



Designing healthier and acceptable diets using data envelopment analysis

Argyris Kanellopoulos^{1,*} , Johanna C Gerdessen¹, Ante Ivancic¹,
Johanna M Geleijnse^{2,3}, Jacqueline M Bloemhof-Ruwaard¹ and Pieter van't Veer²

¹Operations Research and Logistics Group, Wageningen University, 6706 KN Wageningen, The Netherlands:

²Division of Human Nutrition and Health, Wageningen University, 6708 WE Wageningen, The Netherlands:

³Top Institute Food and Nutrition (TiFN), Wageningen, P.O. Box 557, 6700, AN, Wageningen, The Netherlands

Submitted 11 April 2019: Final revision received 29 October 2019: Accepted 12 November 2019: First published online 17 April 2020

Abstract

Objective: The objective of this research is to propose methodology that can be used to benchmark current diets based on their nutrient intakes and to provide guidelines for improving less healthy diets in a way that is acceptable for the studied population.

Design: We discuss important limitations of current diet models that use optimisation techniques to design healthier and acceptable diets. We illustrate how data envelopment analysis could be used to overcome such limitations, and we describe mathematical models that can be used to calculate not only healthier but also acceptable diets.

Setting: We used data from the Nutrition Questionnaires plus dataset of habitual diets of a general population of adult men and women in The Netherlands (n 1735).

Participants: Adult population.

Results: We calculated healthier diets with substantial higher intakes of protein, fibre, Fe, Ca, K, Mg and vitamins, and substantially lower intakes of Na, saturated fats and added sugars. The calculated diets are combinations of current diets of individuals that belong to the same age/gender group and comprise of food item intakes in proportions observed in the sample.

Conclusions: The proposed methodology enables the benchmarking of existing diets and provides a framework for proposing healthier alternative diets that resemble the current diet in terms of foods intake as much as possible.

Keywords
Diet model
Benchmark
Efficiency
Nutrition
DEA
Public health

Rising incomes and urbanisation lead to global dietary changes. Traditional diets are replaced by diets higher in energy, refined sugars, saturated fats and meats. Such dietary changes can increase the incidence of chronic diseases (e.g. type II diabetes and CHD)⁽¹⁾. Replacing the current diets with healthier alternatives will have substantial public health and probably also environmental benefits^(2,3).

Designing healthier diets is a rather complex process. The variety and adequacy of nutrient intake of the various subgroups of the population must be taken into account while considering dietary preferences of individuals.

Diet models have been developed to deal with the complexity of designing such alternative, healthier diets^(4–6). These models are used to determine the optimal quantities of available food items that should be included in a diet of an individual to optimise specific criteria (e.g. improve healthiness or minimise cost) under specific nutritional and acceptability constraints^(4–6).

Combining quantities of different food items to optimise a specific objective accounting only for nutritional constraints can result in optimal diets that are far from what people actually eat. Probably, such diets will not be accepted by individuals, which makes them less relevant. For that reason, in diet models, acceptability constraints and minimisation of the deviation from observed diets (at population, sub-population or individual level) are used to improve the acceptance of optimised diets⁽⁷⁾. Acceptability constraints have the form of upper and lower limits to the intake of specific food items^(5,7,8) and penalties for deviations from current energy intake and serving size⁽⁹⁾. Defining acceptability constraints is a challenging process that often involves expert knowledge and a substantial degree of subjectivity. Similarly, minimising deviation of the calculated diet from an observed diet involves the implicit assumption that food items which are not consumed currently are also not preferred.

*Corresponding author: Email argyris.kanellopoulos@wur.nl

In the current study, to avoid defining such acceptability constraints explicitly and imposing additional implicit assumptions, instead of trying to compose new diets from available food items, we look within the existing diets of a certain population and identify the most healthy ones. These current 'healthier' diets are used as benchmark for the diets of other individuals in the population with less healthy diets. Benchmarking current diets instead of composing new diets from existing food items is definitely a novel viewpoint in diet modelling that can result in calculated diets that are not only healthier but also acceptable alternatives for the population under study.

A commonly used benchmarking technique in operations research literature is data envelopment analysis (DEA)⁽¹⁰⁾. DEA has been used in many fields like banking, health care, agricultural economics, transportation and education^(11–15). In general, the aim of DEA is to identify decision-making units that convert inputs (i.e. less-is-better criteria) into outputs (i.e. more-is-better criteria) in the most efficient way. Within the DEA context, the diet of an individual is seen as a decision-making unit that can be evaluated based on the intake of multiple less-is-better 'unhealthy' nutrients and more-is-better 'healthy' nutrients. Despite the broad variety of applications of DEA, it has not been used to benchmark healthier diets.

The objective of this research is to propose a DEA-based framework that can be used to benchmark current diets based on their nutrient intakes and to provide guidelines for improving less healthy diets in a way that is acceptable by the studied population.

Methods

In this section, first we present the main structure of existing diet models and we discuss important limitations and

challenges. We use an illustrative example to present DEA as a novel approach for benchmarking and re-designing healthier diets. Finally, we use this simple example to propose a mathematical programming model that enables to account for acceptability considerations in diet modelling.

Diet models for designing healthier diets

Existing diet models are often mathematical programming models that aim to compose an alternative optimal diet for a specific individual or group of individuals⁽⁹⁾. A schematic representation of a mathematical programming model that focuses on optimising diets is presented in Fig. 1a.

The main components of mathematical programming diet models are (i) the decision variables ($q_i \forall i \in (1 \dots n)$) which are the quantities of available food items i that should be included in the calculated diet in order to achieve a specific objective, (ii) the objective function which is an indicator to be optimised expressed as a function of the decision variables (e.g. the objective function could be the minimisation of the cost or the CO₂ emissions of the diet; or even the minimisation of the differences between food intakes of an observed and a calculated diet), (iii) the nutritional constraints that are used to either impose the dietary guidelines to the calculated diet (through hard constraints) or to calculate deviations between nutrient (and food) intakes of the calculated diet and the dietary guidelines through soft constraints (i.e. constraints that can be violated but violation of these constraints is associated with a penalty in the objective function which is minimised) and (iv) the acceptability constraints which aim to facilitate acceptance of the diet by imposing restrictions on food item quantities following current consumption patterns. Without acceptability constraints, the calculated diet might comprise of food items that do not necessarily form a realistic diet.

Defining acceptability constraints is often based on expert knowledge and information on current meals.

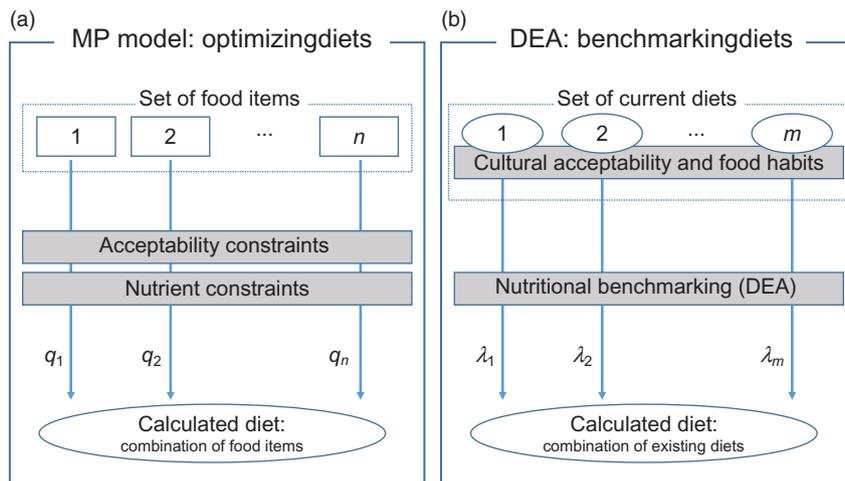


Fig. 1 (colour online) Schematic representation of the differences between current mathematical programming (MP) diet models, which focus on optimising diets by deciding on optimal intakes of available food items, and the proposed benchmarking approach, which focuses on identifying efficient current diets and combines them to healthier alternatives

This restricts the set of feasible diets in an arbitrary and subjective way which introduces an important level of bias to the optimal food item intakes of the calculated diet.

