{\eta_c + 1}} & \text{if } u_i \leq 0.5 \\ \left[\dfrac{1}{2(1 - u_i)}\right]^{\frac{1}{\eta_c + 1}} & \text{otherwise} \end{cases}$$

The children y^1, y^2 are then calculated from parents x^1, x^2 using the following equations:

$$y_i^1 = 0.5\left[(1 + \beta q_i)x_i^1 + (1 - \beta q_i)x_i^2\right], \; y_i^2 = 0.5\left[(1 - \beta q_i)x_i^1 + (1 + \beta q_i)x_i^2\right]$$

Here η_c is the distribution index for crossover.

The mutation operator used by NSGA-II:

The NSGA-II uses polynomial mutation operator, this operator mutes solution x^1 into y^1 as follows:

$$y_i^{(1,t+1)} = x_i^{(1,t+1)} + (x_i^u - x_i^u)\delta_i$$

where t is the generations counter, x_i^u, x_i^l are respectively upper and lower bounds of x_i.

The parameter δ_i is calculated using the following polynomial probability distribution:

$$P(\delta) = 0.5(\eta_m + 1)(1 - |\delta|)^{\eta_m}$$

$$\delta_i = \begin{cases} (2r_i)^{1/(\eta_m + 1)} - 1, & \text{if } r_i \prec 0.5 \\ 1 - [2(1 - r)_i]^{1/(\eta_m + 1)}, & \text{if } r_i \geq 0.5 \end{cases}$$

Here η_m is the distribution index for mutation, r_i is a random number in [0,1] interval.

The crowding distance:

According to Deb et al (2002), the crowding distance of all solutions in a non dominated set is computed using the following procedure:

crowding-distance-assigment (F)

$l = |F|$ number of solutions in front F

for each i, set $F[i]_{\text{distance}} = 0$

for each objective m

$\quad F = \text{sort}(F, m)$

$\quad F[1]_{\text{distance}} = F[l]_{\text{distance}} = \infty$

$\quad$ for $i = 2$ to $(l-1)$

$\quad\quad F[i]_{dis\tan ce} = F[i]_{dis\tan ce} + (F[i+1].m - F[i-1].m) / (f_m^{\max} - f_m^{\min})$

where $F[i].m$ is the m-th criterion value of the i-th solution in the considered set F. The maximal the crowding distance is the better the solution is.

Appendix B. Supplementary material

Supplementary material associated with this article can be found in the online version.

References

Bandyopadhyay, S., Saha, S., Maulik, U., Deb, K. 2008. "A simulated annealing-based multiobjective optimization algorithm: AMOSA." Ieee T. Evolut. Computat. 12, 269-283.

Berman, M.E., DeJong, T.M. 1996. Water stress and crop load effects on fruit fresh and dry weights in peach (*Prunus persica*). Tree Physiol. 16, 859-864.

Boote, K.J., Kropff, M.J., Bindraban, P.S. 2001. Physiology and modelling of traits in crop plants: implications for genetic improvement. Agr. Syst. 70, 395–420.

Buck-Sorlin, G.H., Bachmann, K., 2000. Simulating the morphology of barley spike phenotypes using genotype information. Agronomie. 20, 691–702.

Chapman, S., Cooper, M., Podlich, D., Hammer, G., 2003. Evaluating plant breeding strategies by simulating gene action and dryland environment effects. Agron. J. 95, 99–113.

Chenu K., Chapman, S.C., Tardieu, F., McLean, G., Welcker, C., Hammer, G.L., 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.

Coello, C., Veldhuizen, D.V., Lamont, G. 2002. Evolutionary Algorithms for Solving Multi-Objective Problems. Eds David E. Goldberg and John R. Koza, New York, Springer.

Deb, K., 2001. Multi-Objective Optimization using Evolutionary Algorithms. Chichester, UK, Wiley.

Deb, K., Pratap, A., Agarwal, S., Meyarivan, T. 2002. A fast and elitist multiobjective genetic algorithm: NSGA-II. Ieee T. Evolut. Computat. 6, 182-197.

deVoil, P., Rossing, W.A.H., Hammer, G.L. 2006. Exploring profit-sustainability trade-offs in cropping systems using evolutionary algorithms. Environ. Model. Software 21, 1368-1374.

Escobar-Gutiérrez, A.J., Gaudillère, J.P., 1997. Carbon partitioning in source leaves of peach, a sorbitol-synthesizing species, is modified by photosynthetic rate. Physiol. Plantarum. 100, 353-360.

Escobar-Gutiérrez, A.J., Zipperlin, B., Carbonne, F., Moing, A., Gaudillère, J.P., 1998. Photosynthesis, carbon partitioning and metabolite content during drought stress in peach seedlings. Austr. J. Plant Physiol. 25, 197-205.

Feliciano, A., Feliciano, A.J., Ogawa, J.M., 1987. *Monilinia fructicola* resistance in the peach cultivar Bolinha. Phytopathology. 77, 776-780.

Fishman, S., Génard, M., 1998. Model of fruit growth based on biophysical description of main contributing processes. Simulation of seasonal and diurnal dynamics of weight. Plant Cell Environ. 21, 739-752.

Francisco, S.R., Ali, M., 2006. Resource allocation tradeoffs in Manila's peri-urban vegetable production systems: an application of multiple objective programming. Agric. Syst. 87, 147–168.

Génard, M., Lescourret, F., Gomez, L., Habib, R., 2003. Changes in fruit sugar concentration in response to assimilate supply, metabolism and dilution. A modelling approach applied to peach fruit *Prunus persica* (L.). Tree Physiol. 23, 373-385.

Génard, M., Gouble, B., 2005. ETHY. A theory of fruit climacteric ethylene emission. Plant Physiol. 139, 531-545.

