On the Planning of Healthy and Balanced School Lunches through Multi-Objective Evolutionary Algorithms

Gara Miranda, Eduardo Segredo, Juan-Manuel Ramos-Pérez, Coromoto León, Casiano Rodríguez

Dpto. de Ingeniería Informática y de Sistemas, Universidad de La Laguna
Avda. Astrofísico Francisco Sánchez, s/n, 38200, San Cristóbal de La Laguna, Spain

---

Abstract

In this paper, we propose a novel multi-objective formulation for the *Menu Planning Problem*, which we have termed the *Multi-objective Menu Planning Problem*. Menu planning is of great interest in the health field due to the importance of proper nutrition in today society, and particularly in school cafeterias. Most of the related literature on menu planning deals with single-objective formulations of the problem, mainly focused on minimising the menu cost while satisfying a set of constraints. In this paper, in addition to considering cost as an objective, we also introduce a second objective function aimed at minimising the degree of repetition of the meals that shape the final plan. Generally speaking, people need non-repetitive meals, not only for health reasons, but also in order to avoid detesting certain meals. As a result, the motivation behind our novel formulation is to offer meal planning that is not only affordable, but also varied and balanced from a nutritional standpoint. The planning is designed for a given number of days and ensures that the particular nutritional requirements of children are satisfied. Furthermore, other special circumstances such as allergies, food intolerances, and specific constraints regarding religion or lifestyle habits, for instance, may be considered as well. However, since the aforementioned features are more focused on a particular person rather than on a group of people, their study is beyond the scope of this work. The computational results obtained by a set of well-known multi-objective evolutionary algorithms show that suitable meal plans, in terms of nutritional values, are achieved. In addition, the multi-objective nature of our novel formulation is also demonstrated. We show how the cheapest meal plans are those with the lowest variety of meals, while the largest variety of meals is given by the most expensive plans. Finally, we note that the novel formulation may be adapted, with little effort, to deal with the specific nutritional requirements of other groups of people, such as patients in hospitals and convicts in prisons, among others.

Keywords: Menu planning, multi-objective optimisation, evolutionary algorithms, nutrition and food science

1. Introduction

One of the main problems nowadays is the inappropriate eating habits exhibited by part of the population [1]. The hectic lifestyle of modern society has raised the consumption of fast food and snacks. The increased presence of this food type in our diet has led to its consumption even when there is no real need related to lack of time. In addition, the low frequency of physical activity induces high levels of excess weight and obesity in the population, which have continued to increase in recent years. Moreover, the habit of eating away from home can result in higher energy and cholesterol intakes, mostly from saturated fats, trans fatty acids, added sugar and sodium [3, 20]. On the other hand, the economic crisis in many developed countries has increased the poverty rates of the population, with consequent malnutrition problems. This problem of malnutrition, the lack or excess of nutrients, is especially serious when it arises in children. It is vital, therefore, to have a healthy and balanced diet [39].

Controlling the feeding of children at home is quite an impossible task for us. However, educational institutions can take steps to improve the nutrition of children in school since a vast majority of children have their lunch in school cafeterias. In general, a balanced diet should be manually designed and reviewed by a nutrition expert. An expert—trained professionally in the field of nutrition or food science—is able to define specific meals that are ideally suited to a person nutritional requirements. However, by considering general criteria for children nutrition and a predefined set of meals, it is possible to properly formulate the problem to automatically plan school menus.

In this work, we propose a novel multi-objective formulation for the *Menu Planning Problem* (MPP) [30], which we will refer to as the *Multi-objective Menu Planning Problem* (MMPP). This initial situation entails a series of specifics that make the problem differ from previous single- and multi-objective definitions of the MPP. Firstly, school cafeterias are in use mainly at lunchtime, so our formulation only focuses on recommending lunch menus, thus excluding the rest of the day meals. Secondly, the target audience for this planner are children ages 4 to 13. The recommended daily amount of nutrients for children of these ages differs considerably from that for adults. However, the amount of nutrients recommended for children between 4 and 13 does not differ considerably by gender, according to data from the *White Book on Child Nutrition*[1] of the *Spanish Association of Paediatrics* and the *Spanish Nutrition Foundation.* In addition, a gender distinction is not made in practice in school cafeterias, and meals are provided indistinctly to boys and girls. At the same time, unlike most conventional MPP formulations, the one we are proposing is intended for groups of people, rather than individuals. Consequently, gathering physical

[1]This document can be accessed through `https://www.aeped.es/sites/default/files/documentos/libro_blanco_de_la_nutricion_infantil.pdf`.

data, nutritional goals and user preferences in an effort to design a personalised plan—one that is more in keeping with a particular user needs and features—does not make sense in this case.

Some other aspects [11] that are usually taken into account in other formulations, such as the meals features, their quality and temperature and so on, and even the preparation time, will not be considered in this paper since the meals are prepared for large groups of people by cooking professionals. Instead, some of the characteristics that will be considered are those related to the cost, variety and nutritional quality of meals. In this paper, the cost of the menu and the variety of the meals are taken as the objective functions of the problem. The goal is to minimise the cost of the menu while also minimising how often its meals are repeated. The nutritional features of meals and menus are handled as the problem constraints, i.e., any feasible menu must satisfy a set of nutritional requirements and healthy food recommendations.

Bearing the above in mind, the main contributions of this work are as follows:

- A novel, multi-objective formulation of the MPP, termed MMPP, is proposed that not only considers the cost of the menu, but also the degree of repetition of the meals in the plan, as the two objective functions to be minimised. The particular formulation focuses on automatically designing lunch menus for children in school cafeterias.
- An extensive experimental evaluation is conducted in which well-known *Multi-Objective Evolutionary Algorithms* (MOEAs) are applied to solve different instances of the MMPP.
- Quantitative and qualitative analyses of the resulting meal plans are provided that show their suitability from the standpoint of their nutritional value.
- The multi-objective nature of the novel MMPP is demonstrated.

The remaining content of this work is structured as follows. A review of the literature on the MPP is presented in Section 2. This review focuses mainly on the existing alternatives for the MPP formulation—including single-objective and multi-objective definitions—and the types of methods applied in their solution. Our novel formulation of the MMPP is introduced in Section 3. Then, the experimental evaluation carried out is described in Section 4, which is also devoted to presenting and exhaustively analysing the computational results obtained. Finally, the conclusions and some lines of further research are given in Section 5.

2. State of the art on menu planning

The MPP [30] involves generating a diet plan, usually daily, weekly or monthly, consisting of healthy menus for the main meals of the day. The different meals that comprise every menu in the plan should be fully specified. For example, for lunch, we could specify the starter, the main course, the dessert, the drink and the side dishes. Some interesting considerations in the problem field involve the design of approaches that are not only limited to offering nutritionally adequate diet plans, but that are also able to adapt these plans to the specific needs of the user: personal preferences on the type of food, cost of the menus, variety of the meals, aesthetic characteristics of the food, preparation methods, incompatibilities with certain foods due to allergies, intolerances, diseases or specific lifestyles, and the quality of the food based on its seasonality, among others [30]. In this sense, the problem could involve many other objectives or constraints beyond offering a healthy menu, and consequently, it is necessary to define which objectives are going to be considered or optimised, and which is the set of constraints to be fulfilled by the feasible solutions.

The issue of what aspects to consider as problem objectives and what to consider as constraints may differ among authors. However, there is a clear tendency to consider the cost of the menu as an objective to be minimised [12, 14, 17, 32]. Other authors—who do not take into account the cost of the menu—usually include the user preferences for certain foods, the level of adequacy or the level of acceptance as the objective to be optimised [5, 9, 15, 23].

The most important and common consideration for menu planning problems, in general, is to ensure a healthy and balanced menu from a nutritional point of view. These nutritional requirements are usually grouped into multiple constraints, since they must provide the recommended minimum and maximum amounts of different nutrients [1, 12, 13, 14, 17, 23]. Other common constraints are the variety of the meals, the time required to prepare them, and food that cannot be consumed, among many others [1, 9, 14].

Finally, depending on the level of completeness that is desired, there may be many other aspects to consider. First and foremost is the user data or personal preferences. If a meal plan has to be designed for a particular person, additional features such as their weight, height, physical activity level or health goals could be considered in order to produce a more specific plan [5, 13, 15]. However, that information does not make sense when planning menus intended for groups of people, which is the goal of the novel formulation we are proposing. Another important aspect not taken into consideration in many of the papers reviewed in the literature is whether the user in question suffers from any type of allergy, intolerance, or if their lifestyle simply prevents the person from consuming certain foods [5, 15]. Whether or not the user belongs to one of these groups should be considered as a specific constraint so as to exclude or include the corresponding food in the meal plan.

Authors in the field have proposed different single- and multi-objective formulations of the MPP and have applied different types of meta-heuristics and optimisation methods to obtain a solution [30]. Following an exhaustive review of the literature, Table 1 shows a summary of the most relevant works, including information about the number of objective functions addressed and the types of optimisation methods applied.

As we can see, most existing works in the literature use single-objective optimisation methods to tackle the problem. Among these methods, *Evolutionary*

Table 1: Summary of most relevant works related to the MPP classified according to the number of objectives addressed - Single-Objective (SO) or Multi-Objective (MO) - and the types of solvers. The following approaches have been proposed/applied: *Genetic Algorithm* (GA), *Estimation of Distribution Algorithm* (EDA), *Bacterial Foraging Optimisation Algorithm* (BFOA), *Linear Programming* (LP), *Mixed Integer Linear Programming* (MILP), *Case-based Reasoning* (CBR), *Rule-based Reasoning* (RBR), *Fuzzy Mathematical Programming* (FMP), *Type-2 Fuzzy Ontology* (T2FO), *Mathematical Programming* (MP), *Belief Merging* (BM), *Decision Trees* (DT), and *Multi-Objective Evolutionary Algorithm* (MOEA)

	Meta-heuristics		**Other methods**	
	Reference	**Approach**	**Reference**	**Approach**
SO	Bulka et al. [4]	GA	Stigler [38]	LP
	Gaál et al. [10]	GA	Leung et al. [28]	MILP
	Kahraman and Seven [17]	GA	Petot et al. [33]	CBR and RBR
	Wang et al. [40]	GA	Khan and Hoffmann [24]	CBR
	Kashima et al. [23]	GA	Alfaro [2]	MILP
	Osthus [32]	GA	Noah et al. [31]	RBR
	Funabiki et al. [9]	Greedy	Kashima et al. [22]	FMP
	Gumustekin et al. [12]	EDA, GA	Aberg [1]	Branch&Bound
	Isokawa and Matsui [15]	GA	Lee et al. [26]	T2FO
	Hernández-Ocaña et al. [13]	BFOA	Hsiao and Chang [14]	Branch&Bound
			Jothi et al. [16]	CBR
			Kovásznai [25]	CBR
			Chávez et al. [5]	MP and BM
			Kale and Auti [19]	DT
MO	Kaldrim and Köse [18]	MOEA		
	Seljak [36]	MOEA		
	Moreira et al. [29]	MOEA		

Algorithms (EAs), and more particularly, *Genetic Algorithms* (GAs) are very popular [4, 10, 12, 15, 17, 23, 32, 40]. In some works, several objectives are taken into account, although i most cases multiple objectives are reduced to a single-objective function during the optimisation phase. In spite of the above, we have found three references where multi-objective formulations of the MPP are addressed by multi-objective optimisers. In all of them, MOEAs are applied.

In Seljak [36], a multi-objective variant of the MPP was solved by applying the *Non-dominated Sorting Genetic Algorithm II* (NSGA-II) [6]. The authors proposed a weekly plan consisting of seven daily menus with five meals each. The cost, seasonal quality and other aspects related to the food, such as its flavour and its temperature, for instance, were considered as objectives, while the users personal preferences and nutritional requirements were managed as constraints. Standard and general information for an adult person was considered.

The paper by Kaldrim and Köse [18] also applied the NSGA-II to tackle a multi-objective formulation of the MPP. This formulation considered the cost and personal preferences for food as objectives. It also took into account the gender and age of the user to generate a menu suited to said parameters.

A more recent work that also solves the problem through meta-heuristics and considers a multi-objective formulation is the paper written by Moreira et al. [29]. Once more, the NSGA-II was used to minimise the cost of the menu while also minimising the nutritional error in Brazilian schools, i.e., for children between four and five years old, based on the daily nutritional requirements set by the government. In addition to the application of the NSGA-II, they also solved a single-objective formulation of the problem by means of a GA.

Although they claimed that both algorithms met the objectives of the problem, they emphasised that the best results were provided by the NSGA-II.