To avoid formulating explicit acceptability constraints and imposing the current diet as a reference, we propose an alternative diet optimisation method based on nutritional benchmarking. Nutritional benchmarking is defined as the comparison of diets based on their nutrient (and/or food item) intakes. The objective is to identify for each individual that belongs to the same population group combinations of current diets that are healthier or at least as healthy as the current diet. The proposed method is summarised in Fig. 1b. Because the new calculated diet is a combination of other current diets, acceptability considerations are taken into account implicitly. The resulting diet is not a mix of individual food items but a mix of food items that either have been chosen together in meals or have been declared in dietary surveys by individuals in the sample. The diets that are used to compose all diets of the population under study are identified based on nutritional benchmarking and DEA⁽¹⁰⁾.

DEA with a simple example

This section illustrates the basic concepts of DEA based on a simple two-dimensional example. DEA aims to compare decision-making units based on their capacity to convert multiple inputs into multiple outputs⁽¹⁰⁾. In the specific case of a diet problem, the decision-making units are individual diets and outputs are the more-is-better nutrients like vitamins, fibre and protein, while inputs are the less-is-better nutrients like saturated fat and Na. The objective is to identify those diets that have a higher ratio of more-is-better nutrients content per unit of less-is-better nutrient.

To demonstrate the method graphically, we assume that diets are evaluated based only on two nutrients that is, dietary fibre (as the more-is-better nutrient) and Na (as the less-is-better nutrient). Figure 2 involves six individual

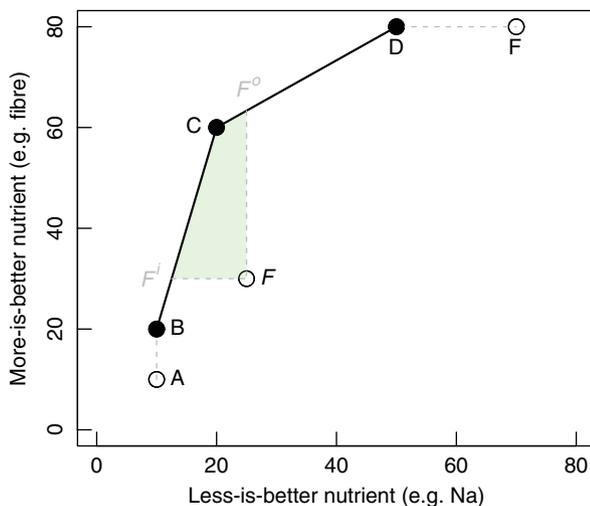


Fig. 2 (colour online) A two-dimensional illustrative example of data envelopment analysis for benchmarking diets

diets with the same energy intakes labelled with letters (A, B, C, D, E and F) and their performance with respect to dietary fibre and Na content (e.g. diet C contains 60 units of dietary fibre and 20 units of Na).

DEA aims to compare each diet with all other diets in the sample and identify those that are efficient that is, those diets that for a certain level of less-is-better nutrients contain the highest (compared to all others) level of more-is-better nutrients or those that for a certain level of more-is-better nutrients contain the lowest level of less-is-better nutrients. From Fig. 2, it is visible that diets B, C and D are DEA-efficient because for their intake of Na there is no linear combination of all other diets with higher or the same intake of fibre. The line segments BC and CD are called the DEA-efficient frontier. The DEA model assigns to DEA-efficient diets an efficiency score of 1. It is assumed that all other (inefficient) diets can be projected to the efficient frontier and either increase more-is-better nutrients for the same amount of less-is-better nutrients, that is, output-oriented DEA model (OO DEA) or decrease the amount of less-is-better nutrients for the same level of more-is-better nutrients, that is, input-oriented DEA model (IO DEA). All inefficient diets (i.e. diets A, E and F) receive an efficiency score of < 1 . The efficiency score can be interpreted as the distance to the efficient frontier. The lower the efficiency score of a certain diet, the larger the distance of the diet to the efficient frontier.

In our illustrative example of Fig. 2, diet A is not efficient because it can be replaced by diet B which has the same Na content but higher dietary fibre content. Diet E is inefficient because it can be replaced by diet D which has a lower Na content for the same dietary fibre content. Finally, diet F is inefficient because by combining diet B and C we create a new diet on the efficient frontier that is, diet F^I that has the same dietary fibre level for substantially less Na (IO DEA model). In the IO DEA model, diets B and C are called the peers of diet F. Similarly, diet F can be replaced by diet F^O and increase level of dietary fibre for the same level of fat (OO DEA model). In the OO DEA model, the peers of diet F are diets C and D.

In practice, the healthiness of a diet is determined by more than one more-is-better nutrient and more than one less-is-better nutrient, which makes an informative graphical representation of the problem impossible. To deal with the multi-dimensionality of the problem and benchmark current diets, we solve a sequence of linear programming models, that is, the IO DEA and the OO DEA models presented in online Appendix A. The decision variables of the DEA models are the proportion of each diet of the sample in the calculated diet, and the objective function is an efficiency score that represents the distance of the evaluated diet from the efficient frontier. The orientation of the DEA model (IO DEA or OO DEA) determines the direction of measuring the distance to the efficient frontier. The result of the optimisation is a set of healthier diets which are combinations of other diets that belong to the same group.



This enables to account for acceptability of the diets without imposing additional acceptability constraints or minimising the deviation from the observed diet. In the next step, we identify the healthier diet that minimises the deviation from the current food item intakes.

Minimising deviation from current food intake (MINDV model)

The IO DEA and OO DEA models identify the subset of diets that are DEA-efficient. For each inefficient diet, an alternative healthier diet is calculated as a linear combination of its peers. The new 'efficient' and healthier alternative is identified by minimising the distance to the DEA-efficient frontier. For example, for diet F , a new diet F^i is calculated that is healthier (less Na for the same dietary fibre level). However, the new healthier diet (F^i) might include food items completely different than the current diet. Using DEA to compare and benchmark diets that belong to the same subgroup (i.e. comparison between similar peers) can reduce this problem.

Actually, all linear combinations of the efficient diets within the shaded area F^iFF^oC of Fig. 2 are healthier alternatives of the current diet F because they contain less Na and more dietary fibre than the current diet F . In line with existing diet optimisation models⁽⁷⁾, we assume that the willingness of an individual to accept an alternative healthier diet increases with the resemblance (in terms of food item intakes) of this alternative healthier diet with the current diet of this individual. Therefore, within the set of healthier diets of the shaded area F^iFF^oC , we search for the diet that is most similar to the current diet in terms of food item intakes.

For this reason, we propose minimum deviation (MINDV) model, which combines the efficient diets, identified with the DEA models and calculate for each individual an alternative healthier diet which resembles as much as possible, in terms of food item intakes, the current diet. It is important to emphasise that similar to the DEA models the MINDV model combines existing diets instead of individual food items (current practice in diet optimisation). Because of this, the calculated diets are as close as possible to the current diet, but they also comprise of food item intakes in proportions that match current diets of individuals within the same group. To achieve this, we minimise the total absolute deviation between the food item intakes of the healthier alternative diet and the food item intakes of the current diet. We make sure that, compared to the current diet, the healthier alternative diets have lower or equal intakes of less-is-better nutrient and more or equal intakes of more-is-better nutrients. The mathematical formulation of the MINDV model is presented in online Appendix B.