Génard, M., Bertin, N., Borel, C., Bussières, P., Gautier, H., Habib, R., Lechaudel, M., Lecomte, A., Lescourret, F., Lobit, P., Quilot, B., 2007. Towards a 'Virtual Fruit' focusing on quality: modelling features and potential uses. J. Exp. Bot. 58, 917-928.

Génard M., Bertin N., Gautier H., Lescourret F., Quilot, B., 2010. Virtual profiling: a new way to analyse phenotypes. Plant J. 62, 344-355.

Gibert, C. 2007. Génération de fissures cuticulaires sur la pêche (*Prunus persica* (L.) Batsch) en réponse à des opérations culturales. Conséquences pour la qualité et la contamination par les monilioses. PhD Thesis, Université D'Avignon et des Pays de Vaucluse, Avignon (FR).

Gibert, C., Chadoeuf, J., Vercambre, G., Génard, M., Lescourret, F. 2007. Cuticular cracking on nectarine fruit surface: spatial distribution and development in relation to irrigation and thinning. J. Amer. Soc. Hort. Sci. 132, 583-591.

Gibert, C., Chadoeuf, J., Nicot, P., Vercambre, G., Génard, M., Lescourret, F., 2009. Modelling the effect of cuticular crack surface area and inoculum density on the probability of nectarine fruit infection by *Monilinia laxa*. Plant Pathol. 58, 1021-1031.

Gibert C., Génard M., Vercambre G., Lescourret F. 2010. Quantification and modelling of the stomatal, cuticular and crack components of peach fruit surface conductance. Funct. Plant Biol. 37, 264–274.

Goel, T., Vaidyanathan, R., Haftka, R.T., Shyy, W., Queipo, N.V., Tucker, K. 2007. Response surface approximation of Pareto optimal front in multi-objective optimization. Comput. Methods Appl. Mech. Engrg. 196, 879-893.

Gradziel, T.M., Thorpe, M.A., Bostock, R.M., Wilcox, S. 1998. Breeding for brown rot (*Monilia fructicola*) resistance in clingstone peach with emphasis on the role of fruit phenolics. Acta Hort., 465, 161-170.

Hammer, G. L., Kropff, M.J., Sinclair, T.R., Porter, J.R., 2002. Future contributions of crop modelling - from heuristics and supporting decision making to understanding genetic regulation and aiding crop improvement. Eur. J. Agron. 18, 15-31.

Hammer, G.L., Sinclair, T.R., Chapman, S.C., vanOosterom, E., 2004. On systems thinking, systems biology, and the in silico plant. Plant Physiol. 134, 909–911.

Hammer, G., Chapman, S., van Oosterom, E., Podlich, D., 2005. Trait physiology and crop modeling as a framework to link phenotypic complexity to underlying genetic systems. Aust. J. Agric. Res. 56, 947–960.

Hammer, G., Cooper, M., Tardieu, F., Welch, S., Walsh, B., van Eeuwijk, F., Chapman, S., Podlich, D., 2006. Models for navigating biological complexity in breeding improved crop plants**. Trends Plant Sci. 11, 587-593.**

Harrison, R.D., Daniell, J.W., Cheshire, J.M. Jr., 1989. Net photosynthesis and stomatal conductance of peach seedlings and cuttings in response to changes in soil water potential. Journal of the American Society for Horticultural Science. 114: 986-990.

Haverkort, A. J., Grashoff, C., 2004. IDEOTYPING-POTATO a modelling approach to genotype performance, in: MacKerron, D.K.L., Haverkort, A.J. (Eds), Decision support systems in potato production: bringing models to practice. Wageningen academic publishers, Wageningen, pp. 199-211.

Herndl, M., Shan, C., Wang, P., Graeff, S., Claupein, W., 2007. A model based ideotyping approach for wheat under different environmental conditions in North China Plain. ASC. 6, 1426-1436.

Isaacson, T., Kosma D.K., Matas, A.J., Buda G.J., He Y., Yu B., Pravitasari A., Batteas J., Stark R., Jenks M.A., Rose, J.K.C., 2009. Cutin deficiency in the tomato fruit cuticle consistently affects resistance to microbial infection and biomechanical properties, but not transpirational water loss. Plant J. 60, 363-377.

Jolliffe, I.T., 2002. Principal Component Analysis. Springer-Verlag, New York.

Lescourret, F., Ben Mimoun, M., Génard, M., 1998. A simulation model of growth at the shoot bearing fruit level. I Description and parameterisation for peach. Eur. J. Agron. 9, 173-188.

Lescourret, F., Génard, M., 2005. A virtual peach fruit model simulating changes in fruit quality during the final stage of fruit growth. Tree Physiol. 25, 1303-1315.

Letort, V., Mahe, P., Cournede, P.H., De Reffye P., Courtois, B., 2008. Quantitative genetics and functional-structural plant growth models: Simulation of quantitative trait loci detection for model parameters and application to potential yield optimization. Ann. Bot. 101, 1243-1254.

Lo Bianco, R., Rieger, M., Sung, S. S., 2000. Effect of drought on sorbitol and sucrose metabolism in sinks and sources of peach. Physiol. Plantarum. 108, 71-78.

Messina, C., Hammer, G., Dong, Z., Podlich, D., Cooper M. 2009. Modelling crop improvement in a G x E x M framework via gene − trait − phenotype relationships. Chapter 10. In Crop physiology. Applications for genetic improvement and agronomy. Eds Victor O. Sadras and Daniel F. Calderini, Elsevier Inc.

Morris, M.D., 1991. Factorial Sampling Plans For Preliminary Computational Experiments. Technometrics. 33, 161-174.

Naor, A., I. Klein, I. Doron, Y. Gal, Z. Ben-David, Bravdo, B. 1997. The effect of irrigation and crop load on stem water potential and apple fruit size. J. Hort. Sci. 72, 765–771.