3. Our formulation of the multi-objective menu planning problem

In a *Multi-objective Optimisation Problem* (MOP), the objective functions are usually in conflict with one another. This means that all the objective functions cannot be optimised simultaneously; consequently, an improvement in one of the objectives usually involves a deterioration in the remaining ones. In this paper, we propose a novel multi-objective formulation of the MPP, which we term the MMPP.

One of the main goals herein is to automatically design meal plans for school cafeterias; therefore, only lunch menus are considered through the plan, which excludes all other daily meals. Moreover, the plans are designed for large groups of children. As previously stated, the recommended daily amount of nutrients for children between 4 and 13 years old does not differ significantly by gender, and in practice, such a distinction is not made in school cafeterias. As a result, it will not be considered herein.

Regarding the problem objectives to optimise, two conflicting functions are considered. The first objective is to minimise the plan cost. In school cafeterias, meal prices should be reasonable. We should note at this point that the total cost of a meal plan is calculated as the sum of the costs of its meal. Moreover, the cost of a meal is calculated as the sum of the costs of all the ingredients—and their respective amounts—required to prepare that meal.

The second objective is the the minimisation of the degree of repetition of the meals included in the plan. The variety of the meals included in a menu designed for a given number of days—usually, weekly or monthly—is mandatory. Children need a varied diet, not only for health reasons [11, 39], by including in the menu all the food groups required from the nutritional standpoint, but also to avoid detesting certain foods, by varying the specific courses included in the plan. Studies have shown that dietary variety and/or diversity are healthy and related to diet quality [7, 8, 21]. Eating a more varied diet is associated with a higher intake of macro and micronutrients, as well as higher nutritional adequacy and diet quality [37].

The above is the main motivation behind the second objective function proposed, which was designed by considering advice given by nutrition experts from the *Intervention Program for the Prevention of Childhood Obesity*[2], endorsed by the *Health Service* of the *Government of the Canary Islands*. As far as we know, the diversity or variety of the menus that a meal plan consists of has only been considered as a constraint in past research [1, 14]. Typically, the repetition of, either some food groups, or specific courses, in the menu involves a penalty value which increases with the degree of repetition defined. Therefore, a non-feasible solution would be that one whose penalty value is large enough depending on

[2]The information of this program can be accessed through `http://www.programapipo.com/`.

the particular constraint established. In addition to the above, the repetition of courses among different days is not usually taken into consideration in those constraints. In our formulation, the variety of a given meal plan not only depends on the degree of repetition of the specific courses, but also on the food groups used in the meals that make up the plan. At the same time, the degree of repetition also depends on the number of days for which the plan is designed, since the time intervals between repetitions are also considered.

The nutritional quality of the meal plans are treated as the problem constraints. As we previously stated, general recommendations on intakes of nutrients, fats, carbohydrates and proteins are considered in this formulation of the MMPP, for which we followed the recommendations given in the *White Book on Child Nutrition*. Some other consensus documents and guides on school diets and allergens were also consulted, particularly, those endorsed by the Spanish Government *Ministry of Education, Culture and Sports* and the *Ministry of Health, Social Services and Equality*[3]. Since all these recommendations will be treated as constraints, the proposed formulation ensures that every meal plan obtained will fulfil the entire set of nutritional recommendations. Other very important constraints that may be considered in the problem definition involve the inability of certain people to consume some types of food. These restrictions are mainly due to food intolerances, allergies, illnesses, and other types of food incompatibilities, such as those associated with religion or lifestyle habits. However, since the aforementioned constraints are more focused on a particular person rather than on a group of people, they were not considered when defining the formulation of the MMPP that we are proposing here.

Finally, we note that in an effort to design appropriate meal plans, it was necessary to design a database of courses and ingredients with information on their respective nutritional specifications and nutritional advice, as well as on their allergens. The aforementioned documents were also considered to this end.

3.1. Encoding of individuals

An individual represents a meal plan in which there are as many meals—consisting of a starter, main course and dessert—as there are days for which the plan is going to be designed. Computationally, an individual is represented as a one-dimensional vector of integers I with length $|I| = 3 \cdot n$, where n is the number of days considered in the plan. Each element $i_{q=1,\ldots,3 \cdot n} \in I$ corresponds to the identification number (id) of a course in the course database. For each course considered, all of its related information, i.e., its name, price, amount in grams, nutritional data, food groups, allergens and food incompatibilities, is also available in the database. Figure 1 graphically represents the above explanation.

[3] These consensus documents can be accessed through `http://www.aecosan.msssi.gob.es/AECOSAN/docs/documentos/nutricion/educanaos/documento_consenso.pdf` and `https://www.aepnaa.org/`.

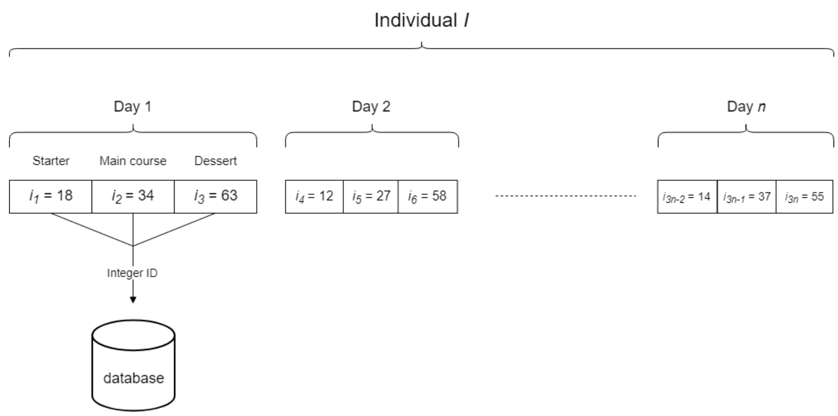

Figure 1: Example of individual encoding

Considering a given number of starters l_{st}, main courses l_{mc} and desserts l_{ds}, the size of the search space S is given by Equation 1. We note that the size of the search space exponentially increases with the number of days that the plan is designed for. For instance, by only considering $l_{st} = l_{mc} = l_{ds} = 10$ and a weekly meal plan, i.e., $n = 5$ days, the number of potential solutions $|S|$ would be equal to 10^{15}.

$$|S| = (l_{st} \cdot l_{mc} \cdot l_{ds})^n \tag{1}$$

3.2. Objective functions

The total cost of a meal plan for n days is calculated as follows:

$$C = \sum_{j=1}^{n} c_{st_j} + c_{mc_j} + c_{ds_j} \tag{2}$$

where c_{st_j}, c_{mc_j} and c_{ds_j} are the costs for the starter, main course and dessert, respectively, for day j. In individual I, st_j, mc_j, and ds_j correspond to elements $i_{3\cdot j-2}$, $i_{3\cdot j-1}$ and $i_{3\cdot j}$, respectively. The cost for a given course is calculated as the sum of the costs of its ingredients. For each ingredient, the database stores its price per kilogram, and for each course, the number of grams of a given ingredient required to prepare that course is also stored.

The novel objective function modelling the degree of repetition of courses and food groups is calculated as:

$$R = \sum_{j=1}^{n} v_{MC_j} + \frac{p_{st}}{d_{st_j}} + \frac{p_{mc}}{d_{mc_j}} + \frac{p_{ds}}{d_{ds_j}} + v_{FG_j} \tag{3}$$

where v_{MC_j} represents the compatibility, in terms of food groups, among courses st, mc and ds for day j; p_{st}, p_{mc} and p_{ds} are the penalty constants, one per

Table 2: Types of penalties defined to calculate the second objective function

Penalty	Description	Value
p_1	penalty for repeating *other* food group	0.1
p_2	penalty for repeating *meat* food group	3
p_3	penalty for repeating *cereal* food group	0.3
p_4	penalty for repeating *fruit* food group	0.1
p_5	penalty for repeating *dairy* food group	0.3
p_6	penalty for repeating *legume* food group	0.3
p_7	penalty for repeating *shellfish* food group	2
p_8	penalty for repeating *pasta* food group	1.5
p_9	penalty for repeating *fish* food group	0.5
p_{10}	penalty for repeating *vegetable* food group	0.1
p_{11}	penalty for repeating the same food group one day before	3
p_{12}	penalty for repeating the same food group two days before	2.5
p_{13}	penalty for repeating the same food group three days before	1.8
p_{14}	penalty for repeating the same food group four days before	1
p_{15}	penalty for repeating the same food group five days before	0.2
$p_{16} = p_{st}$	penalty for repeating a starter	8
$p_{17} = p_{mc}$	penalty for repeating a main course	10
$p_{18} = p_{ds}$	penalty for repeating a dessert	2

course type; d_{st_j}, d_{mc_j} and d_{ds_j} are the number of days since the corresponding course last appeared in previous days with respect to day j; and v_{FG_j} is the penalty value for repeating food groups in the last five days with respect to day j. The food groups considered for the available meals in this work are $G =$ {*other, meat, cereal, fruit, dairy, legume, shellfish, pasta, fish, vegetable*} by following the suggestions given by the *Intervention Program for the Prevention of Childhood Obesity*.

A healthy meal plan should be well-balanced, and therefore, specific courses and food groups should not be repeated frequently. Penalties are operations performed on the second objective function in order to determine the quality and variety of a meal plan. They directly affect the way a meal plan is obtained and can be used by the decision-maker in order to establish preferences. Penalties are represented by float constants. The higher the values, the greater the penalties. The different types of penalties considered are shown in Table 2.

As it can be observed, the penalties are determined by the repetition of food groups (p_1–p_{10}), the repetition of the same food group from one to five days prior to the current day (p_{11}–p_{15}), and the repetition of specific courses ($p_{16} = p_{st}$, $p_{17} = p_{mc}$, and $p_{18} = p_{ds}$). Penalty values were set by performing a preliminary study where the features of the solutions obtained were analysed in terms of repetition of courses and food groups.

In the case of penalties for repeating food groups (p_1–p_{10}), if the penalty value of a given food group is very large in comparison to the remaining food group penalty values, then a plan with a lower number of courses belonging to that food group will be provided. For instance, we have given preference to those courses consisting primarily of vegetables ($p_{10} = 0.1$) over courses composed primarily of meat ($p_2 = 3$). These penalties are used in the computation of v_{MC} and v_{FG} in Equation 3.

In the case of penalties for repeating the same food group in previous days (p_{11}–p_{15}), the more days that have passed since a food group was repeated, the lower the penalty. A time window $T = 5$ days was set because a weekly scholar menu typically involves the specification of five meals, one per day. Penalties for repeating food groups are less restrictive in comparison to penalties for repeating specific courses. A particular course should not be repeated in short periods of time. A particular food group, however, will be likely repeated since different courses could belong to the same food group. These penalties are considered in the calculation of v_{FG} in Equation 3.

Finally, in the case of the penalties for repeating a specific course, i.e., $p_{16} = p_{st}$, $p_{17} = p_{mc}$, and $p_{18} = p_{ds}$, their value is set depending on the quantity of a course type, i.e., the number of starters, main courses or desserts, available in the database. Since there are fewer desserts than starters and main courses, desserts will inevitably be repeated more often and thus their penalty value will be lower in comparison to the penalty values for repeating starters and main courses. We note that if a course of a given day j has not been repeated on any previous day, then the corresponding value for d_{st_j}, d_{mc_j} or d_{ds_j} will be infinite. As a result, the corresponding fraction in Equation 3 will be equal to zero.