Case study: designing alternative healthier diets in The Netherlands

To demonstrate how DEA can be used to benchmark diets and calculate alternative healthier diets, we used data from

the Nutrition Questionnaires plus dataset of habitual diets of a general population of adult men and women in The Netherlands⁽¹⁶⁾. Habitual diet was assessed by means of a FFQ from which the intake of foods and nutrients was calculated.

To make diets comparable, we defined six groups of individuals: (F1) females 20–40 years of age (n 206), (F2) females 41–50 years of age (n 203), (F3) females > 50 years of age (n 425), (M1) males 20–40 years of age (n 121), (M2) males 41–50 years of age (n 155) and (M3) males > 50 years of age (n 625).

The nutrient and food intakes of the diets are standardised to a diet of 8368 kJ (i.e. 2000 kcal). For example to normalise the nutrient and food intakes of a diet of 12 552 kJ (i.e. 3000 kcal), we multiplied average nutrient and food intakes of this diet with a factor of $\frac{2}{3}$. In the current study, for assessing the healthiness of a diet, we used the nine encouraged and the three discouraged nutrients used to calculate the Nutrient Rich Diet (9,3) index^(17,18). The nine more-is-better nutrients are protein, fibre, Ca, Fe, Mg, K and vitamins A, D and E. The three less-is-better nutrients are Na, saturated fat and added sugars. Na intake was estimated from foods only, ignoring discretionary salt use during cooking and dining.

We defined sufficient levels of intake for all nutrients included in Nutrient Rich Diet 9,3 based on dietary reference intakes for Europe⁽¹⁹⁾ and the USA⁽²⁰⁾. The sufficient levels for more-is-better nutrients that we chose to use in our model would be adequate to meet the nutrient requirements of 97.5 % of adult individuals in the population. The sufficient daily intake of protein was assumed to be 75 g, of Ca 1200 mg, of Fe 18 g, of Mg 420 mg, of K 4700 mg, of vitamin A 750 retinol equivalents (RE), of vitamin C 110 mg and of vitamin E 15 mg. We considered that increasing intakes above that sufficient level is of less importance, and thus improving the intakes of the other nutrients becomes a priority. Therefore, we capped nutrients at the sufficient intake level in the DEA model. However, we did not cap the intakes of dietary fibre assuming that more is better. For less-is-better nutrients, the lowest possible intake level was considered desirable. The mean and SD of nutrient intakes of selected nutrients, per group of individuals, are presented in Table 1.

DEA was used to benchmark diets that belong to the same of the six consumer groups, that is, we only compared diets that belong to the same consumer group. We used both the IO DEA and the OO DEA models to identify the efficient diets. The MINDV model was used to combine DEA-efficient diets and compose for each current diet a new alternative diet that is healthier than the current diet but at the same time as similar as possible to the food group item (based on the Dutch food composition table⁽²¹⁾) intakes, expressed in percentage of the diet's mass of the current diet.

Furthermore, to quantify the importance of DEA-efficient diets, we calculated the frequency that each

Table 1 Used Nutrition Questionnaires plus variables per gender-age group of individuals

	Pr (>F)	Female groups						Male groups					
		F1		F2		F3		M1		M2		M3	
		Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Age (years)	0.000	32	6	46	3	59	5	33	5	47	3	61	5
BMI (kg/m ²)	0.000	25	5	25	5	26	4	25	3	26	4	27	3
Total energy intake (kJ/d)	0.000	7979	1749	7602	1753	7272	1615	9950	2648	9506	2402	8673	2247
Standardised nutrient intakes to a 8368 kJ (i.e. 2000 kcal) diet													
Protein (g/d)	0.091	77	15	79	16	81	13	78	14	77	13	80	13
Dietary fibre (g/d)	0.000	23	7	23	7	23	6	21	6	21	6	21	6
Ca (mg/d)	0.000	928	296	1010	304	1080	306	857	281	860	319	921	283
Fe (mg/d)	0.000	11	3	12	2	12	3	10	2	11	2	11	2
Mg (mg/d)	0.007	338	79	367	67	375	62	339	68	341	61	348	62
K (mg/d)	0.033	3095	726	3339	662	3554	647	2934	747	3128	581	3231	651
Vitamin A (RE/d)	0.760	920	728	995	707	995	581	852	678	792	451	973	576
Vitamin C (mg/d)	0.000	104	65	103	50	115	60	70	45	81	48	83	48
Vitamin E (mg/d)	0.000	12	5	11	4	11	4	10	4	10	3	10	4
Added sugars (g/d)	0.000	58	28	56	23	47	18	55	23	52	22	47	19
Saturated fat (g/d)	0.564	29	7	29	6	29	7	29	7	29	6	29	6
Na (mg/d)	0.230	2455	566	2403	620	2400	573	2533	617	2451	627	2463	550

The *P* values of pairwise *t* test comparisons between the groups of individuals are presented in online Appendix C.

DEA-efficient diet is used as peer of inefficient diets using the following equation (1):

$$FRQ_k = 100 \frac{\sum_{l \neq k} \lambda_{l,k}^*}{\sum_{l,q} \lambda_{l,q}^*} \% \quad \forall k \quad (1)$$

where FRQ_k is the frequency (%) of the efficient diet k as a peer, and $\lambda_{l,k}^*$ is the weight (value from 0 to 1) of diet k in the efficient alternative of diet l .

Results

In this section, we present the results and main outputs of the proposed DEA-based framework which was applied on the Nutrition Questionnaires plus dataset⁽¹⁶⁾. The variables of the Nutrition Questionnaires plus dataset which are used in the current study are summarised in Table 1.

We present the potential improvements on the nutrient intakes of the diets calculated with the three different models, and we present the corresponding changes in terms of food item intakes. We analyse the peers, which are the DEA-efficient diets that have been combined to design alternative healthier diets.

The potential to decrease less-is-better nutrients or increase more-is-better nutrients by replacing current diets with healthier alternatives as calculated by the IO DEA, OO DEA and MINDV models is presented in Fig. 3.

The alternative diets calculated by the different models result in substantial improvements. All models resulted in alternative diets with lower levels of less-is-better nutrients and higher levels of more-is-better nutrients for all six consumer groups. Exception is the intake of vitamin A in the alternative diets of older females (group F3) which decreases marginally. The reason for this decrease is

related to the capping of the intakes of more-is-better nutrients to their optimal levels and the high intakes of vitamin A in current diets. The vitamin A intake of the alternative healthier diets exceeds the optimal level (i.e. intake used for capping) of vitamin A, but it is, on average, lower than the intakes of the current diets.

As expected, the MINDV model results in the smallest improvements. The MINDV model aims to calculate diets that are as close as possible to current diets as long as they have at least the same level of more-is-better nutrients and at most the same level of less-is-better nutrients with the current diet. The results of all models show that the largest improvement can be achieved in diets of younger females and males (i.e. group of individuals F1 and M1).

To assess the nutrient adequacy of the calculated diets, we calculated the deviations of important nutrient intakes from dietary recommendations. All three models resulted in calculated improved diets with intakes of protein, vitamin A and vitamin C much higher than the intake which is considered adequate to meet the nutrient requirements of 97.5% of adult individuals in the population (Fig. 4). The calculated diets, even though they comprise higher intakes of Fe than the observed diets, do not reach the dietary-specific requirements of Fe (Fig. 4). For some groups, the intakes of vitamin E, K and Ca are higher than the observed diets but still below the recommendations.