Podlich, D.W., Winkler, C.R., Cooper, M., 2004. Mapping as you go: an effective approach for marker-assisted selection of complex traits. Crop Sci. 44, 1560–1571.

Qi, R., Ma, Y.T., Hu, B.G., de Reffye, P., Cournede, P.H., 2010. Optimization of source-sink dynamics in plant growth for ideotype breeding: A case study on maize. Comput. Electron. Agr. 71, 96-105.

Quilot, B., Genard, M., Lescourret F., Kervella, F., 2005a. Simulating genotypic variations of fruit quality in an advanced peach x *Prunus davidiana* cross. J. Exp. Bot. 56, 3071-3081.

Quilot, B., Kervella, J., Genard, M., Lescourret, F., 2005b. Analysing the genetic control of peach fruit quality through an ecophysiological model combined with a quantitative trait loci approach. J. Exp. Bot. 56, 3083-3092.

Raju, K.S., Kumar, D.N., 1999. Multicriterion decision making in irrigation planning. Agric. Syst. 62, 117–129.

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 Physiol. 131, 664–675.

R Development Core Team, 2008. R: A Language and Environment for Statistical Computing. Vienna, Austria: R Foundation for Statistical Computing (http://www.r-project.org).

Saltelli, A., Tarantola, S., Chan, K.P.S., 1999. A quantitative model-independent method for global sensitivity analysis of model output. Technometrics. 41, 39-56.

Saltelli, A., Ratto, M., Andres, T., Campolongo, F., Cariboni, J., Gatelli, D., Saisana, M., Tarantola, S., 2008. Global Sensitivity Analysis: The Primer. WileyBlackwell.

Tardieu F., 2003. Virtual plants: modelling as a tool for the genomics of tolerance to water deficit. Trends Plant Sci. 8, 9-14.

Wallach, D., Goffinet, B., Bergez, J.E., Debaeke, P., Leenhardt, D., Aubertot, J.N., 2001. Parameter estimation for crop models: a new approach and application to a corn model. Agron. J. 93, 757–766.

Wang, Z., Tang K., Yao, X. 2010. Multi-Objective Approaches to Optimal Testing Resource Allocation in Modular Software Systems. Ieee T. Evolut. Computat. 59, 563-575.

White, W.J., Hoogenboom G., 2003. Gene-based approaches to crop simulation: past experiences and future opportunities. Agron. J. 95, 52-64.

Yin, X., Kropff, M.J., 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., Chasalow, S.C., Dourleijn, C.J., Stam, P., Kropff, M.J. 2000. Coupling estimated effects of QTLs for physiological traits to a crop growth model: predicting yield variation among recombinant inbred lines in barley. Heredity. 85, 539-549.

Yin, X., Struik, P.C., Tang, J., Qi, C., Liu, T., 2005. Model analysis of flowering phenology in recombinant inbred lines of barley. J. Exp. Bot. 56, 959-965.

Table 1

Genetic parameters subjected to the optimization process. Reference values and boundary values of the range to be explored are given.

Parameters	Unit	Definition	Processes	Reference value	Minimum value	Maximum value
r_1	dimensionless	Leaf structural mass/leafy shoot structural mass	Vegetative structure	0.766	0.672 [a]	0.842 [a]
RGR_{ini}	degree-day^{-1}	Initial relative fruit growth rate	Fruit growth	0.0035	0.001 [b]	0.01 [b]
λ_{ph}	dimensionless	Proportion of carbon as sucrose in phloemic sap	Sugar metabolism	0.347	0.23 [c]	0.54 [c]
$share_1$	g	Empirical coefficient relating stone dry mass to fruit dry mass	Stone-pulp partition	6.78	6.10 [b]	7.46 [b]
cut_2	dimensionless	Parameter of the equation of the relative expansion rate of the cuticle surface area	Transpiration	1.26	1.134 [a]	1.386 [a]
Lx	g cm^{-2} bar^{-1} h^{-1}	Hydraulic conductance per unit of fruit surface	Water fluxes	0.2333	0.0769 [b]	0.667 [b]

[a] ± 10 % variations

[b] Observed variability (Quilot et al., 2005a)

[c] Extreme values from literature

Table 2

Correlations between parameters and variables for the WI_LC scenario

	r_1	RGR_{ini}	λ_{ph}	$share_1$	cut_2	Lx	Fruit mass	Sweetness
RGR_{ini}	0.01							
λ_{ph}	0.02 *	-0.21 ***						
$share_1$	-0.11 *	0.24 ***	-0.16 ***					
cut_2	0.01	0.43 ***	-0.25	0.28 ***				
Lx	0.04 *	**0.53** ***	-0.12 *	0.12	0.18 ***			
Fruit mass	0.09	0.38 ***	-0.10	0.08	0.20 ***	**0.86** ***		
Sweetness	-0.09	**-0.74** ***	0.19 ***	-0.21 ***	-0.48 ***	**-0.87** ***	**-0.77** ***	
Crack density	0.08	-0.11 *	0.14 **	-0.18 ***	**-0.60** ***	0.53 ***	**0.64** ***	-0.24 ***

*, **, *** significant at P≤0.05, P≤0.01 and P≤0.001, respectively.

Table 3
Particular solutions selected among the optimized population in the WI_LC scenario displaying contrasted trade-off between the three criteria. Optimized values for the 6 parameters and the corresponding simulated criteria are given.