In order to compute v_{FG_j}, Equation 4 is used, where $T = 5$ days is the number of previous days considered, $|G|$ is the number of food groups, $x_{g_i} \in \{0, 1\}$ indicates whether the food group g is repeated at day $j - i$ ($x_{g_i} = 1$) or not ($x_{g_i} = 0$) with respect to day j, $y_i \in \{0, 1\}$ indicates whether any food group was repeated i day(s) before the current day j ($y_i = 1$) or not ($y_i = 0$), and p_g and $p_{|G|+i}$ are the corresponding penalty values.

$$v_{FG_j} = \sum_{i=1}^{min(j-1,T)} (\sum_{g=1}^{|G|} x_{g_i} \cdot p_g) + (y_i \cdot p_{|G|+i}) \tag{4}$$

Equation 5 allows the value of v_{MC_j} to be calculated, where $|G|$ is the number of food groups, $x_g \in \{0, 1, 2, 3\}$ is the number of times a particular food group is contained in the three courses (starter, main course and dessert) of the menu for day j, and p_g is the corresponding penalty value for repeating the food group g.

$$v_{MC_j} = \sum_{g=1}^{|G|} x_g \cdot p_g \tag{5}$$

3.3. Constraints

Every feasible meal plan must comply with a recommended amount of nutrients. The nutrients considered in a plan are those given by $H = \{$*folic acid, calcium, energy, phosphorus, total fat, iron, magnesium, potassium, proteins, selenium, sodium, vitamin A, vitamin B1, vitamin B2, vitamin B6, vitamin B12, vitamin C, vitamin D, vitamin E, iodine, zinc*$\}$ by following the suggestions given by the *Intervention Program for the Prevention of Childhood Obesity.*

For each meal plan generated, all of its nutrients will be calculated. For each of the nutrients, its value has to be within the acceptable range of its recommendation; otherwise, the individual will be considered non-feasible. At this point, we should note that, since we are only considering lunch in the meal plan, recommendation ranges are the acceptable ones by only considering lunch, instead of taking into account the meals for a whole day. There is therefore a set R of pairs (r_{min}, r_{max}) and size $|H|$, where element r_{min_h} represents the minimum amount of nutrient h allowed, and r_{max_h} represents the maximum amount of nutrient h allowed, in grams, milligrams, micrograms or kilocalories, and multiplied by the number of days for which the meal plan is being designed. Formally, a feasible individual—from a nutritional point of view—should fulfil the following constraints:

$$\forall h \in H : r_{min_h} \leq r_h \leq r_{max_h} \tag{6}$$

4. Experimental evaluation

In this section we describe the experiments designed to deal with the particular formulation of the MMPP proposed in this paper. The MOEAs we considered were the well-known *Non-dominated Sorting Genetic Algorithm II* (NSGA-II) [6], the *Improved Strength Pareto Evolutionary Algorithm* (SPEA2) [44], and the adaptive version of the *Indicator-based Evolutionary Algorithm* (IBEA) [43]. The NSGA-II was selected since it has been one of the most widely applied algorithms to deal with multi-objective variants of the MPP proposed in past research. For that reason, it would be interesting to know how it performs with the novel MMPP. The NSGA-II is a Pareto-based MOEA, i.e., it applies Pareto optimality concepts to guide the search. The SPEA2, which is another Pareto-based MOEA, was also selected in order to compare its performance with respect to the NSGA-II. Finally, the IBEA was also selected, in order to include in the comparison a MOEA that applies a quality indicator, particularly, the *binary additive* ϵ*-indicator* [45], to assign a fitness value to individuals.

Other types of MOEAs have been proposed, such as the *Multi-objective Evolutionary Algorithm based on Decomposition* (MOEA/D) [41] and the *Global Weighting Achievement Scalarizing Function Genetic Algorithm* (Global WASF-GA) [34], among others. We have to say, however, that a significant number of those algorithmic schemes are aggregation-based MOEAs aimed to solve a particular MOP by decomposing it in a set of scalarizing—single-objective—subproblems, which is an approach completely different to that applied by more traditional MOEAs. At the same time, although they can be applied to solve bi-objective problems, aggregation-based MOEAs are much more suitable to deal with many-objective optimisation, i.e., when three or more objective functions are addressed, which is not the case herein. Finally, we are much more interested in analysing if the novel formulation of the MMPP we are proposing is appropriate, not only from the point of view of its multi-objective nature, but also from the nutritional standpoint of the solutions achieved by solving it, instead of focusing our efforts on providing the best possible solutions. In the end, we are proposing a novel multi-objective formulation of the MPP, and therefore, providing the best-known solutions has no sense at this moment. Bearing the above in mind, the application of more recent MOEAs to our formulation of the MMPP will be addressed as a future line of work, with the aim of potentially improving the quality of the solutions presented in the current work.

In order to complete the definition of the MOEAs selected for comparison, the crossover, mutation and repair operators must be described. The crossover method aims to uniformly combine a pair of individuals or solutions I and I' to create a new descendent. Given a crossover rate p_c, which is the probability of applying the operator, and $q = \{1, ..., |I|\}$, all the elements of both individuals will be traversed at the same time, and for each element $i_q \in I$ and $i'_q \in I'$ there will be a 50% probability that the said elements will be exchanged.

Through the mutation operator, every gene of a particular individual, i.e., a daily meal consisting of a starter, a main course and a dessert, will be randomly replaced by another daily meal with probability p_m.

The repair method is responsible for assessing each individual according to the nutritional value constraints and for determining if the individual is feasible, based on the definition of feasibility given in Section 3.3. Specifically, a daily meal for an infeasible individual will be modified by following the same procedure implemented in the mutation. However, no rate is considered by the repair operation, meaning the modification will always occur. The above steps will be repeated for each daily meal until the individual becomes feasible.

Experimental method. All the algorithms, as well as the MMPP, were implemented through the *Meta-heuristic-based Extensible Tool for Cooperative Optimisation* (METCO) proposed by León et al. [27]. The experiments were executed on a Debian GNU/Linux computer with four AMD® Opteron™ processors (model number 6348 HE) at 2.8 GHz and 64 GB RAM. Unless the opposite is established expressly, each run was repeated 30 times, since all the approaches considered are stochastic. Furthermore, with the aim of statistically supporting the conclusions extracted, the following statistical testing procedure, which was formerly used in a previous work by the authors [35], was applied to conduct comparisons between experiments. First, a *Shapiro-Wilk test* was performed to check whether the values of the results followed a normal (Gaussian) distribution. If so, the *Levene test* checked for the homogeneity of the variances. If the samples had equal variance, an ANOVA *test* was done; if not, a *Welch test* was performed. For non-Gaussian distributions, the non-parametric *Kruskal-Wallis* test was used. For all tests, a significance level $\alpha = 0.05$ was considered. The *hypervolume* [42] was selected as the metric to perform the comparisons among the different approaches. To calculate it, the definition of a reference point is required. We set the reference point to the *nadir point*, i.e., the worst—maximum—values that each of the two objective functions attained considering all the runs performed in each of the experiments. Additionally, the hypervolume was normalised in the range $[0, 1]$. The higher its value, the better the performance of the algorithm in question.

Instances. In order to define an instance of the problem, the number of days for which the meal plan will be designed has to be specified. Furthermore, a

```
<name>, <price>, <amount>, <allergens>, <incompatibilities>,
<amount of nutrients>, <food groups>
```

Figure 2: Comma-sepparated values (CSV) format of each line of the files that contain information on the starters, main courses and desserts. Each line specifies the information of a particular starter, main course or dessert. At the same time, each field consists of different subfields whose values have to be provided.

starter, a main course and a dessert have to be selected for each day. As a result, all potential starters, main courses and desserts that can be chosen have to be somehow described as well. In the particular case of our implementation, a total number of 67 courses consisting of $l_{st} = 19$ starters, $l_{mc} = 34$ main courses and $l_{ds} = 14$ desserts are provided in three different files containing information on each type of course[4]. The format designed for those files, and for each of their lines in particular, is depicted in Figure 2, where each field must be filled in with the following information[5]:

- `<name>`: name of the course.
- `<price>`: price of the course.
- `<amount>`: amount in grams.
- `<allergens>`: cereal, nuts, legumes, shellfish, fish, egg protein, dairy protein. Each of the above fields takes the value 0 if the course does not include the allergen, and the value 1 if it does.
- `<incompatibilities>`: caeliac, diabetes, semi-vegetarianism, vegetarianism, veganism. Each of the aforementioned fields takes the value 0 if the course is not incompatible, and the value 1 if it is.
- `<amount of nutrients>`: folic acid (μg), calcium (mg), energy (kcal), phosphorus (mg), total fat (g), iron (mg), magnesium (mg), potassium (mg), proteins (g), selenium (μg), sodium (mg), vit. A (μg), vit. B1 (mg), vit. B2 (mg), vit. B6 (mg), vit. B12 (μg), vit. C (mg), vit. D (μg), vit. E (mg), iodine (μg), zinc (mg). Each of the above fields has to be specified as a decimal number indicating the amount of the nutrient in question by considering its corresponding unit of measure.
- `<food groups>`: others (0), meat (1), cereal (2), fruit (3), dairy (4), legume (5), shellfish (6), pasta (7), fish (8), vegetable (9). Only those food groups that the course in question contains have to be specified through their corresponding integer numbers.

[4]Considering, for instance, a meal plan for 5 days, the size of the search space would be $|S| = (19 \cdot 34 \cdot 14)^5 = 6.051 \cdot 10^{19}$ potential solutions.

[5]Further information about the source code of the MMPP, the specific instances defined for the current work, and our results can be found at `https://github.com/Tomas-Morph/On-the-Design-of-Healthy-Menus-through-Evolutionary-Computation`.

Table 3: Parameter values tested for the NSGA-II, the SPEA2 and the IBEA

Parameter	Value	Parameter	Value
Stopping criterion	$4 \cdot 10^8$ evals.	Mutation prob. (p_m)	0.05, 0.1, 0.15
Population size (n)	25, 50, 100, 250 inds.	Crossover prob. (p_c)	0.8, 0.9, 1

4.1. First experiment: parameter setting

The main goal of this first experiment is to conduct a study involving the values that parameters belonging to the MOEAs selected for comparison should take, since we are dealing with a novel MMPP. To do so, we tested different values for the parameters of each algorithm, as Table 3 shows. Particularly, a total number of 36 different configurations of each MOEA were applied in order to yield a 20-day meal plan. Those configurations were obtained by combining four values for the population size (n), three values for the mutation rate (p_m) and three values for the crossover rate (p_c). In the particular case of parameter p_m, test values were selected based on the operation of the mutation operator. The reader should recall that a mutation operator specifically designed for the MMPP is considered herein. Specifically, for each day, the operator modifies the three courses with probability p_m. Since this experiment requires a 20-day meal plan, $p_m = 0.05$ is the minimum probability that allows, at least, a day courses to be changed. Bearing the above in mind, the values 0.05, 0.1 and 0.15 were selected, and therefore, the courses of one, two or three days were altered by the mutation operator, respectively. So as to avoid the overly disruptive behaviour of the operator, larger values for p_m were not considered. In the case of the crossover rate, high values are usually set. That is the reason why the test values for parameter p_c were 0.8, 0.9 and 1. In order to complement the information given in Table 3, we should note that all configurations based on the IBEA applied a scaling factor equal to 0.002, while for each configuration of the SPEA2, the archive size was set to the same value considered for the population size. Finally, note that a sufficiently long stopping criterion was set, consisting of $4 \cdot 10^8$ function evaluations. The idea was to study whether the MOEAs selected converged prematurely or not, and thus determine a proper value for the stopping criterion to be used in subsequent experiments. Furthermore, for this first analysis, and due to the large number of configurations tested, the executions were only repeated 10 times. Bearing the above in mind, a total of 1,080 runs were carried out, which involved more than 14,500 computational hours.

Table 4 shows some statistical data related to the hypervolume values attained by each configuration of the NSGA-II at the end of the executions. The last three columns allow a ranking to be established among the different configurations tested. In particular, the number of times a given configuration was able to statistically outperform other configurations (W), and the number of times a particular configuration was statistically outperformed by other configurations (L) are shown. Configuration A statistically outperforms configuration B if the p-value, obtained after performing a pairwise comparison of both approaches by following the statistical testing procedure described at the beginning of this

Table 4: Performance of the NSGA-II in terms of the hypervolume attained at the end of the runs. Configurations are sorted in descending order by considering their ranking as the first criterion, and the mean and the median of the hypervolume as the second and third ones, respectively