The proposed alternative healthier diets can be further analysed to the level of intakes of specific food items. The food group item intake expressed as a percentage of the total intake (in g) of the observed diets, and the diets calculated with the IO DEA, OO DEA and MINDV models are presented in Fig. 5. We present result as a percentage of the total intake to enable for comparisons of the changes between the food items. The mass of the calculated diets might be different than the observed diets, but the energy

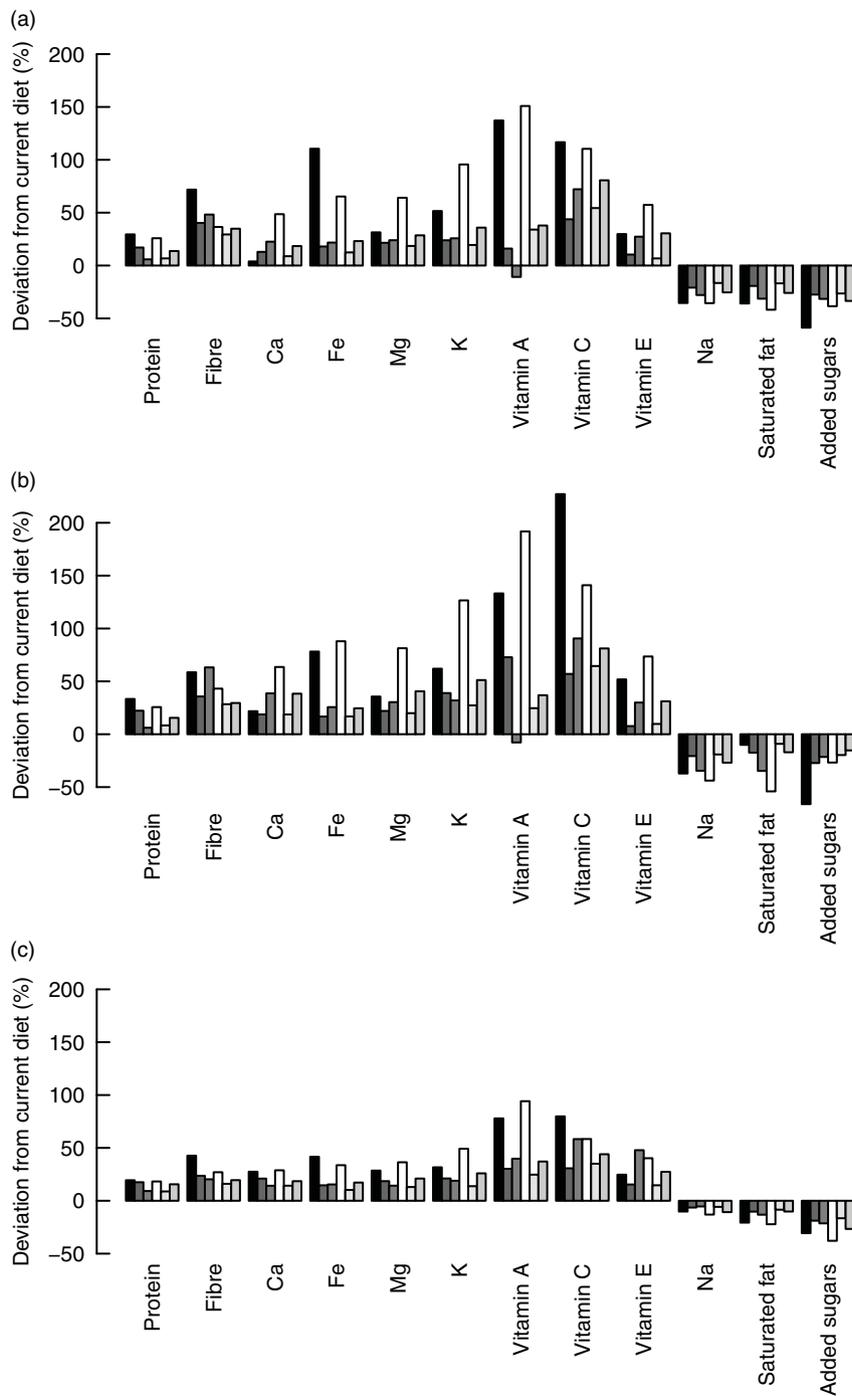


Fig. 3 Difference between efficient and observed nutrient intakes as calculated with the (a) input-oriented data envelopment analysis (DEA), (b) output-oriented DEA and (c) MINDV models. ■, F1; ■, F2; ■, F3; □, M1; □, M2; □, M3

intake of the diets is constant and equal to 8368 kJ (i.e. 2000 kcal).

In general, compared to the current diets, the alternative healthier diets calculated with the three models contain higher intakes of fruits and vegetables. Calculated healthier diets of younger males contain higher intakes of potatoes, milk and dairy, while healthier diets of females include lower intakes of potatoes and higher intakes of eggs. The quantities

of fish increase in the improved diets of all groups except from those of younger males and females. This is because of the lower quantities of fish in current diets of younger individuals and different types of dishes. As expected, the most similar to the current diets in terms of food group item intakes are the diets calculated with the MINDV model.

Healthier diets also contain lower intakes of bread; nuts and snacks; fats and oils; soups, pastry and cakes; and