Solutions	r_1	RGR_{ini}	λ_{ph}	$share_1$	cut_2	Lx	Fruit mass	Sweetness	Crack density
1	0.792	0.007	0.538	6.175	1.331	0.642	300.0	8.3	12.9
2	0.832	0.009	0.507	6.108	1.378	0.432	125.3	5.0	0.2
3	0.682	0.007	0.540	6.164	1.169	0.077	130.8	20.0	5.6
4	0.839	0.005	0.540	6.111	1.383	0.077	112.3	16.7	0.0
5	0.818	0.007	0.540	6.126	1.136	0.608	257.9	9.8	18.1
6	0.840	0.007	0.540	6.102	1.194	0.236	199.8	12.8	10.2
7	0.817	0.007	0.539	6.137	1.379	0.163	210.9	12.1	4.9

Fig. 1. Schematic representation of the relationships between sub-models as considered in the 'Virtual Fruit' (adapted from Bertin et al., 2010). The sub models simulate carbon balance of a fruit bearing stem, sugars metabolism within the fruit, fruit water balance, skin conductance and microcracking and fruit respiration. The inputs of the model are weather data, stem water potential, and the state of the system at the beginning of the simulation (number of leafy shoots and fruits, initial values of state variables). The outputs are flesh and stone masses, sugar, skin microcracking and emission of gases.
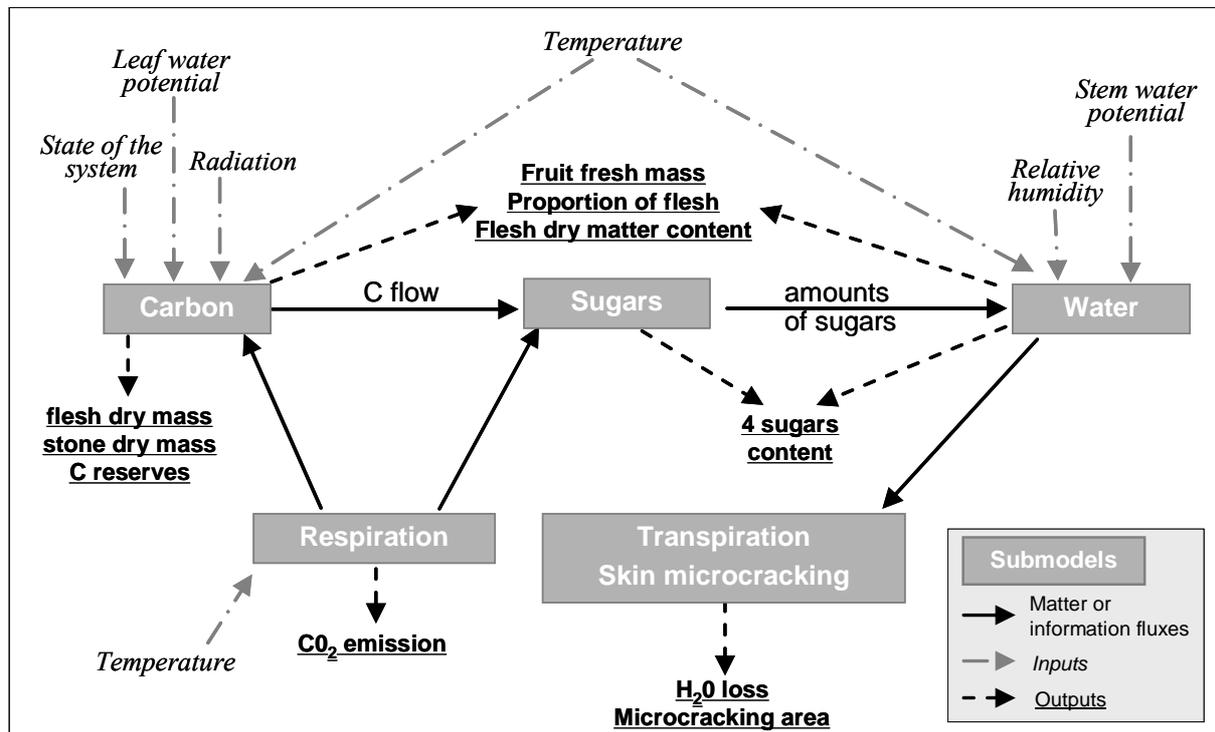
Fig. 2. Daily mean water potential of stem (a) and leaf (b) along fruit growth, entered as inputs in the 'Virtual Fruit'. Same pattern is used for the two well-irrigated (WI) treatments whatever the crop load but patterns depend on the crop load (low (LC) and high (HC)) for the water-deficit (WD) treatment.
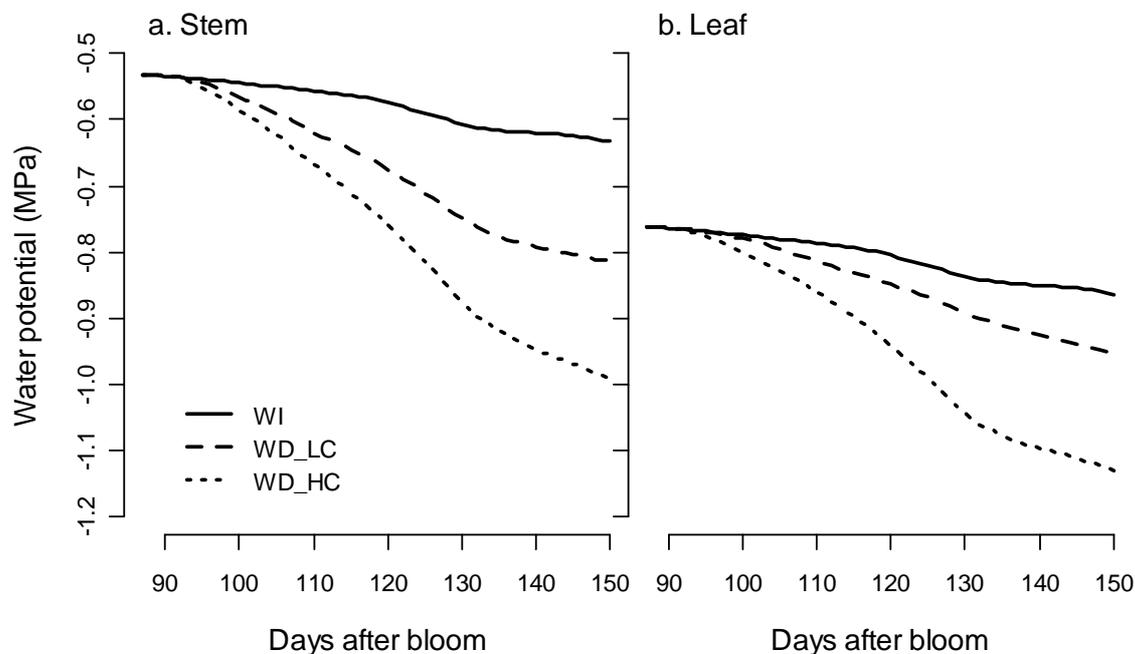
Fig. 3. Schematic representation of the procedure NSGA-II, adapted from Deb et al. (2002), implemented to design peach genotypes
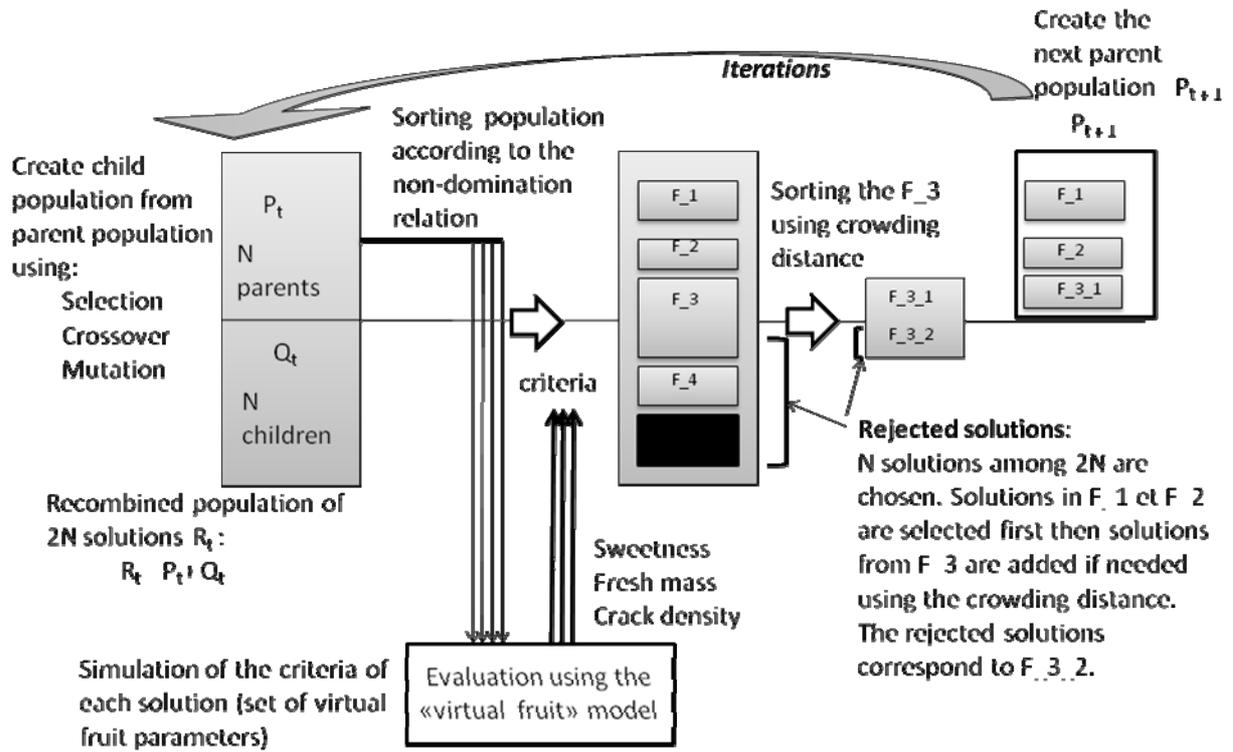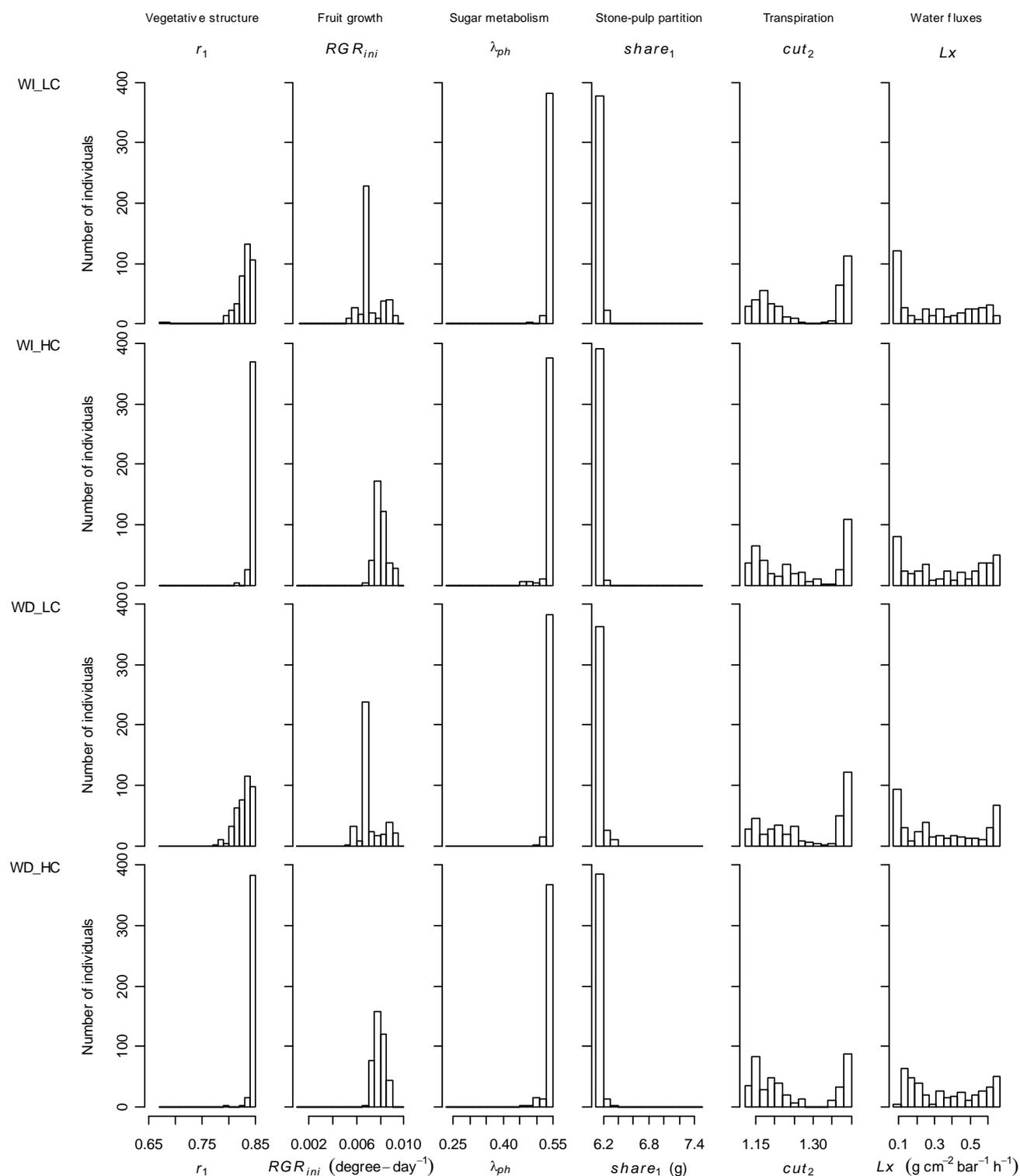
Fig. 4. Distribution of the values of the 6 parameters within the optimized populations (400 genotypes) for each of the 4 scenarios.