Configuration	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.	W	L	R
NSGA-II_ps_250_pm_0.1_pc_0.8	0.8634	0.8666	0.8696	0.8691	0.8713	0.8743	28	0	28
NSGA-II_ps_250_pm_0.05_pc_0.8	0.8598	0.8654	0.8675	0.8679	0.8703	0.8767	24	0	24
NSGA-II_ps_250_pm_0.15_pc_0.8	0.8541	0.8619	0.8675	0.8678	0.8728	0.8818	18	0	18
NSGA-II_ps_50_pm_0.1_pc_0.8	0.8553	0.8589	0.8619	0.8660	0.8694	0.8871	11	0	11
NSGA-II_ps_50_pm_0.15_pc_0.8	0.8474	0.8596	0.8668	0.8643	0.8686	0.8740	10	0	10
NSGA-II_ps_100_pm_0.1_pc_0.9	0.8463	0.8581	0.8632	0.8640	0.8720	0.8770	10	0	10
NSGA-II_ps_100_pm_0.05_pc_0.9	0.8525	0.8571	0.8647	0.8639	0.8709	0.8728	10	0	10
NSGA-II_ps_250_pm_0.1_pc_0.9	0.8492	0.8597	0.8640	0.8632	0.8678	0.8782	10	0	10
NSGA-II_ps_250_pm_0.05_pc_0.9	0.8567	0.8587	0.8629	0.8629	0.8653	0.8709	11	2	9
NSGA-II_ps_50_pm_0.05_pc_1.0	0.8501	0.8568	0.8620	0.8627	0.8701	0.8725	9	1	8
NSGA-II_ps_100_pm_0.05_pc_0.8	0.8528	0.8588	0.8613	0.8624	0.8687	0.8706	10	2	8
NSGA-II_ps_100_pm_0.1_pc_0.8	0.8480	0.8565	0.8616	0.8619	0.8659	0.8785	9	1	8
NSGA-II_ps_50_pm_0.05_pc_0.8	0.8514	0.8588	0.8611	0.8615	0.8657	0.8694	9	2	7
NSGA-II_ps_50_pm_0.1_pc_0.9	0.8470	0.8521	0.8638	0.8612	0.8686	0.8762	5	1	4
NSGA-II_ps_100_pm_0.15_pc_0.8	0.8415	0.8600	0.8609	0.8599	0.8650	0.8696	6	2	4
NSGA-II_ps_50_pm_0.05_pc_0.9	0.8472	0.8553	0.8594	0.8599	0.8637	0.8745	6	2	4
NSGA-II_ps_25_pm_0.05_pc_0.9	0.8483	0.8580	0.8597	0.8598	0.8632	0.8676	7	3	4
NSGA-II_ps_250_pm_0.05_pc_1.0	0.8541	0.8550	0.8572	0.8595	0.8625	0.8687	7	3	4
NSGA-II_ps_25_pm_0.15_pc_0.8	0.8493	0.8586	0.8598	0.8587	0.8601	0.8635	7	3	4
NSGA-II_ps_250_pm_0.15_pc_1.0	0.8480	0.8559	0.8569	0.8587	0.8616	0.8722	6	3	3
NSGA-II_ps_100_pm_0.15_pc_0.9	0.8333	0.8547	0.8598	0.8576	0.8647	0.8671	4	2	2
NSGA-II_ps_50_pm_0.15_pc_0.9	0.8457	0.8511	0.8634	0.8592	0.8653	0.8684	4	3	1
NSGA-II_ps_50_pm_0.1_pc_1.0	0.8442	0.8530	0.8614	0.8586	0.8643	0.8677	4	3	1
NSGA-II_ps_250_pm_0.15_pc_0.9	0.8409	0.8481	0.8627	0.8594	0.8687	0.8757	1	1	0
NSGA-II_ps_25_pm_0.05_pc_0.8	0.8449	0.8558	0.8578	0.8574	0.8624	0.8647	4	5	-1
NSGA-II_ps_25_pm_0.1_pc_0.8	0.8371	0.8509	0.8564	0.8571	0.8671	0.8749	0	3	-3
NSGA-II_ps_250_pm_0.1_pc_1.0	0.8303	0.8425	0.8501	0.8518	0.8576	0.8808	0	10	-10
NSGA-II_ps_100_pm_0.1_pc_1.0	0.8427	0.8487	0.8505	0.8534	0.8569	0.8686	0	13	-13
NSGA-II_ps_25_pm_0.05_pc_1.0	0.8385	0.8461	0.8558	0.8528	0.8577	0.8647	0	13	-13
NSGA-II_ps_25_pm_0.1_pc_0.9	0.8422	0.8477	0.8560	0.8535	0.8582	0.8614	0	16	-16
NSGA-II_ps_25_pm_0.15_pc_0.9	0.8355	0.8456	0.8523	0.8512	0.8556	0.8634	0	20	-20
NSGA-II_ps_100_pm_0.05_pc_1.0	0.8297	0.8502	0.8543	0.8515	0.8557	0.8585	0	21	-21
NSGA-II_ps_25_pm_0.1_pc_1.0	0.8294	0.8452	0.8516	0.8496	0.8570	0.8634	0	22	-22
NSGA-II_ps_50_pm_0.15_pc_1.0	0.8371	0.8486	0.8524	0.8512	0.8553	0.8586	0	24	-24
NSGA-II_ps_25_pm_0.15_pc_1.0	0.8411	0.8465	0.8489	0.8501	0.8540	0.8598	0	24	-24
NSGA-II_ps_100_pm_0.15_pc_1.0	0.8339	0.8392	0.8506	0.8476	0.8554	0.8582	0	25	-25

section, is lower than the significance level α , and if at the same time, A provides a higher mean and median of the hypervolume at the end of the runs. The ranking (R) was calculated as the difference between the number of times a configuration outperformed others, and the number of times a configuration was outperformed by others ($R = W - L$). Note that the configurations are shown in descending order based on their position in the ranking as the first criterion, and the mean and the median of the hypervolume as the second and third criteria. With the aim of differentiating the configurations, the nomenclature $Algorithm\_ps\_n\_pm\_p_m\_pc\_p_c$ was used. Finally, Tables 5 and 6 show the same information but for the IBEA and SPEA2 approaches, respectively[6].

With respect to the NSGA-II, we observed that, as a general fact, larger

[6]The same nomenclature was used to differentiate the configurations, but adding some slight modifications to include information on the scaling factor (*sFactor*), in the case of the IBEA, and the archive size (*as*), in the case of the SPEA2.

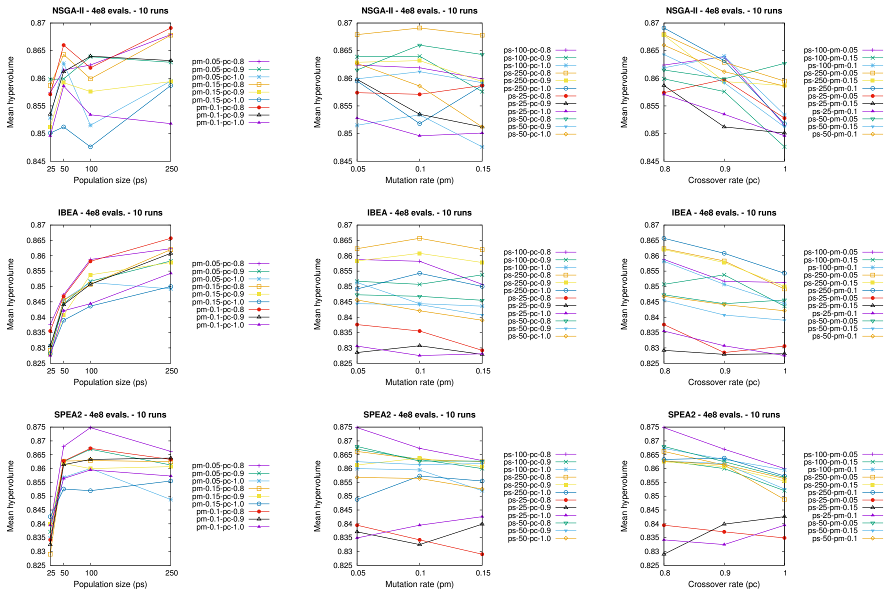

Figure 3: Mean hypervolume values achieved by the different configurations of the NSGA-II (first row), the IBEA (second row) and the SPEA2 (third row) at the end of the runs with respect to the population size (first column), the mutation rate (second column) and the crossover rate (third column)

population sizes yielded better performance in terms of hypervolume values, in comparison to smaller populations. For instance, the values for the first-ranked configuration were $p_m = 0.1$, $p_c = 0.8$, with $n = 250$ individuals, which was the largest population tested. If the same values are set for parameters p_m and p_c, the worst-ranked configuration contained 25 individuals, which was the smallest population considered. In order to better analyse the above fact, Figure 3 shows the mean hypervolume values attained by the the different configurations of the NSGA-II (first row), the IBEA (second row) and the SPEA2 (third row) at the end of the executions with respect to the population size (first column), the mutation rate (second column) and the crossover rate (third column). In the case of the NSGA-II, note that for almost every possible combination of values for parameters p_m and p_c (first row, first column), the best results were achieved by applying the largest population considered ($n = 250$), while the worst results were provided by using the smallest population tested ($n = 25$).

Regarding the crossover rate, we should note that, generally speaking, lower values yielded better hypervolume values, and therefore better performance, while higher values of the crossover rate decreased the performance of the NSGA-II significantly. As proof of the above, the last five configurations in the ranking were applied with $p_c = 1$, while the first five configurations in the ranking were executed with $p_c = 0.8$. Furthermore, as Figure 3 shows (first row, third column), in the general case, the lowest hypervolume values were achieved with

Table 5: Performance of the IBEA in terms of the hypervolume attained at the end of the runs. Configurations are sorted in descending order by considering their ranking as the first criterion, and the mean and the median of the hypervolume as the second and third ones, respectively

Configuration	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.	W	L	R
IBEA_ps_250_pm_0.1_pc_0.8_sFactor_0.002	0.8594	0.8640	0.8657	0.8657	0.8689	0.8703	32	0	32
IBEA_ps_250_pm_0.05_pc_0.8_sFactor_0.002	0.8489	0.8580	0.8645	0.8623	0.8690	0.8707	28	0	28
IBEA_ps_250_pm_0.1_pc_0.9_sFactor_0.002	0.8519	0.8566	0.8593	0.8608	0.8636	0.8710	28	1	27
IBEA_ps_250_pm_0.15_pc_0.8_sFactor_0.002	0.8346	0.8589	0.8651	0.8620	0.8700	0.8739	25	0	25
IBEA_ps_100_pm_0.05_pc_0.8_sFactor_0.002	0.8488	0.8538	0.8614	0.8588	0.8630	0.8672	25	1	24
IBEA_ps_100_pm_0.1_pc_0.8_sFactor_0.002	0.8461	0.8547	0.8586	0.8582	0.8631	0.8702	24	1	23
IBEA_ps_250_pm_0.15_pc_0.9_sFactor_0.002	0.8444	0.8537	0.8576	0.8578	0.8629	0.8704	23	1	22
IBEA_ps_250_pm_0.05_pc_0.9_sFactor_0.002	0.8344	0.8486	0.8620	0.8583	0.8666	0.8787	20	0	20
IBEA_ps_250_pm_0.1_pc_1.0_sFactor_0.002	0.8428	0.8532	0.8549	0.8543	0.8554	0.8614	19	3	16
IBEA_ps_100_pm_0.15_pc_0.9_sFactor_0.002	0.8401	0.8514	0.8557	0.8538	0.8570	0.8610	19	3	16
IBEA_ps_100_pm_0.05_pc_1.0_sFactor_0.002	0.8257	0.8516	0.8528	0.8513	0.8565	0.8601	19	5	14
IBEA_ps_100_pm_0.05_pc_0.9_sFactor_0.002	0.8303	0.8514	0.8529	0.8517	0.8567	0.8671	12	3	9
IBEA_ps_100_pm_0.1_pc_0.9_sFactor_0.002	0.8391	0.8472	0.8511	0.8507	0.8563	0.8600	14	7	7
IBEA_ps_100_pm_0.15_pc_0.8_sFactor_0.002	0.8385	0.8456	0.8509	0.8507	0.8575	0.8593	13	7	6
IBEA_ps_250_pm_0.15_pc_1.0_sFactor_0.002	0.8355	0.8445	0.8492	0.8500	0.8561	0.8666	10	6	4
IBEA_ps_250_pm_0.05_pc_1.0_sFactor_0.002	0.8300	0.8453	0.8509	0.8492	0.8556	0.8609	10	7	3
IBEA_ps_50_pm_0.05_pc_0.8_sFactor_0.002	0.8336	0.8396	0.8478	0.8473	0.8549	0.8621	9	8	1
IBEA_ps_50_pm_0.1_pc_0.8_sFactor_0.002	0.8293	0.8436	0.8482	0.8468	0.8500	0.8618	9	11	-2
IBEA_ps_50_pm_0.05_pc_1.0_sFactor_0.002	0.8299	0.8404	0.8481	0.8456	0.8510	0.8596	9	11	-2
IBEA_ps_50_pm_0.15_pc_0.8_sFactor_0.002	0.8314	0.8421	0.8455	0.8455	0.8505	0.8555	9	11	-2
IBEA_ps_100_pm_0.1_pc_1.0_sFactor_0.002	0.8338	0.8403	0.8466	0.8444	0.8475	0.8531	9	12	-3
IBEA_ps_50_pm_0.05_pc_0.9_sFactor_0.002	0.8317	0.8374	0.8463	0.8444	0.8507	0.8540	8	11	-3
IBEA_ps_50_pm_0.1_pc_0.9_sFactor_0.002	0.8287	0.8417	0.8431	0.8441	0.8484	0.8599	8	11	-3
IBEA_ps_100_pm_0.15_pc_1.0_sFactor_0.002	0.8370	0.8396	0.8424	0.8436	0.8466	0.8543	9	14	-5
IBEA_ps_50_pm_0.1_pc_1.0_sFactor_0.002	0.8243	0.8335	0.8419	0.8421	0.8520	0.8541	7	13	-6
IBEA_ps_50_pm_0.15_pc_0.9_sFactor_0.002	0.8234	0.8302	0.8409	0.8407	0.8487	0.8583	7	14	-7
IBEA_ps_50_pm_0.15_pc_1.0_sFactor_0.002	0.8237	0.8330	0.8408	0.8390	0.8461	0.8505	7	16	-9
IBEA_ps_25_pm_0.05_pc_0.8_sFactor_0.002	0.8277	0.8324	0.8376	0.8376	0.8430	0.8479	7	22	-15
IBEA_ps_25_pm_0.1_pc_0.8_sFactor_0.002	0.8304	0.8324	0.8345	0.8355	0.8376	0.8460	4	24	-20
IBEA_ps_25_pm_0.1_pc_0.9_sFactor_0.002	0.8198	0.8271	0.8302	0.8307	0.8359	0.8398	0	28	-28
IBEA_ps_25_pm_0.05_pc_1.0_sFactor_0.002	0.8188	0.8250	0.8319	0.8306	0.8331	0.8456	0	28	-28
IBEA_ps_25_pm_0.15_pc_0.8_sFactor_0.002	0.8137	0.8219	0.8295	0.8292	0.8371	0.8452	0	28	-28
IBEA_ps_25_pm_0.05_pc_0.9_sFactor_0.002	0.8151	0.8242	0.8286	0.8285	0.8304	0.8424	0	29	-29
IBEA_ps_25_pm_0.15_pc_1.0_sFactor_0.002	0.8191	0.8225	0.8280	0.8281	0.8339	0.8355	0	29	-29
IBEA_ps_25_pm_0.15_pc_0.9_sFactor_0.002	0.8153	0.8265	0.8288	0.8279	0.8301	0.8376	0	29	-29
IBEA_ps_25_pm_0.1_pc_1.0_sFactor_0.002	0.8122	0.8226	0.8282	0.8275	0.8347	0.8376	0	29	-29