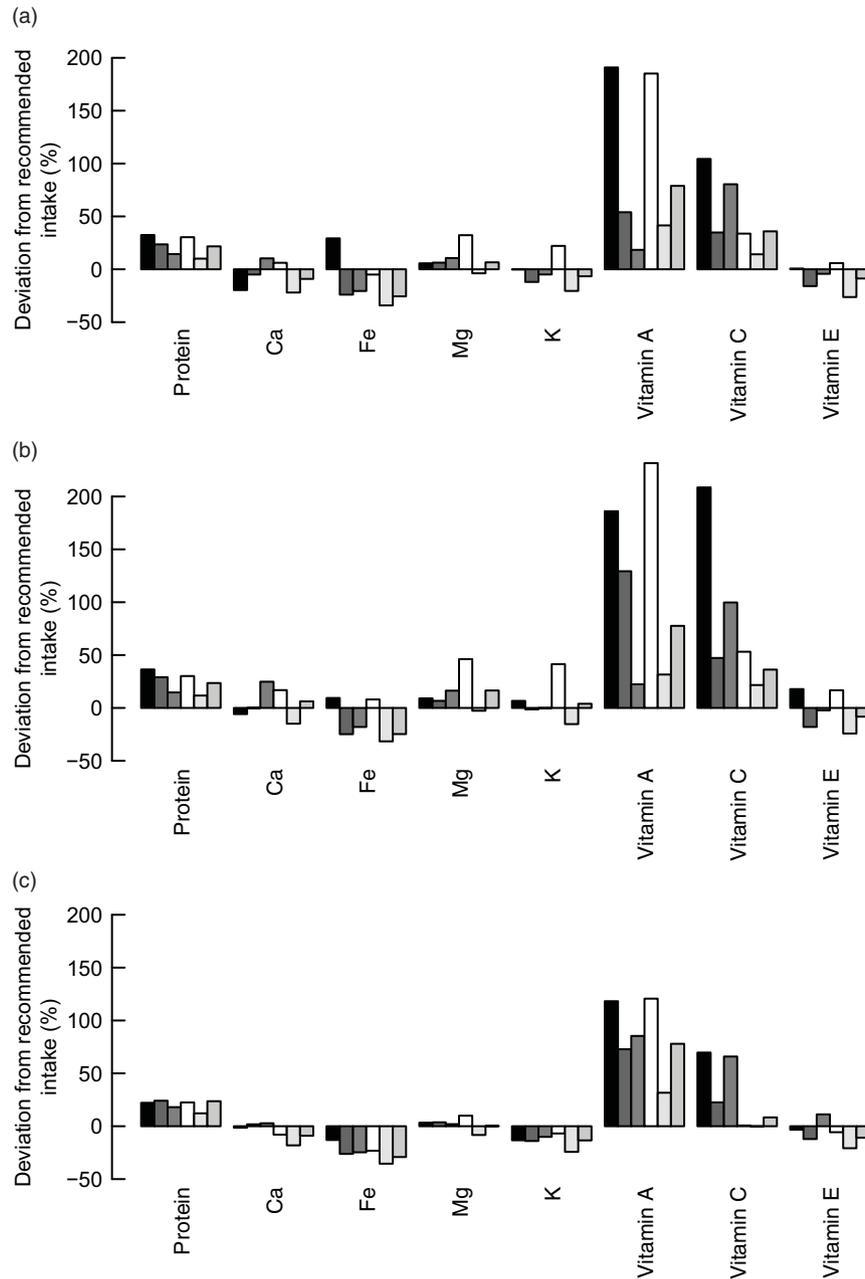


Fig. 4 Difference between the dietary nutrient requirements of 97.5 % of adult individuals in the population and the calculated improved nutrient intakes with the (a) input-oriented data envelopment analysis (DEA), (b) output-oriented DEA and (c) MINDV models. ■, F1; ■, F2; ■, F3; □, M1; □, M2; □, M3

sugar/sweets. Decrease of intakes of bread and the nuts, seed and snacks groups is not in line with existing dietary guidelines that recommend the increase of nuts, seeds and whole wheat bread. To investigate this contradiction between model's results and dietary recommendations, we investigated the diet composition of the nuts, seed and snacks group. Figures 6 and 7 represent the average food item intake of the nuts, seeds and snacks food item group for females and males, respectively. In general and in line with the current dietary recommendations calculated, diets recommend decreased quantities of snacks

but increased quantities of unsalted nuts and food items with high content of unsaturated (healthy) fats.

Identifying alternative diets that are important for improving the current diets of the population as a whole can provide valuable information to researchers and policy makers for designing healthier diets. To quantify the importance of an efficient diet, we used as indicator the frequency of this diet as a peer for other inefficient diets (Fig. 8). Benchmarked diets with higher frequency scores (calculated according to eq. 1) can provide guidelines for defining nutritional and policy goals.

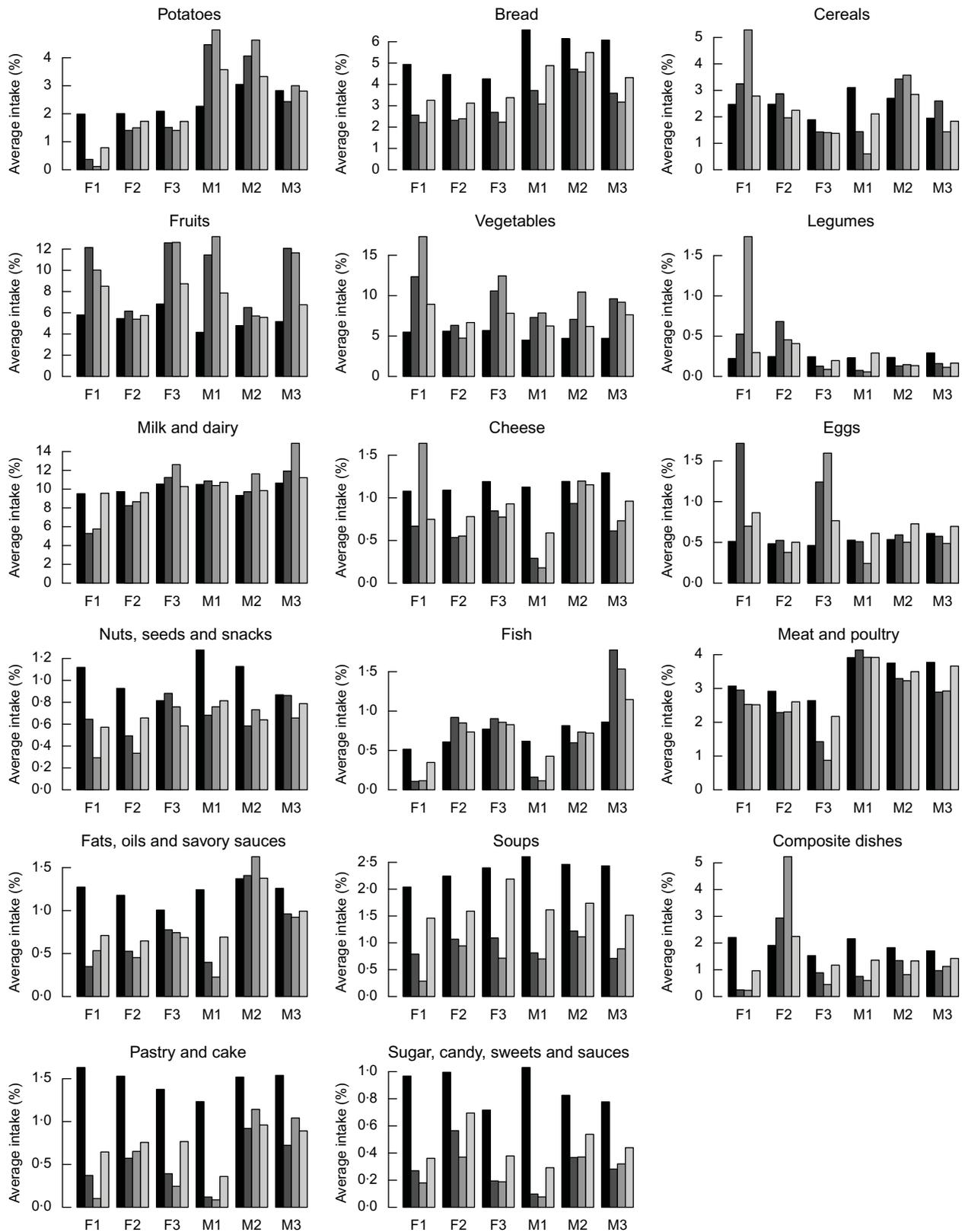


Fig. 5 Average daily intakes of important food groups (expressed in percentage of total weight of the diet) in the current (■) and the calculated diets with the input-oriented data envelopment analysis (DEA) (■), output-oriented DEA (■) and MINDV (■) models

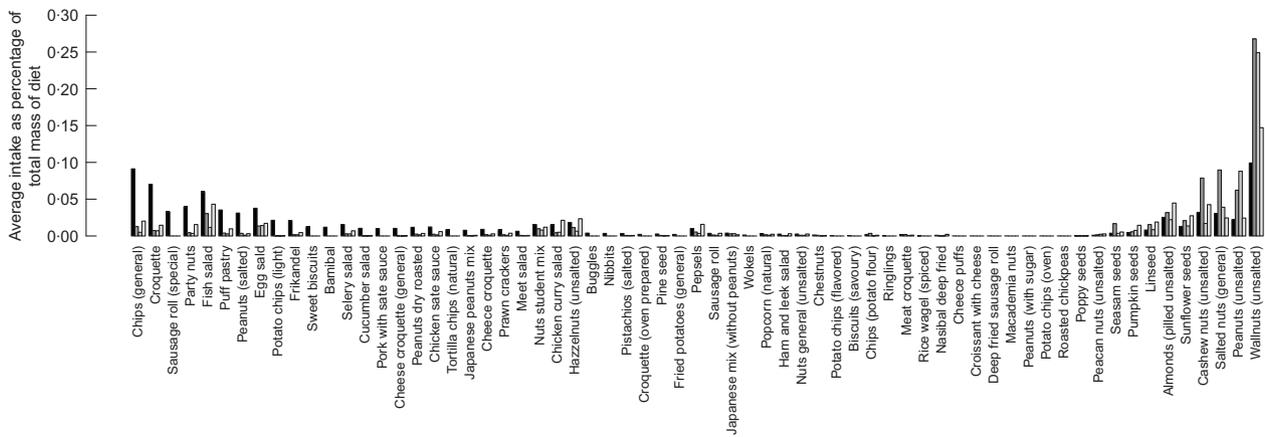


Fig. 6 Detailed average intake of food items that belong to the nuts, seeds and snack group for female consumer of the dataset. ■, Current diet; ■, diet calculated with input-oriented data envelopment analysis (DEA) model; □, diet calculated with output-oriented DEA model; □, diet calculated with MINDV model