Fig. 5. Repartition of 400 random individuals (grey square) and the 400 optimized genotypes (black circle) on the first two PC plan of a PCA analysis carried out (a) on the parameter values and (b) on the simulated criteria of the random population, for the WI_LC scenario only. The optimized population was added as supplementary data on the PC plan. The random population was built by sampling at random the 6 parameters 400 times, within their range of variations and the corresponding criteria values obtained by running the corresponding simulations.
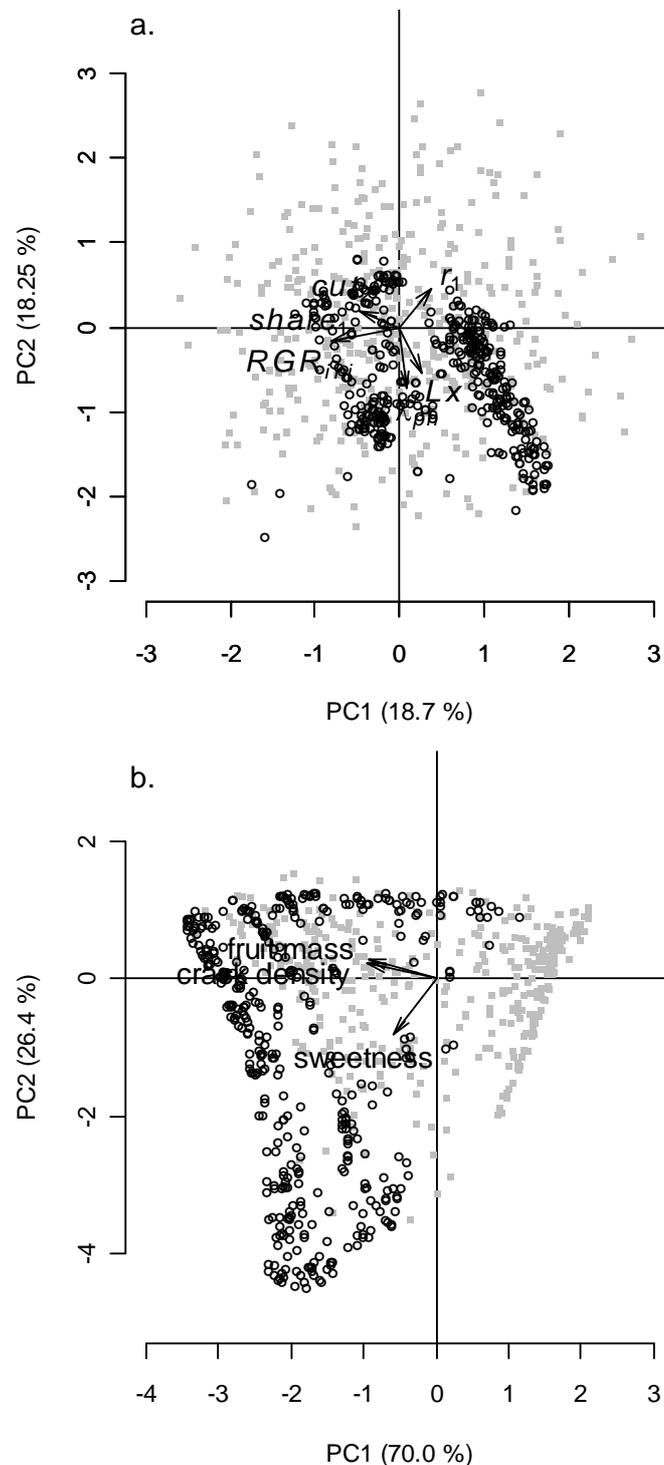
Fig. 6. Relationships between fruit mass and sweetness or crack density for the optimized population (400 phenotypes) in the WI_LC