$p_c = 1$, while the highest hypervolume values were attained with $p_c = 0.8$.

In regard to the mutation rate, however, a clear conclusion cannot be extracted. It is true that, for instance, the first three configurations of the NSGA-II in the ranking were applied with the same population size and crossover rate, while each one of them was run with a different value for the mutation rate. Bearing the above in mind, we may claim that the mutation rate parameter does not seem to have a significant effect on the performance of the NSGA-II, at least considering the values we selected for parameter p_m in this study. However, as Figure 3 shows (first row, second column), in some cases, lower values of the mutation rate provided better results, while higher values of the mutation rate yielded better performance over other cases. As a result, more experiments should be conducted to shed more light on the behaviour of the NSGA-II with respect to the mutation rate, which seems to be its most sensitive parameter when dealing with this particular problem.

Finally, it is important to remark that the first configuration of the NSGA-II

Table 6: Performance of the SPEA2 in terms of the hypervolume attained at the end of the runs. Configurations are sorted in descending order by considering their ranking as the first criterion, and the mean and the median of the hypervolume as the second and third ones, respectively

Configuration	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.	W	L	R
SPEA2_ps_100_as_100_pm_0.05_pc_0.8	0.8642	0.8699	0.8752	0.8748	0.8792	0.8880	34	0	34
SPEA2_ps_50_as_50_pm_0.05_pc_0.8	0.8513	0.8636	0.8710	0.8680	0.8742	0.8763	18	0	18
SPEA2_ps_100_as_100_pm_0.1_pc_0.8	0.8527	0.8638	0.8663	0.8673	0.8744	0.8777	18	1	17
SPEA2_ps_100_as_100_pm_0.05_pc_0.9	0.8583	0.8620	0.8661	0.8670	0.8701	0.8831	18	1	17
SPEA2_ps_250_as_250_pm_0.05_pc_0.8	0.8553	0.8591	0.8653	0.8662	0.8723	0.8790	17	1	16
SPEA2_ps_250_as_250_pm_0.1_pc_0.9	0.8537	0.8595	0.8641	0.8638	0.8692	0.8721	15	1	14
SPEA2_ps_100_as_100_pm_0.1_pc_0.9	0.8549	0.8596	0.8611	0.8633	0.8677	0.8761	14	1	13
SPEA2_ps_100_as_100_pm_0.15_pc_0.8	0.8496	0.8605	0.8646	0.8629	0.8652	0.8719	14	1	13
SPEA2_ps_50_as_50_pm_0.15_pc_0.8	0.8494	0.8618	0.8639	0.8627	0.8672	0.8687	14	1	13
SPEA2_ps_50_as_50_pm_0.1_pc_0.8	0.8548	0.8583	0.8620	0.8627	0.8644	0.8735	14	1	13
SPEA2_ps_250_as_250_pm_0.15_pc_0.8	0.8559	0.8585	0.8609	0.8625	0.8657	0.8719	14	1	13
SPEA2_ps_50_as_50_pm_0.15_pc_0.9	0.8547	0.8569	0.8620	0.8618	0.8660	0.8698	13	1	12
SPEA2_ps_250_as_250_pm_0.1_pc_0.8	0.8478	0.8575	0.8636	0.8633	0.8694	0.8779	12	1	11
SPEA2_ps_50_as_50_pm_0.05_pc_0.9	0.8463	0.8564	0.8632	0.8625	0.8699	0.8820	10	1	9
SPEA2_ps_50_as_50_pm_0.1_pc_0.9	0.8466	0.8558	0.8609	0.8614	0.8684	0.8766	10	1	9
SPEA2_ps_250_as_250_pm_0.05_pc_0.9	0.8505	0.8580	0.8620	0.8613	0.8636	0.8730	10	1	9
SPEA2_ps_250_as_250_pm_0.15_pc_0.9	0.8491	0.8553	0.8619	0.8607	0.8631	0.8750	10	1	9
SPEA2_ps_100_as_100_pm_0.05_pc_1.0	0.8464	0.8494	0.8608	0.8600	0.8664	0.8766	10	1	9
SPEA2_ps_100_as_100_pm_0.15_pc_0.9	0.8483	0.8576	0.8613	0.8600	0.8618	0.8712	10	4	6
SPEA2_ps_100_as_100_pm_0.1_pc_1.0	0.8556	0.8563	0.8580	0.8595	0.8623	0.8676	10	5	5
SPEA2_ps_250_as_250_pm_0.1_pc_1.0	0.8460	0.8526	0.8550	0.8573	0.8596	0.8753	9	5	4
SPEA2_ps_50_as_50_pm_0.05_pc_1.0	0.8432	0.8521	0.8576	0.8568	0.8598	0.8679	9	6	3
SPEA2_ps_50_as_50_pm_0.1_pc_1.0	0.8490	0.8521	0.8565	0.8564	0.8597	0.8653	9	12	-3
SPEA2_ps_250_as_250_pm_0.15_pc_1.0	0.8422	0.8513	0.8557	0.8555	0.8590	0.8691	9	12	-3
SPEA2_ps_50_as_50_pm_0.15_pc_1.0	0.8333	0.8471	0.8537	0.8526	0.8606	0.8737	8	12	-4
SPEA2_ps_100_as_100_pm_0.15_pc_1.0	0.8336	0.8433	0.8529	0.8520	0.8627	0.8703	8	13	-5
SPEA2_ps_250_as_250_pm_0.05_pc_1.0	0.8323	0.8391	0.8492	0.8488	0.8567	0.8659	5	20	-15
SPEA2_ps_25_as_25_pm_0.15_pc_1.0	0.8228	0.8373	0.8402	0.8426	0.8502	0.8597	2	24	-22
SPEA2_ps_25_as_25_pm_0.15_pc_0.9	0.8224	0.8310	0.8388	0.8399	0.8511	0.8559	0	26	-26
SPEA2_ps_25_as_25_pm_0.1_pc_1.0	0.8172	0.8287	0.8416	0.8395	0.8501	0.8599	0	26	-26
SPEA2_ps_25_as_25_pm_0.05_pc_0.8	0.8264	0.8322	0.8408	0.8395	0.8464	0.8503	0	26	-26
SPEA2_ps_25_as_25_pm_0.05_pc_0.9	0.8187	0.8275	0.8345	0.8371	0.8471	0.8552	0	27	-27
SPEA2_ps_25_as_25_pm_0.05_pc_1.0	0.8196	0.8299	0.8342	0.8349	0.8406	0.8508	0	27	-27
SPEA2_ps_25_as_25_pm_0.1_pc_0.8	0.8165	0.8283	0.8345	0.8342	0.8377	0.8561	0	27	-27
SPEA2_ps_25_as_25_pm_0.1_pc_0.9	0.8114	0.8330	0.8338	0.8325	0.8395	0.8428	0	28	-28
SPEA2_ps_25_as_25_pm_0.15_pc_0.8	0.7998	0.8191	0.8338	0.8290	0.8409	0.8499	0	28	-28

in the ranking, which was run with parameters $n = 250$, $p_m = 0.1$ and $p_c = 0.8$, was able to statistically outperform 28 out of 35 configurations (80%), while it was not statistically outperformed by any other configuration. Furthermore, that configuration achieved the highest mean and median of the hypervolume at the end of the executions, taking into account all the parameterisations of the NSGA-II tested.

Similar conclusions to those given for the NSGA-II can be extracted for the IBEA (Table 5 and Figure 3). In the general case, the application of larger population sizes and lower crossover rates yielded better performance, while the mutation rate seems to be its most sensitive parameter, as in the case of the NSGA-II. In fact, the configuration of the IBEA that yielded the best performance, i.e. the first configuration in the ranking, was executed with parameters $n = 250$, $p_m = 0.1$ and $p_c = 0.8$, which was also the best-performing parameterisation of the NSGA-II. The best-performing configuration of the IBEA was able to statistically outperform 32 out of 35 configurations (91.4%), while it was not

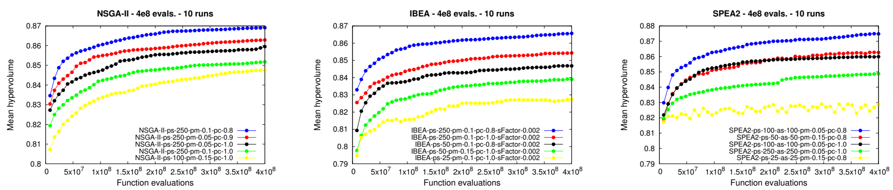

Figure 4: Evolution of the mean hypervolume values achieved by the NSGA-II, IBEA and SPEA2

statistically outperformed by any other configuration. Moreover, we should note that the best-performing configuration of the IBEA provided the highest mean and median of the hypervolume at the end of the runs, considering all of the parameterisations tested, as was the case with the NSGA-II.

The behaviour of the SPEA2, however, was slightly different in terms of its parameterisation (Table 6 and Figure 3), and particularly of the population size, with respect to the other two approaches. The best-performing configuration of the SPEA2 had $n = 100$ individuals, in contrast to the best-performing configurations of the NSGA-II and the IBEA, which were run with 250 individuals. The above may be explained by the fact that the SPEA2 makes use of an archive, whose size is usually set equal to the value selected for the population size, and which allows the non-dominated solutions considered so far to be stored during the executions. The use of that archive may avoid the need to employ larger population sizes. In fact, the first-ranked configuration of the SPEA2 was able to attain the highest mean and median of the hypervolume in comparison to the best-performing configurations of the NSGA-II and the IBEA, with the former having a smaller population (and archive) than the population used by the latter. Furthermore, statistically significant differences appeared when the best-performing configuration of the SPEA2 was compared to the first-ranked configurations of the NSGA-II and the IBEA.

With regard to the remaining parameters (p_m and p_c), similar conclusions to those drawn previously for the NSGA-II and IBEA can be also extracted for the SPEA2. Finally, we note that the first configuration of the SPEA2 in the ranking statistically outperformed 34 out of 35 configurations (97.1%), while it was not outperformed by any other configuration. As was the case with the NSGA-II and the IBEA, that first-ranked configuration provided the highest mean and median of the hypervolume at the end of the runs, considering all the configurations of the SPEA2.