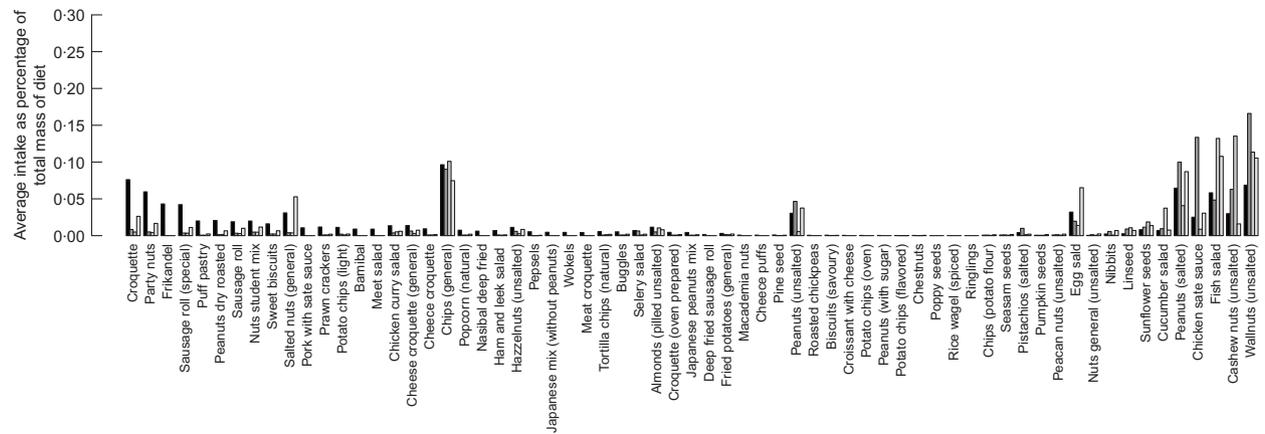


Fig. 7 Detailed average intake of food items that belong to the nuts, seeds and snack group for male consumer of the dataset. ■, Current diet; ■, diet calculated with input-oriented data envelopment analysis (DEA) model; □, diet calculated with output-oriented DEA model; □, diet calculated with MINDV model

The highest frequency score (calculated according to eq. 1) in all three models is achieved by the most frequent peer of the M1 group. The food group intakes of this diet are presented in Fig. 9. This diet comprises larger shares of fruits and vegetables and lower intakes of meat, dairy products and bread compared to the average diet of the same consumer group.

Discussion and concluding remarks

The proposed DEA-based algorithm enables the identification of existing healthier diets and provides a framework for quantification of potential improvement of the current diets. The advantage of the method, compared to existing diet models, is that the calculated diets are combinations of actual diets of peers. Because of this, the calculated diets are closer to the actual diets of the individuals that belong to the same group without the need to specify explicitly acceptability or other meal-related constraints.

Only as a next step, we demonstrated how the MINDV model can be used to identify for each current diet a healthier alternative diet that is as close as possible to the food item intakes of the current diet. In contrast to other diet optimisation models⁽⁷⁾ that also minimise deviations between the food item intakes of the calculated and the observed diet, the MINDV model combines whole DEA-efficient diets instead of individual food items. The main decision variable of the MINDV model is the weight of a DEA-efficient diet in the calculated diet. This ensures that the proportions of food item intakes of the calculated diet are observed in the diets of others in the same group. In existing diet optimisation models that also minimise deviations between the food item intakes of the calculated and observed diets, the decision variables are the intakes of available food items. This implies that the food item intakes can be chosen in any proportion. In most cases where the current diet does not meet specific nutrient requirements (or for any reason the current diet cannot be reproduced), then the quantities of food items in the calculated diet are

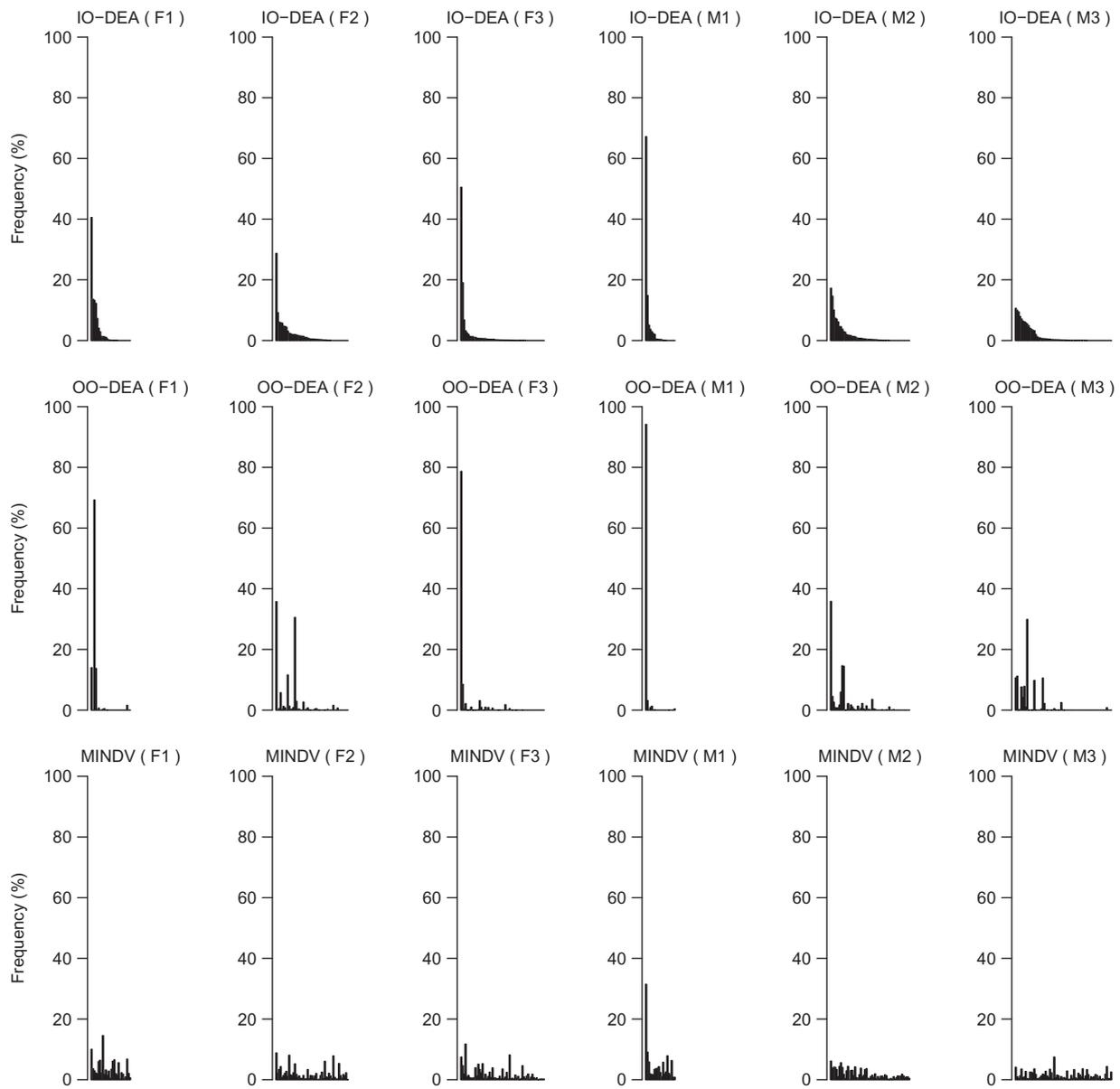


Fig. 8 Frequency of efficient diets (i.e. peers) as calculated by eq. 1 for different groups of individuals and model (i.e. input-oriented data envelopment analysis (IO DEA), output-oriented DEA and MINDV). To allow comparison between models, the peers have been sorted based on their frequency score in the IO DEA model

deviating from the current diet. In these cases, the calculated diet will not comprise necessarily of food items in proportions of actual diets unless acceptability constraints are added explicitly. The proposed DEA-based approach that we present in this manuscript extracts these constraints from the diets of (other) individuals that belong to the same group in an objective, transparent and reproducible way, which is a clear added value of the proposed method.