scenario. Colors (symbols) from light grey (square), grey (circle) to black (triangle) describe increasing values of (a) crack density, (b) sweetness and (c) fruit mass, as mentioned in the title. Numbers refer to the particular solutions display in table 3
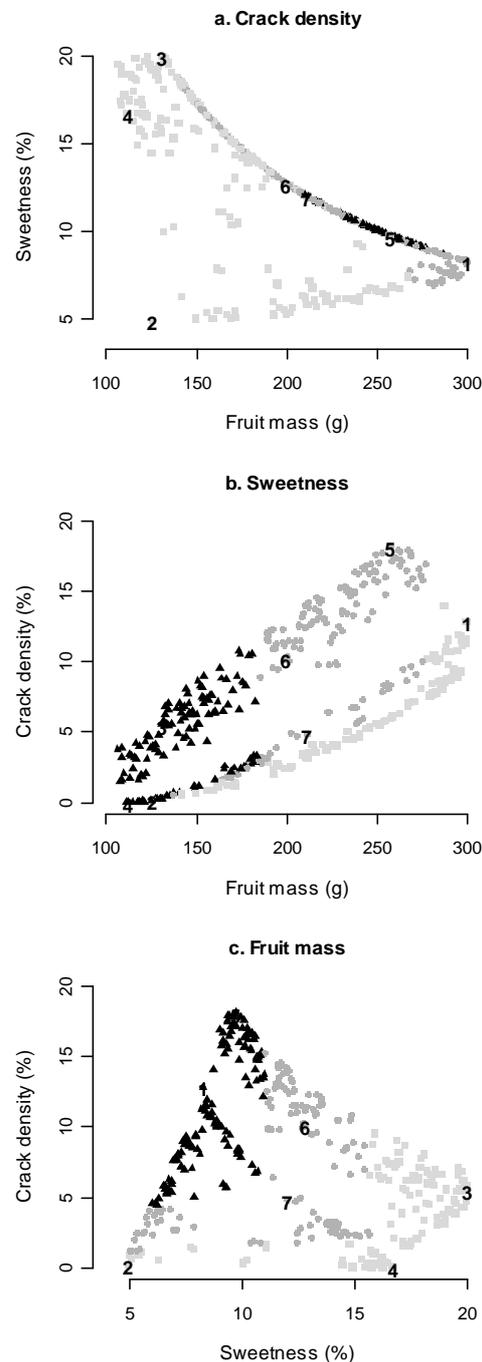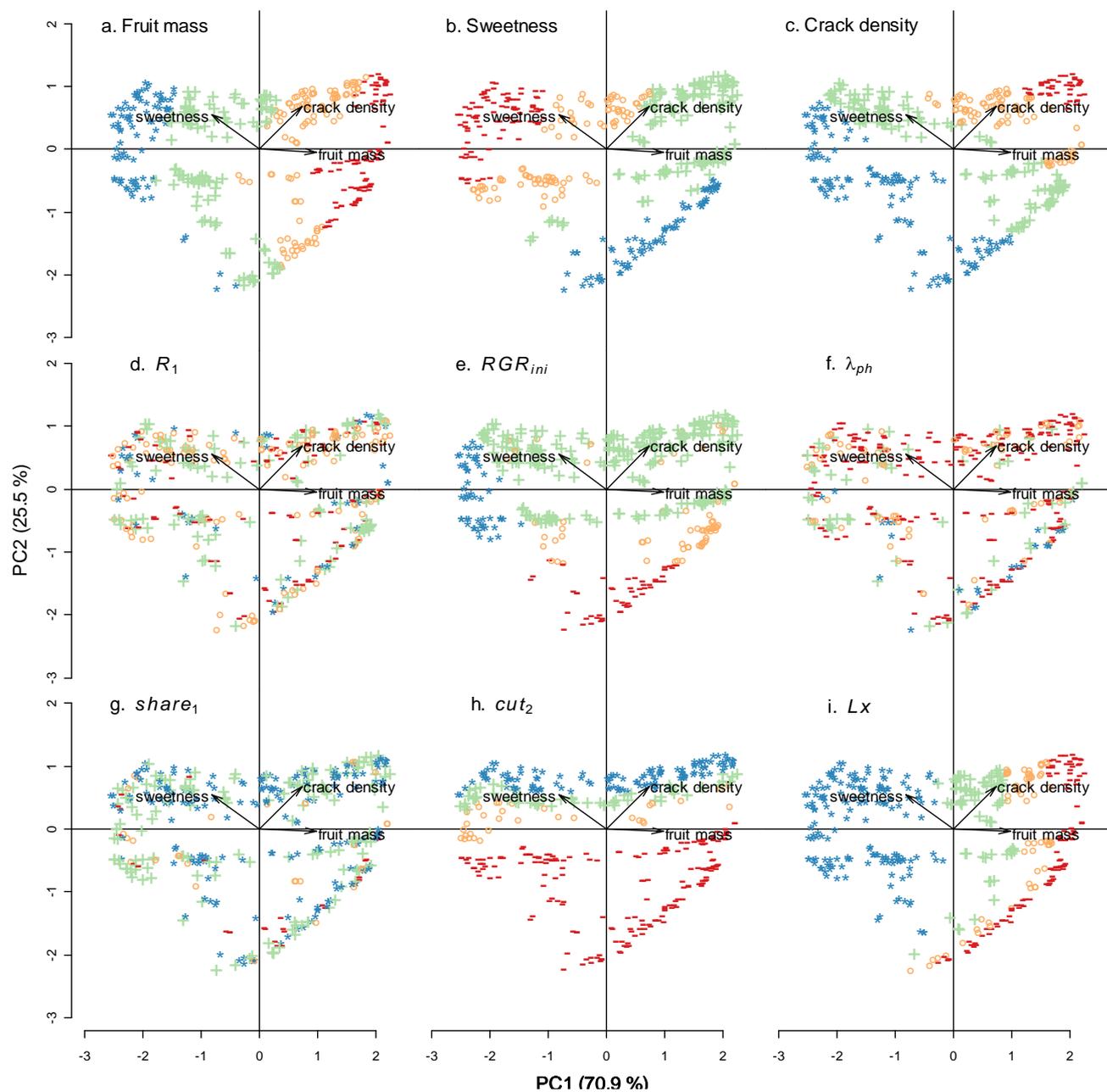
Fig. 7. Repartition of the ideotypes according to the parameter and variable values on the first two PC plan of a PCA analysis carried out on the criteria simulated for the optimized populations (400 phenotypes) in the WI_LC scenario. Colors from blue (star), green (cross), orange (circle) to red (dash) describe increasing values of the parameter or criteria mentioned in the title of each plot.