This analysis was carried out by considering the results obtained at the end of the executions. However, it would be interesting to study the behaviour of the different approaches during the runs. As we said at the beginning of this section, a long stopping criterion was set with the aim of analysing if any of the MOEAs selected for comparison converged prematurely. Figure 4 shows the evolution of the mean hypervolume achieved by some configurations of the NSGA-II, the IBEA and the SPEA2. Specifically, for each of the approaches tested, the first and last configurations, i.e. the best and worst-performing parameterisations, respectively, given their corresponding rankings, were selected for comparison purposes. The configurations ranked 9, 18 and 27 were also included in the comparison. As we can see, for the three approaches, significant differences arose in terms of the performance yielded during the runs by the parameterisations applied, something that we had already stated previously. Also, despite having set a long stopping criterion consisting of $4 \cdot 10^8$ function evaluations, the mean hypervolume was improved even at the end of the runs, allowing us to conclude that none of the MOEAs converged prematurely. Bearing the above in mind, runs may be executed for even longer in order to study the behaviour of the approaches in the long term. We decided to apply the same stopping criterion in subsequent experiments, however, because of computational time restrictions. The reader should keep in mind that $4 \cdot 10^8$ function evaluations take almost 14 hours on average, given the characteristics of the machine where the experiments were conducted.

4.2. Second experiment: designing meal plans for a variable number of days

The main goal of the second experiment is to analyse the performance of the different MOEAs when they are applied to produce meal plans for a variable number of days. Taking into consideration the conclusions extracted from the first experiment, we decided to apply the NSGA-II, IBEA and SPEA2 with the same parameter values used by their corresponding first-ranked configurations. Hence, the population size was set to $n = 250$ individuals in the case of the NSGA-II and the IBEA, and $n = 100$ individuals for the SPEA2. The crossover rate was set to $p_c = 0.8$ for all of the approaches. In addition, the scaling factor of the IBEA was set to 0.002, and the archive size of the SPEA2 was set to 100 individuals. Finally, since the first experiment provided no clear conclusions on the mutation rate p_m, three different values were tested in the second experiment for this parameter. Since for this experiment we decided to design meal plans for 5, 10, 20 and 40 days, p_m was set such that the meals for one, two or three days are changed by the mutation operator. For instance, values 0.2, 0.4 and 0.6 were selected when dealing with meal plans for five days, and therefore, the meals for one, two and three days were altered by the mutation operator, respectively[7]. When designing meal plans for 40 days, however, the values 0.025, 0.05 and 0.075 were tested.

Bearing the above in mind, three different configurations of each MOEA were applied to design meal plans for 5, 10, 20 and 40 days. The stopping criterion was also set to $4 \cdot 10^8$ function evaluations, as in the first experiment, but this time every run was repeated 30 times rather than 10 times, in an effort to provide much more statistically significant conclusions. This gives a total number of 1,080 runs, which involved more than 15,500 computational hours.

[7]The reader should recall that, in this particular case, the chromosome would consist of five genes, i.e. the meal plan length, with each gene encoding a starter, main course and dessert.

Table 7: Performance of the different MOEAs in terms of the hypervolume attained at the end of 30 runs. For each menu planning length, configurations are sorted in descending order by considering their ranking as the first criterion, and the mean and the median of the hypervolume as the second and third ones, respectively

Menu plannings for 5 days									
Configuration	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.	W	L	R
NSGA-II_ps_250_pm_0.2_pc_0.8	0.9603	0.9706	0.9752	0.9734	0.9773	0.9801	7	0	7
NSGA-II_ps_250_pm_0.4_pc_0.8	0.9482	0.9696	0.9731	0.9719	0.9750	0.9792	6	0	6
NSGA-II_ps_250_pm_0.6_pc_0.8	0.9561	0.9688	0.9714	0.9708	0.9741	0.9783	6	1	5
IBEA_ps_250_pm_0.4_pc_0.8_sFactor_0.002	0.9401	0.9604	0.9639	0.9625	0.9670	0.9697	3	3	0
IBEA_ps_250_pm_0.2_pc_0.8_sFactor_0.002	0.9474	0.9560	0.9619	0.9610	0.9666	0.9701	0	3	-3
SPEA2_ps_100_as_100_pm_0.2_pc_0.8	0.9435	0.9562	0.9617	0.9602	0.9652	0.9719	0	3	-3
SPEA2_ps_100_as_100_pm_0.6_pc_0.8	0.9396	0.9583	0.9612	0.9595	0.9637	0.9664	0	4	-4
IBEA_ps_250_pm_0.6_pc_0.8_sFactor_0.002	0.9412	0.9568	0.9580	0.9592	0.9656	0.9700	0	4	-4
SPEA2_ps_100_as_100_pm_0.4_pc_0.8	0.9407	0.9554	0.9588	0.9580	0.9638	0.9665	0	4	-4
Menu plannings for 10 days									
Configuration	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.	W	L	R
NSGA-II_ps_250_pm_0.1_pc_0.8	0.9378	0.9508	0.9549	0.9544	0.9597	0.9674	7	0	7
NSGA-II_ps_250_pm_0.2_pc_0.8	0.9278	0.9481	0.9523	0.9523	0.9594	0.9642	5	0	5
IBEA_ps_250_pm_0.3_pc_0.8_sFactor_0.002	0.9337	0.9453	0.9491	0.9487	0.9537	0.9611	4	1	3
NSGA-II_ps_250_pm_0.3_pc_0.8	0.9371	0.9420	0.9482	0.9487	0.9532	0.9630	4	1	3
IBEA_ps_250_pm_0.2_pc_0.8_sFactor_0.002	0.9367	0.9423	0.9497	0.9479	0.9524	0.9586	4	2	2
SPEA2_ps_100_as_100_pm_0.1_pc_0.8	0.9279	0.9397	0.9440	0.9442	0.9482	0.9593	1	5	-4
IBEA_ps_250_pm_0.1_pc_0.8_sFactor_0.002	0.9272	0.9379	0.9448	0.9441	0.9496	0.9623	1	5	-4
SPEA2_ps_100_as_100_pm_0.2_pc_0.8	0.9299	0.9390	0.9432	0.9440	0.9491	0.9574	1	5	-4
SPEA2_ps_100_as_100_pm_0.3_pc_0.8	0.9093	0.9308	0.9396	0.9385	0.9443	0.9591	0	8	-8
Menu plannings for 20 days									
Configuration	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.	W	L	R
SPEA2_ps_100_as_100_pm_0.1_pc_0.8	0.9166	0.9316	0.9358	0.9359	0.9428	0.9498	5	0	5
SPEA2_ps_100_as_100_pm_0.05_pc_0.8	0.9130	0.9278	0.9328	0.9335	0.9390	0.9530	4	0	4
NSGA-II_ps_250_pm_0.05_pc_0.8	0.9043	0.9263	0.9346	0.9332	0.9418	0.9475	4	0	4
IBEA_ps_250_pm_0.05_pc_0.8_sFactor_0.002	0.9144	0.9275	0.9320	0.9318	0.9382	0.9430	3	0	3
NSGA-II_ps_250_pm_0.1_pc_0.8	0.9169	0.9255	0.9310	0.9303	0.9357	0.9393	1	1	0
IBEA_ps_250_pm_0.1_pc_0.8_sFactor_0.002	0.8966	0.9239	0.9306	0.9282	0.9340	0.9412	0	3	-3
NSGA-II_ps_250_pm_0.15_pc_0.8	0.9084	0.9225	0.9265	0.9273	0.9330	0.9417	0	4	-4
SPEA2_ps_100_as_100_pm_0.15_pc_0.8	0.9137	0.9223	0.9285	0.9272	0.9324	0.9419	0	4	-4
IBEA_ps_250_pm_0.15_pc_0.8_sFactor_0.002	0.9080	0.9208	0.9259	0.9256	0.9297	0.9398	0	5	-5
Menu plannings for 40 days									
Configuration	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.	W	L	R
SPEA2_ps_100_as_100_pm_0.025_pc_0.8	0.9111	0.9180	0.9254	0.9261	0.9340	0.9414	7	0	7
SPEA2_ps_100_as_100_pm_0.05_pc_0.8	0.9117	0.9180	0.9225	0.9234	0.9276	0.9418	6	0	6
SPEA2_ps_100_as_100_pm_0.075_pc_0.8	0.9023	0.9143	0.9197	0.9194	0.9243	0.9335	6	1	5
NSGA-II_ps_250_pm_0.05_pc_0.8	0.9012	0.9106	0.9147	0.9150	0.9209	0.9263	2	3	-1
NSGA-II_ps_250_pm_0.025_pc_0.8	0.8937	0.9071	0.9134	0.9140	0.9208	0.9287	0	3	-3
NSGA-II_ps_250_pm_0.075_pc_0.8	0.8874	0.9074	0.9135	0.9123	0.9180	0.9250	0	3	-3
IBEA_ps_250_pm_0.05_pc_0.8_sFactor_0.002	0.8942	0.9050	0.9091	0.9109	0.9170	0.9327	0	3	-3
IBEA_ps_250_pm_0.025_pc_0.8_sFactor_0.002	0.8947	0.9058	0.9108	0.9106	0.9168	0.9243	0	4	-4
IBEA_ps_250_pm_0.075_pc_0.8_sFactor_0.002	0.8975	0.9035	0.9096	0.9098	0.9147	0.9314	0	4	-4

Table 7 shows, for each meal plan length, statistical information related to the hypervolume values provided by each configuration of the different MOEAs at the end of the runs. The last three columns show data on the statistical ranking achieved by each configuration, which was calculated by following the same procedure explained in the first experiment. Note how, for a small number of days, i.e. 5 and 10 days, the MOEA that yielded the best performance at the end of the executions was the NSGA-II. When designing a meal plan for 5 days, the three highest-ranking configurations were the three parameterisations of the NSGA-II. In fact, each of those three configurations statistically outperformed every IBEA and SPEA2 configuration. When designing a meal plan for 10 days,

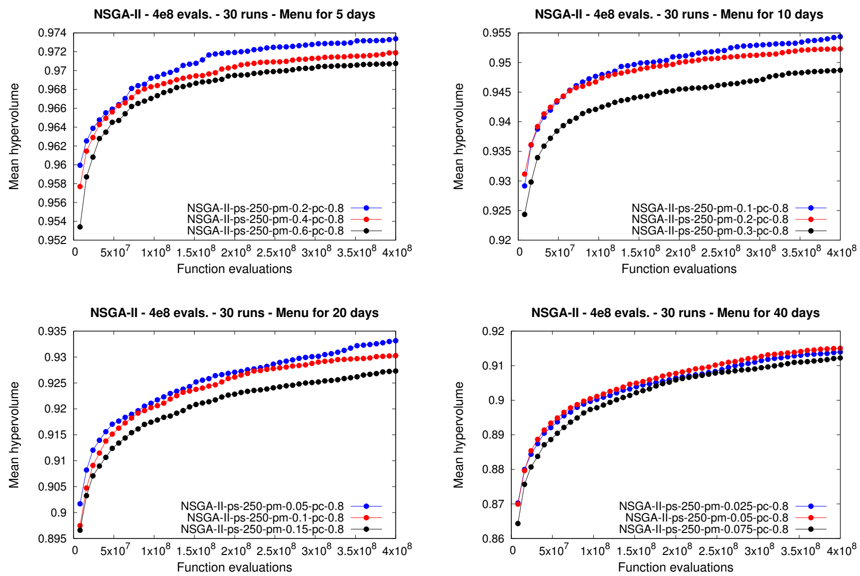

Figure 5: Evolution of the mean hypervolume values achieved by the NSGA-II when obtaining meal plans for 5, 10, 20 and 40 days

the three NSGA-II configurations also obtained the first three positions in the ranking. Although each of its configurations statistically outperformed every SPEA2 parameterisation, some of its configurations did not exhibit statistically significant differences in comparison to some parameterisations of the IBEA.

For longer meal plans, however, i.e. 20 and 40 days, the best performance at the end of the runs was attained by the SPEA2. When producing meal plans for 20 days, two configurations of the SPEA2 where the highest ranked. Those two configurations were able to statistically outperform or did not exhibit statistically significant differences with respect to all of the NSGA-II and IBEA configurations. Finally, when dealing with 40-day meal plans, the three SPEA2 configurations obtained the top three positions in the ranking, with each of them statistically outperforming each NSGA-II and IBEA configuration.

Regarding the mutation rate, note how small values yielded better performance in comparison to larger ones. In fact, when designing meal plans for 5 and 10 days, where the NSGA-II was the best-performing MOEA, those configurations that used the smallest mutation rate tested obtained the first position in their corresponding rankings. When dealing with longer meal plans, i.e. for 20 and 40 days, the SPEA2 behaved similarly, with lower mutation rates providing better performance in comparison to higher values of this parameter.