Similar to the combination of DEA-efficient diets calculated with the MINDV model, other combinations of the DEA-efficient diets can be used to identify healthier alternative diets that, instead of maximising acceptance, maximise environmental performance or minimise total cost. The resulting diets which maximise the environmental

performance or minimise the cost of the diet will be also combinations of DEA-efficient diets and not combinations of unrelated available food items. Consequently, they will reflect the current cultural food habits and preferences. Because of these, the proposed method is an advanced decision support tool that can be used to address questions related to the quest for designing sustainable, healthy affordable and reliable diets⁽⁹⁾.

As mentioned, the proposed diet modelling approach calculates diets as linear combinations of other observed diets. As a result, if the method is applied in datasets where unhealthy diets dominate, the calculated diets will be nutritionally improved but not necessarily nutritionally adequate (i.e. meeting all nutrient requirements). This was

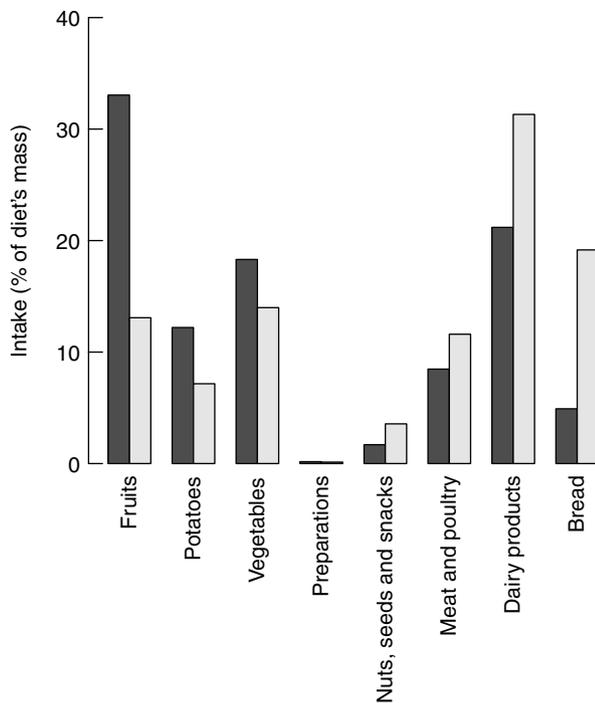


Fig. 9 Food group item intakes (percentage of total diet's weight) of the diet that is used most frequently as peer of inefficient diets (i.e. the most frequent peer of group M1). ■, Frequent peer; □, average of group

also shown by our case study for some nutrients (e.g. Fe) and groups of individuals. Of course, if nutritionally adequate diets are observed within the sample of benchmarked diets, these diets will become peers of other less healthy diets and the calculated diets will be also adequate from a nutritional point of view. Applying this method continuously over time can result gradually to diets that meet all dietary requirements. This is different than current diet optimisation models, where often nutritional adequacy of the calculated diets can be imposed by using hard constraints. With these models, a target diet is calculated, but it remains a question if such calculated nutritionally adequate diets will be acceptable by a population where such diets have not been chosen.

The proposed DEA-based approach can also be used to evaluate scenarios in future studies. Addressing what-if questions related to changes in prices or technologies or policies would be quite straightforward. New combinations of the benchmarked diets will be calculated for different scenarios. For example, a price change scenario of one or more food items can be evaluated by comparing the calculated diet for the current and future sets of prices. Assessing scenarios that involve new food items or diets will require to pre-design diets with different levels of the evaluated food items. The 'healthiness' of these hypothetical diets can then be evaluated using the proposed DEA-based approach.

The proposed method can exploit efficiently existing empirical datasets of the average nutrient intakes of individuals. It results in calculated diets that are linear combinations of other existing healthier diets. This can improve

the potential for acceptance. In the case study presented in this manuscript, we aim mainly to demonstrate the proposed method, and for that reason we chose age and gender as the main grouping variables of the population. It has been shown that socio-economic factors such as age and gender explain food intakes of populations across Europe⁽²²⁾. Moreover, such grouping resulted in adequate numbers of observations in all groups. Adding additional grouping variables like socio-economic status and physical activity will result in quite homogeneous groups. The calculated diets with the proposed DEA method will still include higher intakes of more-is-better nutrients and lower intakes of less-is-better nutrients than the current diets, and they will have a fair chance of adoption. However, the improvement of the calculated diets will be more modest (not enough variation and space to optimise) compared to the current application of the model. Moreover, larger datasets would be required to ensure enough observations per group. On the contrary, applying the DEA method to the whole population (i.e. no sub-grouping) will result in much larger improvements, but the improved diets will be the result of combining diets from different age/gender groups which might decrease the acceptance rate. Because of this trade-off between magnitude of improvement and homogeneity of the groups, the choice of grouping variables should be research question and case specific. Sensitivity analysis on the impact of grouping variables can provide interesting quantitative evidence about the importance of certain grouping variables to the potential improvements.

The nutrients and food items that should be included in the evaluation of the quality of a diet are an open debate in the field of health and nutrition. The focus of this manuscript is to propose a new concept for diet modelling based on benchmarking of existing diets instead of composing diets from available food items. To avoid entering the discussion on which nutrients must be included in such a model, we decided to use the nutrients included in Nutrient Rich Diet 9.3 (without additions or exclusions), which has been commonly used to evaluate quality of diets in the field of public health and nutrition⁽¹⁸⁾. The set of nutrients included in the DEA models allows to tailor the results of the analysis to the specific context of the application (i.e. type of disease, geographical region, etc.). However, the proposed method and the concepts introduced in this manuscript would remain the same. Adding additional nutrients in the DEA models would result in a larger set of efficient diets and consequently less potential improvement at group level. A selection of important nutrients (instead of using all available nutrients) is required to allow for substantial level of improvement. However, similar to the rationale of selecting specific nutrients in diet quality indices (e.g. food habits, available foods and type of disease), the justification of the nutrients that are used in DEA applications should also be provided.

The main dataset used in the current study includes information from FFQ and the Dutch food composition tables. We



used the Nutrition Questionnaires plus dataset⁽¹⁶⁾ which is specific for The Netherlands. However, similar datasets are available in many other European Union countries⁽²³⁾. We do not use any additional dataset or expert knowledge to define acceptability constraints. This enables applications of the method to larger geographical scales (like the European Union) which is an important advantage of the proposed method. We clustered the individual diets in six groups, and we applied the model within each of these groups. Even though the numbers of observations per group are higher than DEA applications in other fields of science⁽¹⁰⁾, a larger sample size per group would enable a better representation of the variability within the population. This would imply even more possibilities for improvement in the calculated diets. Moreover, additional recalls per individual would enable a better representation of current diets and would make possible the estimation of CI around the observed intakes. This would allow the application of extensions of the basic DEA models to account for uncertainty and improve the robustness of the analysis⁽²⁴⁾. Such information would also be beneficial for the majority of diet optimisation models which are currently by nature deterministic.