Fig. 8. Distribution of the three criteria (a. fruit mass, b. sweetness and c. crack density) within the optimized populations (400 phenotypes) for each of the 4 scenarios (well irrigated (WI); water deficit (WD); low (LC) and high (HC) crop loads).
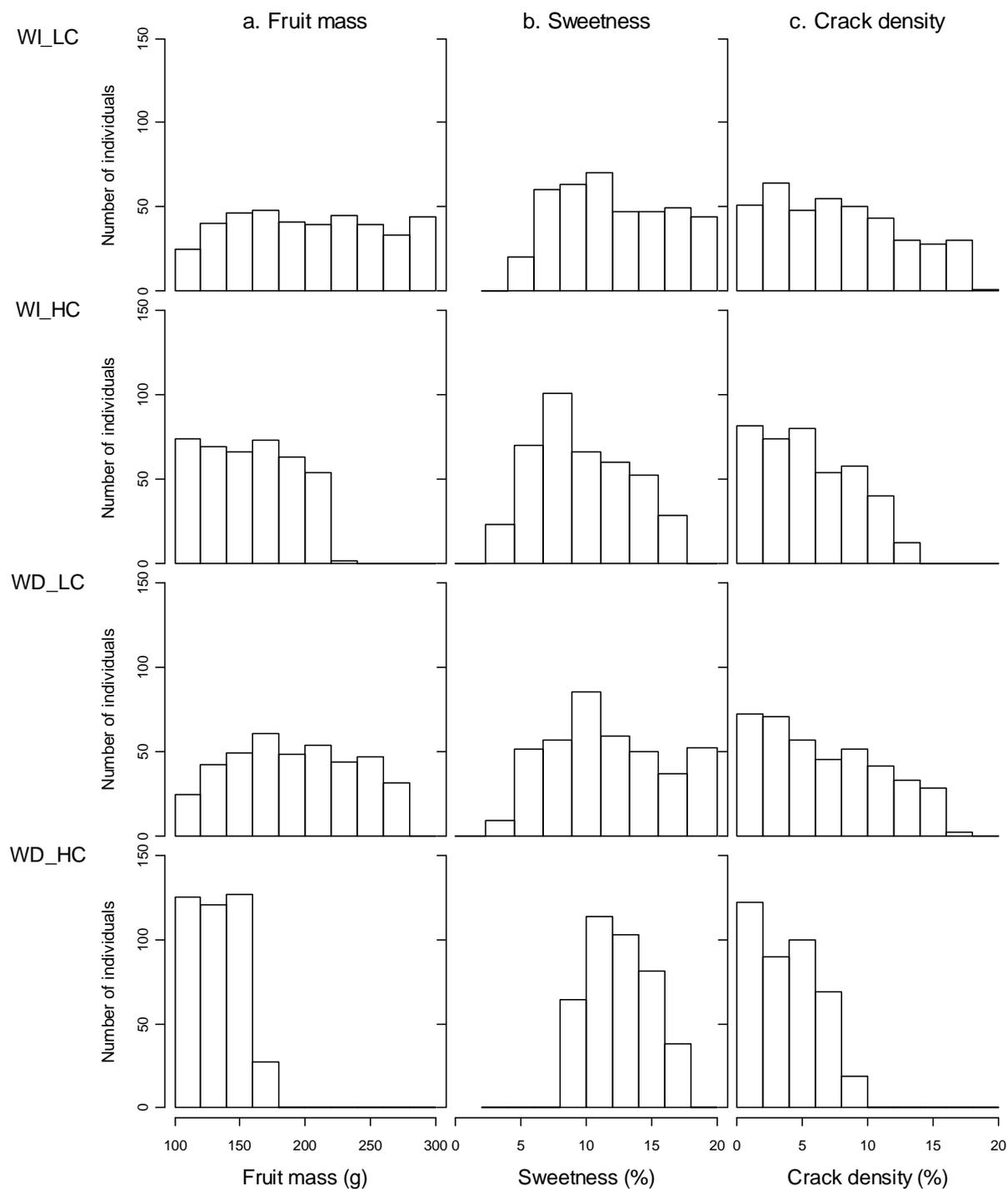
Fig. 9. Dispersal of the optimized populations (400 phenotypes) for each of the 4 scenarios added on the first two PC plan of a PCA analysis carried out on the criteria of the optimized population in the WI_LC scenario only.