As already done in the first experiment, it is interesting to study the behaviour of the MOEAs both during and at the end of the executions. Figure 5 shows, for each meal plan length considered in this experiment, the evolution of the mean hypervolume attained by the NSGA-II. Figures 6 and 7 show the same

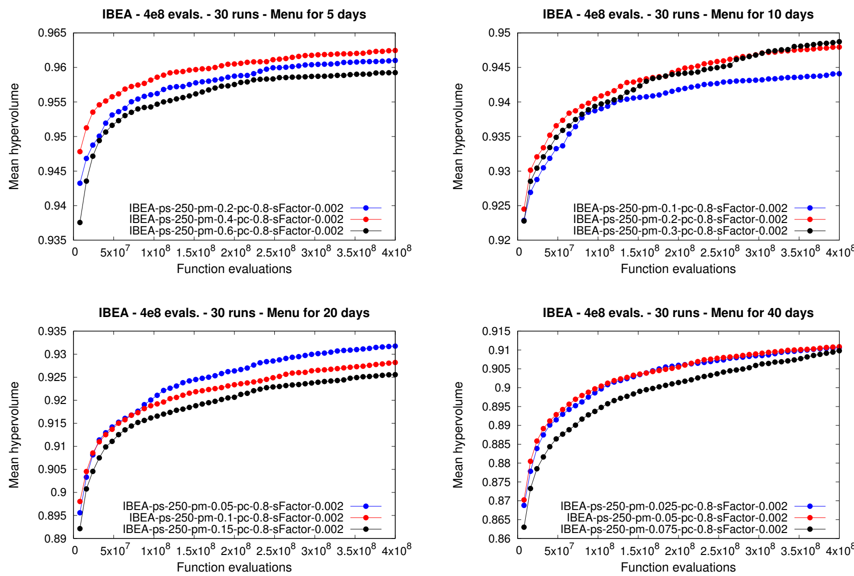

Figure 6: Evolution of the mean hypervolume values achieved by the IBEA when obtaining meal plans for 5, 10, 20 and 40 days

information for the IBEA and the SPEA2, respectively. Note how the three MOEAs exhibited differences during the runs depending on the specific value selected for the mutation rate. Although a few exceptions arose for some instances, particularly in the case of the IBEA, in general every approach that used lower values for the mutation rate was able to attain a higher mean of the hypervolume, and therefore, better performance, during almost every execution, in comparison to using larger mutation rates.

4.3. Third experiment: qualitative and quantitative analysis of solutions

In this third experiment, the main goal is to analyse the solutions attained by the MOEAs considered in this paper, not only from a quantitative point of view, but also from a qualitative one. Additionally, we would like to validate the novel multi-objective formulation of the MMPP we are providing herein in terms of the features of the solutions achieved. In order to do this, we considered, for each meal plan length, the solutions obtained in the second experiment by the best- and worst-performing configurations of each MOEA, i.e. the configurations obtaining the first and last positions in their corresponding ranking.

Figure 8 shows, for each meal plan length, examples of front approximations achieved at the end of the executions by those best- and worst-performing parameterisations. First of all, we should note that the shape of the different front approximations obtained demonstrates that the two objective functions are in conflict with each other. A decrease in the cost of a meal plan entails an increase in the frequency of course repetitions, and vice-versa. Bearing the above

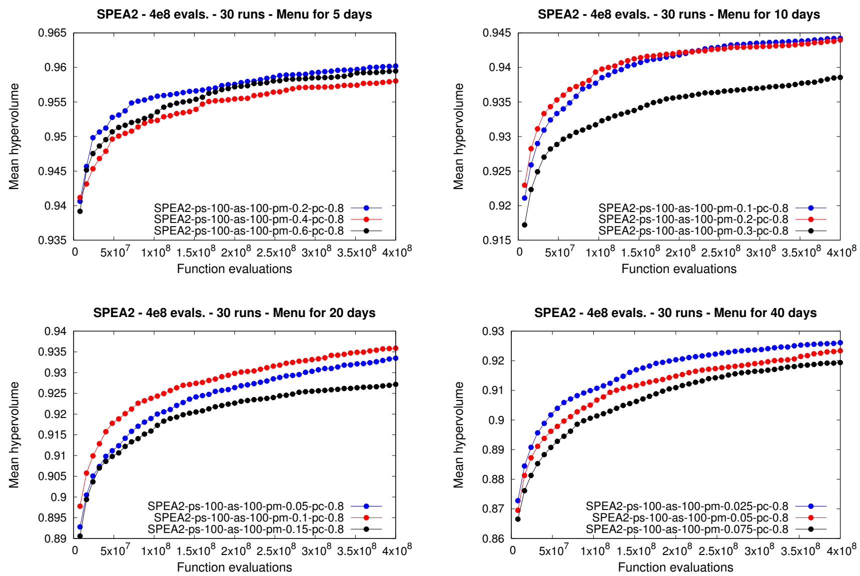

Figure 7: Evolution of the mean hypervolume values achieved by the SPEA2 when obtaining meal plans for 5, 10, 20 and 40 days

in mind, we can state that the novel formulation of the MMPP we are proposing herein is valid from the standpoint of its multi-objective nature.

Furthermore, it can be observed how differences among the front approximations attained by the best-performing and worst-performing configurations arose for each meal plan length. Generally speaking, for each case, the best-performing configuration was able to provide a better front approximation not only in terms of its convergence, but also in terms of its spread and uniformity, features that are closely related to each other and are usually referred to as the diversity of a front. For instance, when dealing with meal plans for 10 and 20 days, it holds that the corresponding best-performing configuration is able to provide a significant number of solutions that dominate a large number of solutions attained by the corresponding worst-performing configuration. The improvement in performance provided by the best configuration with respect to the worst one was even more noticeable when designing 40-day meal plans. The best-performing configuration was able to provide solutions located in regions of the front that are not covered by those solutions attained by the worst-performing configuration.

Finally, it would be interesting to analyse one of those solutions in depth, not only from the point of view of the different meals that it offers, but also from the point of view of its nutritional value. In order to do that, we selected one of the 10-day meal plans obtained by the corresponding best-performing configuration for which the values for the cost and the degree of repetition were 8.16 and 16.145, respectively. Table 8 enumerates the different meals

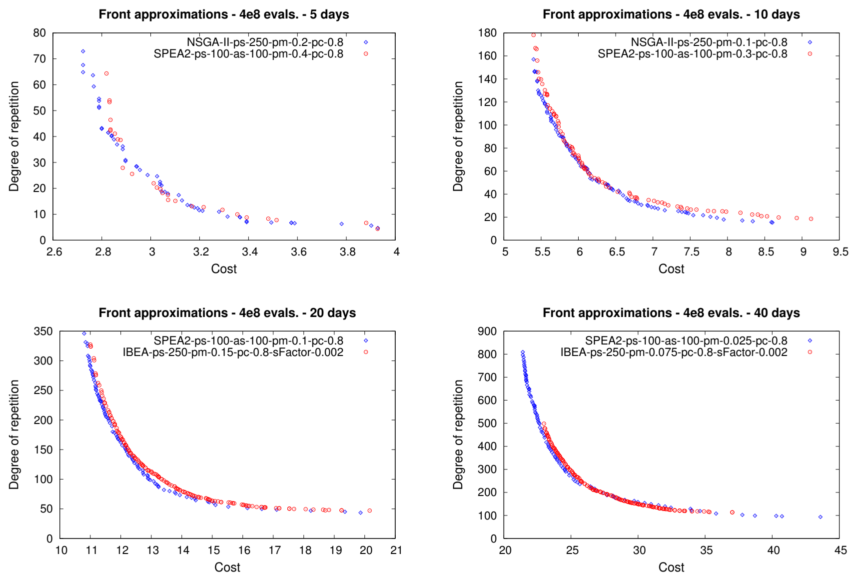

Figure 8: Examples of front approximations attained by the NSGA-II, IBEA and SPEA2 when dealing with meal plans for 5, 10, 20 and 40 days

in said meal plan, while Table 9 describes the total amount of nutrients it provides. According to the *White Book on Child Nutrition*, the main nutrients to consider in the case of lunch are carbohydrates, fats and proteins. The total number of kilocalories recommended for a child is 2000 *kcal* per day. Lunch accounts for approximately 35% of the total daily nutrient intake. As a result, the recommended number of kilocalories in a lunch should be approximately 700 *kcal*. In the meal plan depicted in Table 8 there is a total of 7269.34 *kcal*, which accounts for the recommended kilocalorie intake for 10.4 days. Regarding the total recommended amount of fats, they should contribute no more than 35% of the total amount of energy, meaning that for a lunch of 700 *kcal*, 245 *kcal* should come from fats. One gram of fat provides 9 *kcal*, so one lunch should contain 27.2 *g* of fat. The 10-day meal plan in our example provides 282.05 *g* of fat, which corresponds to the total amount of fats for 10.4 days. The recommended amount of carbohydrates is 50% of the total energy intake provided by one lunch. Consequently, 350 *kcal* should come from carbohydrates. One gram of carbohydrate provides 4 *kcal*, and as a result, a lunch should contain 87.5 *g* of carbohydrates. Our meal plan provides 769.24 *g* of carbohydrates, which corresponds to the total amount for 8.8 days. Finally, the recommended amount of proteins is 15% of the total energy provided by one lunch, which means that 105 *kcal* should come from proteins. One gram of protein provides 4 *kcal*, and therefore a lunch should contain 26.25 *g* of protein. Our meal plan example provides 291.08 *g* of proteins, which is equivalent to 11.1 days in terms of the intake recommendation. In conclusion, although the recommended amounts

Table 8: Example of menu planning for 10 days

Day	Starter	Main course	Dessert
1	White bean salad	Hawaiian Pizza	Fruits of the forest yogurt
2	Mashed potatoes	Tuna patties	Banana
3	Country salad	Cuba style rice	Tangerine
4	Rice soup	Rice three delights	Fried milk
5	Cooked soup	Stuffed eggs	Flan
6	White beans with clams	Ravioli with tomato	Apple
7	Chickpea salad	Breaded chicken thighs	Fruits of the forest yogurt
8	White bean salad	Tuna pizza	Strawberry and banana liquid yogurt
9	Chicken salad with mayonnaise	Potato omelette	Tangerine
10	Mashed vegetables	Marinera potatoes	Watermelon

Table 9: Example of the total amount of nutrients of a menu planning for 10 days. For each nutrient h, its corresponding amount r_h must be between the minimum (r_{min_h}) and maximum (r_{max_h}) amounts allowed. All amounts are for 10 days

Nutrient (h)	**Amount** (r_h)	**Min. amount** (r_{min_h})	**Max. amount** (r_{max_h})
Folic acid	2097.52 μg	450 μg	2295 μg
Calcium	2177.83 mg	1950 mg	9945 mg
Carbohydrates	769.24 g	291.66 g	1487.5 g
Energy	7269.34 kcal	2333.33 kcal	11900 kcal
Phosphorus	4679.52 mg	1875 mg	9562.5 mg
Total fat	282.05 g	91 g	464.1 g
Iron	105.02 mg	28.5 mg	145.35 mg
Magnesium	1283.6 mg	375 mg	1912.5 mg
Potassium	13053.8 mg	6750 mg	34425 mg
Proteins	291.08 g	116.66 g	595 g
Selenium	196.08 μg	85.83 μg	437.75 μg
Sodium	3020.12 mg	2900 mg	14790 mg
Vitamin A	2937.55 μg	1500 μg	7650 μg
Vitamin B1	4.6 mg	1.36 mg	6.97 mg
Vitamin B2	4.73 mg	2.1 mg	10.71 mg
Vitamin B6	7.97 mg	1.8 mg	9.18 mg
Vitamin B12	11.84 μg	7.6 μg	38.76 μg
Vitamin C	248.9 mg	90 mg	459 mg
Vitamin D	24.08 μg	15.5 μg	79.05 μg
Vitamin E	101.27 mg	21 mg	107.1 mg
Iodine	633.70 μg	225 μg	1147.5 μg
Zinc	41.21 mg	22.5 mg	114.75 mg

of nutrients may vary slightly from one recommendation to another, and in practical terms there exists a margin of excess and deficiency of nutrients above and below the recommended amounts, the meal plan selected for this study provides, in general terms, an adequate and balanced amount of nutrients.

5. Conclusions and further research

Conclusions. Most approaches found in the literature for menu planning deal with single-objective formulations of the problem. However, in this paper we propose a novel multi-objective formulation, termed MMPP, which considers two conflicting goals: minimising the menu cost and minimising the degree of repetition of the courses consumed. Our meal plans are designed for school cafeterias; as a result, the menu prices should not exceed reasonable amounts. At the same time, the courses should be varied to keep the students interested in the food.