An important underlying assumption of DEA models is that data should be positive numbers. In the context of diet modelling, this will imply that diets with 0 intake of certain nutrients cannot be evaluated. This was not a problem in our dataset, but in case where some diets have zero intake of certain nutrient, data transformations can be applied^(25,26).

In the current study, we capped the intakes of more-is-better nutrients to their sufficient level to account for an upper bound on the comparisons between diets. By capping, we impose that increasing nutrient intakes above the sufficient levels (i.e. RDAs) is of less importance than increasing the intake of the nutrients that are not yet at their RDA levels. This approach is valid for cases where intake above a certain level is not necessarily better but it is also not harmful. In cases where higher intakes would be harmful, existing extensions of the basic DEA models should be used⁽²⁷⁾.

The outputs of the models provide specific guidelines for the appropriate food item intake that will result in a healthier alternative diet. At the same time, acceptance (or environmental or economic) considerations can be taken into account by identifying combinations of existing healthier diets that optimise specific indicators (e.g. deviation from the current diet). Such quantitative analysis can be used to evaluate ex-ante environmental and nutrition/public health-related policies at population or even higher (e.g. European Union) level.

Acknowledgements

Financial support: Financial support for this study was obtained from funding from the European Union's H2020 Programme under Grant Agreement number 633692 (SUStainable Food And Nutrition Security

(SUSFANS): Metrics, models and foresight for European sustainable food and nutrition security), from TiFN under Project Agreement number 15SD01 (SHARP-BASIC), from ZonMw under Grant Agreement number 91110030 (Nutrition Questionnaires plus study). *Conflict of interest:* There are no conflicts of interest. *Authorship:* A.K.: conception and design, develop model, analysis and interpretation of results, writing manuscript, final approval. J.C.G.: analysis and result interpretation, writing manuscript, final approval. A.I.: analysis and result interpretation, writing manuscript, final approval. J.M.G.: design research, nutritional verification of results, result interpretation, commenting/writing, final approval. J.M.B.-R.: design research, comment on model, result interpretation, commenting/writing, final approval. P.V.: design research, nutritional verification of results, result interpretation, commenting/writing, final approval. *Ethics of human subject participation:* None.

Supplementary material

For supplementary material accompanying this paper visit <https://doi.org/10.1017/S1368980019004774>.

References

1. Tilman D, Yang Y, Chen Y *et al.* (2014) Global diets link environmental sustainability and human health. *Nature* **515**, 518–522.
2. Green R, Milner J, Dangour AD *et al.* (2015) The potential to reduce greenhouse gas emissions in the UK through healthy and realistic dietary change. *Clim Change* **129**, 253–265.
3. Springmann M, Godfray HCJ, Rayner M *et al.* (2016) Analysis and valuation of the health and climate change co-benefits of dietary change. *Proc Natl Acad Sci* **113**, 4146–4151.
4. Buttriss JL, Briand A, Darmon N *et al.* (2014) Diet modelling: how it can inform the development of dietary recommendations and public health policy. *Nutr Bull* **39**, 115–125.
5. Gerdessen JM & de Vries JHM (2015) Diet models with linear goal programming: impact of achievement functions. *Eur J Clin Nutr* **69**, 1272–1278.
6. Ribal J, Fenollosa ML, Garca-Segovia P *et al.* (2016) Designing healthy, climate friendly and affordable school lunches. *Int J Life Cycle Assess* **21**, 631–645.
7. Gazan R, Brouzes CMC, Vieux F *et al.* (2018) Mathematical optimization to explore tomorrow's sustainable diets: a narrative review. *Adv Nutr* **9**, 602–616.
8. Maillot M, Vieux F, Amiot MJ *et al.* (2010) Individual diet modeling translates nutrient recommendations into realistic and individual specific food choices. *Am J Clin Nutr* **91**, 421–430.
9. Mertens E, van't Veer P, Hiddink GJ *et al.* (2017) Operationalising the health aspects of sustainable diets: a review. *Public Health Nutr* **20**, 739–757.
10. Cooper WW, Seiford LM & Kaoru T (2007) *Data Envelopment Analysis*. New York: Springer.
11. Zhou H, Yang Y, Chen Y *et al.* (2018) Data envelopment analysis application in sustainability: the origins, development and future directions. *Eur J Oper Res* **264**, 1–16.
12. Barr RS, Seiford LM & Siems TF (1993) An envelopment-analysis approach to measuring the managerial efficiency of banks. *Ann Oper Res* **45**, 1–19.



13. Liu JS, Lu LYY, Lu WM *et al.* (2013) A survey of DEA applications. *Omega* **41**, 893–902.
14. Thanassoulis E, Kortelainen M, Johnes G *et al.* (2011) Costs and efficiency of higher education institutions in England: a DEA analysis. *J Oper Res Soc* **62**, 1282–1297.
15. Chilingirian JA & David Sherman H (1996) Benchmarking physician practice patterns with DEA: a multi-stage approach for cost containment. *Ann Oper Res* **67**, 83–116.
16. Brouwer-Brolsma EM, van Lee L, Streppel MT *et al.* (2018) Nutrition Questionnaires plus (NQplus) study, a prospective study on dietary determinants and cardiometabolic health in Dutch adults. *BMJ Open* **8**, e020228.
17. Drewnowski A (2009) Defining nutrient density: development and validation of the nutrient rich foods index. *J Am Coll Nutr* **28**, 421S–426S.
18. Fulgoni VL, Keast DR & Drewnowski A (2009) Development and validation of the nutrient-rich foods index: a tool to measure nutritional quality of foods. *J Nutr* **139**, 1549–1554.
19. EFSA (2018) *Dietary Reference Values and Dietary Guidelines*. Parma, Italy: European Food Safety Authority (EFSA). <https://www.efsa.europa.eu/en/topics/topic/dietary-reference-values> (accessed April 2020).
20. Institute of Medicine (IOM) (2018) *Nutrient Recommendations: Dietary Reference Intakes (DRI)*. Washington, DC, USA: The National Academies Press. https://ods.od.nih.gov/Health_Information/Dietary_Reference_Intakes.aspx (accessed April 2020).
21. NEVO (2016) *Nutrient Recommendations: Dietary Reference Intakes (DRI)*. Bilthoven, The Netherlands: National Institute for Public Health and the Environment. <https://www.rivm.nl/en/dutch-food-composition-database/access-nevo-data/nevo-online> (accessed April 2020).
22. Mertens E, Kuijsten A, Dofkova M *et al.* (2019) Geographic and socioeconomic diversity of food and nutrient intakes: a comparison of four European countries. *Eur J Nutr* **58**, 1475–1493.
23. EFSA (2020) *The EFSA Comprehensive European Food Consumption Database*. Parma, Italy: European Food Safety Authority (EFSA). <https://www.efsa.europa.eu/en/food-consumption/comprehensive-database> (accessed April 2020).
24. Dyson RG & Shale EA (2010) Data envelopment analysis, operational research and uncertainty. *J Oper Res Soc* **61**, 25–34.
25. Sarkis J (2007) *Preparing your Data for DEA*. Boston, MA: Springer US.
26. Cook WD & Seiford LM (2009) Data envelopment analysis (DEA), thirty years on. *Eur J Oper Res* **192**, 1–17.
27. Chen Y, Cook WD, Du J *et al.* (2017) Bounded and discrete data and Likert scales in data envelopment analysis: application to regional energy efficiency in China. *Ann Oper Res* **255**, 347–366.