The computational results attained by a set of well-known MOEAs show that suitable meal plans—in terms of their nutritional values—are provided. For short plans—5 and 10 days—the algorithm achieving the highest hypervolume values at the end of the executions is the NSGA-II. For longer plans—20 and 40 days—the best-performing algorithm is the SPEA2. In addition, the experimental analysis also demonstrates the multi-objective nature of our novel formulation. The cheapest meal plans are those with the lowest variety of courses, while the largest variety of courses is given by the most expensive plans.

Further research. This work provides an initial approach to the MMPP. During and after its development, we identified some aspects worthy of further research, and other features to consider for inclusion. A larger number of courses could be added to the database in order to provide a greater variety in the production of meal plans. The above, however, would increase the search space significantly. Seasonal products could also be considered as a determining factor in the price and quality of the food. Regarding the formulation of the problem, it would be interesting to carry out more tests with additional variation operators and even other MOEAs, including aggregation-based multi-objective optimisers, such as MOEA/D and Global WASF-GA, among others. Additionally, the repair method could be improved in such a way that it can obtain a feasible individual in a much smarter manner. Since the time invested in a single execution was significant, parallel approaches could be proposed aimed to speed up the computation time. For the plan generation, more objective functions and/or extra constraints based on other nutritional recommendations may be considered. Furthermore, the development of a tool that can help nutritionists to automatically evaluate pre-designed meal plans could be very promising. This tool could automatically calculate costs, the level of food repetition, nutritional values and other aspects. Finally, since different algorithms yielded the best results depending on the particular instance of the problem being solved, it would be interesting to apply machine learning techniques in order to classify instances in terms of their features. Hence, when solving a novel instance, the best-performing approach for that particular instance could be recommended in light of its characteristics.

Acknowledgements

This work was supported by the Spanish Ministry of Economy, Industry and Competitiveness as part of the programme "I+D+i Orientada a los Retos de la Sociedad", with contract number TIN2016-78410-R.

The problem formulation and restrictions follow the recommendations of the school cafeteria programme of the Government of the Canary Islands: "Programa de Eco-comedores Escolares de Canarias", which is a joint priority action of the "Instituto Canario de Calidad Agroalimentaria", the "Dirección General de Salud Pública del Gobierno de Canarias" and the "Dirección General de Ordenación, Innovación y Promoción Educativa".

References

[1] Aberg, J., 2009. An evaluation of a meal planning system: Ease of use and perceived usefulness. In: Proceedings of the 23rd British HCI Group Annual Conference on People and Computers: Celebrating People and Technology. BCS-HCI '09. British Computer Society, Swinton, UK, UK, pp. 278–287.

[2] Alfaro, H., Oct 2003. Menu planning using the exchange diet system. In: Systems, Man and Cybernetics, 2003. IEEE International Conference on. Vol. 3. pp. 3044–3049 vol.3.

[3] Bernardo, G., Proenca, R., Calvo, M., Fiates, G., Hartwell, H., 01 2015. Assessment of the healthy dietary diversity of a main meal in a self-service restaurant: A pilot study. British Food Journal 117, 286–301.

[4] Bulka, J., Izworski, A., Koleszynska, J., Lis, J., Wochlik, I., Feb 2000. Automatic meal planning using artificial intelligence algorithms in computer aided diabetes therapy. In: 2009 4th International Conference on Autonomous Robots and Agents. pp. 393–397.

[5] Chávez, O., Marchi, J., Pozos, P., 2014. Nutritional menu planning: A hybrid approach and preliminary tests. Research in Computing Science 82, 93–104.

[6] Deb, K., Pratap, A., Agarwal, S., Meyarivan, T., 2002. A fast and elitist multiobjective genetic algorithm: NSGA-II. IEEE Transactions on Evolutionary Computation 6, 182–197.

[7] Dixon, L. B., Cronin, F. J., Krebs-Smith, S. M., 2001. Let the pyramid guide your food choices: Capturing the total diet concept. The Journal of Nutrition 131 (2), 461S–472S.

[8] Drewnowski, A., Renderson, S. A., Driscoll, A., Rolls, B. J., 1997. The dietary variety score: Assessing diet quality in healthy young and older adults. Journal of the American Dietetic Association 97 (3), 266 – 271.

[9] Funabiki, N., Taniguchi, S., Matsushima, Y., Nakanishi, T., June 2011. A proposal of a menu planning algorithm for two-phase cooking by busy persons. In: 2011 International Conference on Complex, Intelligent, and Software Intensive Systems. pp. 668–673.

[10] Gaál, B., Vassányi, I., Kozmann, G., 2005. A novel artificial intelligence method for weekly dietary menu planning. Methods of Information in Medicine 44 (5), 655–664.

[11] Ginani, V. C., Zandonadi, R. P., Araujo, W. M. C., Botelho, R. B. A., 2012. Methods, instruments, and parameters for analyzing the menu nutritionally and sensorially: A systematic review. Journal of Culinary Science & Technology 10 (4), 294–310.

[12] Gumustekin, S., Senel, T., Cengiz, M., 2014. A comparative study on bayesian optimization algorithm for nutrition problem. Journal of Food and Nutrition Research 2 (12), 952–958.

[13] Hernández-Ocaña, B., Chávez-Bosquez, O., Hernández-Torruco, J., Canul-Reich, J., Pozos-Parra, P., 2018. Bacterial foraging optimization algorithm for menu planning. IEEE Access 6, 8619–8629.

[14] Hsiao, J., Chang, H., Oct 2010. Smartdiet: A personal diet consultant for healthy meal planning. In: 2010 IEEE 23rd International Symposium on Computer-Based Medical Systems (CBMS). pp. 421–425.

[15] Isokawa, T., Matsui, N., May 2015. Performances in ga-based menu production for hospital meals. In: 2015 IEEE Congress on Evolutionary Computation (CEC). pp. 2498–2501.

[16] Jothi, N., Husain, W., Damanhoori, F., Nov 2011. Myhealthcentral.com: Integrating knowledge based system in personalized diet plan and menu construction. In: 2011 International Conference on Research and Innovation in Information Systems. pp. 1–6.

[17] Kahraman, A., Seven, H., 2005. Healthy daily meal planner. In: Proceedings of the 7th Annual Workshop on Genetic and Evolutionary Computation. GECCO '05. ACM, New York, NY, USA, pp. 390–393.

[18] Kaldrim, E., Köse, Z., 2006. Application of a multi-objective genetic algorithm to the modified diet problem. Ph.D. thesis, Istanbul Technical University, Turkey.

[19] Kale, A., Auti, N., 2015. Automated menu planning algorithm for children: Food recommendation by dietary management system using id3 for indian food database. Procedia Computer Science 50, 197 – 202, big Data, Cloud and Computing Challenges.

[20] Kant, A., Graubard, B., 2004. Eating out in america, 1987-2000: trends and nutritional correlates. Preventive Medicine 38, 243–249.

[21] Kant, A., Schatzkin, A., Harris, T., Ziegler, R., Block, G., 1993. Dietary diversity and subsequent mortality in the first national health and nutrition examination survey epidemiologic follow-up study. The American Journal of Clinical Nutrition 57 (3), 434–440.

[22] Kashima, T., Han, H., Matsumoto, S., Ishii, H., June 2008. On a rough set with fuzzy weights in well-balanced menu planning system. In: 2008 3rd International Conference on Innovative Computing Information and Control. pp. 504–504.

[23] Kashima, T., Matsumoto, S., Ishii, H., 2009. Evaluation of menu planning capability based on multi-dimensional 0/1 knapsack problem of nutritional management system. IAENG International Journal of Applied Mathematics 39 (3).

[24] Khan, A., Hoffmann, A., 2003. Building a case-based diet recommendation system without a knowledge engineer. Artificial Intelligence in Medicine 27 (2), 155 – 179.

[25] Kovásznai, G., May 2011. Developing an expert system for diet recommendation. In: 2011 6th IEEE International Symposium on Applied Computational Intelligence and Informatics (SACI). pp. 505–509.

[26] Lee, C.-S., Wang, M.-H., Hagras, H., April 2010. A type-2 fuzzy ontology and its application to personal diabetic-diet recommendation. Fuzzy Systems, IEEE Transactions on 18 (2), 374–395.

[27] León, C., Miranda, G., Segura, C., 2009. METCO: a parallel plugin-based framework for multi-objective optimization. International Journal on Artificial Intelligence Tools 18 (04), 569–588.

[28] Leung, P., Wanitprapha, K., Quinn, L., 1995. A recipe-based, diet-planning modelling system. British Journal of Nutrition 74 (2), 151–162.

[29] Moreira, R., Wanner, E., Martins, F., Sarubbi, J., 2017. The menu planning problem: A multiobjective approach for brazilian schools context. In: Proceedings of the Genetic and Evolutionary Computation Conference Companion. GECCO '17. ACM, New York, NY, USA, pp. 113–114.

[30] Ngo, H., Cheah, Y., Goh, O., Choo, Y., Basiron, H., Kumar, Y. J., 01 2016. A review on automated menu planning approaches. Journal of Computer Science 12, 582–596.

[31] Noah, S. M., Abdullah, S. N., Shahar, S., Abdul-Hamid, H., Khairudin, N., Yusoff, M., Ghazali, R., Mohd-Yusoff, N., Shafii, N. S., Abdul-Manaf, Z., 02 2004. Dietpal: A web-based dietary menu-generating and management system. Journal of medical Internet research 6, e4.

[32] Osthus, D., 2011. A genetic algorithm approach to optimize planning of food fortification. Ph.D. thesis, Department of Statistics, ISU.

[33] Petot, G., Marling, C., Sterling, L., 1998. An artificial intelligence system for computer-assisted menu planning. Journal of the American Dietetic Association 98 (9), 1009–1014.

[34] Saborido, R., Ruiz, A. B., Luque, M., Jun. 2017. Global WASF-GA: An evolutionary algorithm in multiobjective optimization to approximate the whole pareto optimal front. Evol. Comput. 25 (2), 309–349.

[35] Segura, C., Coello, C., Segredo, E., Aguirre, A. H., Dec 2016. A novel diversity-based replacement strategy for evolutionary algorithms. IEEE Transactions on Cybernetics 46 (12), 3233–3246.

[36] Seljak, B. K., June 2006. Dietary menu planning using an evolutionary method. In: 2006 International Conference on Intelligent Engineering Systems. pp. 108–113.

[37] Steyn, N., Nel, J., Nantel, G., Kennedy, G., Labadarios, D., 2006. Food variety and dietary diversity scores in children: are they good indicators of dietary adequacy? Public Health Nutrition 9 (5), 644–650.

[38] Stigler, G., 1945. The cost of subsistence. Journal of Farm Economics 27 (2), 303–314.

[39] Veiros, M. B., Da Costa Proença, R. P., Kent-Smith, L., Hering, B., De Sousa, A. A., 2006. How to analyse and develop healthy menus in foodservice. Journal of Foodservice 17 (4), 159–165.

[40] Wang, F., Yuan, Y., Pan, Y., Hu, B., Dec 2008. Study on the principles of the intelligent diet arrangement system based on multi-agent. In: 2008 Second International Symposium on Intelligent Information Technology Application. Vol. 1. pp. 264–268.

[41] Zhang, Q., Li, H., Dec 2007. Moea/d: A multiobjective evolutionary algorithm based on decomposition. IEEE Transactions on Evolutionary Computation 11 (6), 712–731.

[42] Zitzler, E., Deb, K., Thiele, L., July 1999. Comparison of Multiobjective Evolutionary Algorithms on Test Functions of Different Difficulty. In: Wu, A. S. (Ed.), Proceedings of the 1999 Genetic and Evolutionary Computation Conference. Workshop Program. Orlando, Florida, pp. 121–122.

[43] Zitzler, E., Künzli, S., 2004. Indicator-Based Selection in Multiobjective Search. In: VIII Conference on Parallel Problem Solving from Nature. Vol. 3242 of LNCS. Springer, pp. 832–842.

[44] Zitzler, E., Laumanns, M., Thiele, L., 2002. SPEA2: Improving the Strength Pareto Evolutionary Algorithm for Multiobjective Optimization. In: Evolutionary Methods for Design, Optimization and Control. CIMNE, Barcelona, Spain, pp. 19–26.

[45] Zitzler, E., Thiele, L., Laumanns, M., Fonseca, C. M., da Fonseca, V. G., April 2003. Performance assessment of multiobjective optimizers: an analysis and review. IEEE Transactions on Evolutionary Computation 7 (2), 117–132.