Initial model for designing SHARP diets

Deliverable No. D7.3

Argyris Kanellopoulos, Ante Ivancic, Jacqueline Bloemhof-Ruwaard (WUR ORL), Marianne Geleijnse, Anneleen Kuijsten (WUR HNE)

SUSFANS DELIVERABLES

Version	Release date	Changed	Status	Distribution
V1	28/02/2018	-	,,,	...

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 633692

SUSFANS Deliverable document information

Project name	SUSFANS
Project title:	Metrics, Models and Foresight for European SUStainable Food And Nutrition Security
Project no	633692
Start date:	April 2015
Report:	D7.3
Work package	WP7
WP title (acronym):	Modelling SHARP diets for EU-consumers
WP leader:	WUR, Marianne Geleijnse
Period, year:	2, 2018
Responsible Authors:	Argyris Kanellopoulos (WUR), Ante Ivancic (WUR), Marianne Geleijnse (WUR), Jacqueline Bloemhof-Ruwaard (WUR), Anneleen Kuijsten (WUR)
Participant acronyms:	WUR
Dissemination level:	Public
Version	V1
Release Date	30/03/2018
Planned delivery date:	30/03/2018
Status	Final
Distribution	Public

Dissemination level of this report

Public

ACKNOWLEDGMENT & DISCLAIMER

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 633692. Neither the European Commission nor any person acting on behalf of the Commission is responsible for how the following information is used. The views expressed in this publication are the sole responsibility of the author and do not necessarily reflect the views of the European Commission.

TABLE OF CONTENT

DELIVERABLE SHORT SUMMARY FOR USE IN MEDIA

Rising incomes and urbanization leads to global dietary changes. Traditional diets are replaced by diets higher in refined sugars, fats, oils and meats. Such dietary changes can increase the incidence of chronic diseases (like type II diabetes, coronary heart disease and other chronic non-communicable diseases). Indeed, high intakes of sodium and alcohol, and low intakes of fruit and vegetables, whole grains, and nuts and seeds ranked among the leading risk factors for early death and disability in European population in 2015. Replacing the current diets with healthier alternatives can have substantial public health and environmental benefits. Designing healthier diets can be a rather complex process. Diet models have been developed to deal with the complexity of designing such alternative, healthier or more sustainable diets. Their main objective is to determine the optimal quantities of available food items that should be included in a diet of an individual, so as to optimize specific criteria (e.g. minimize cost, increase diet healthiness etc.), while taking into account nutritional and acceptability constraints. Combining quantities of different food items to optimize a specific objective accounting only for nutritional constraints can result to optimal diets that are far from what people actually eat. The objective of this research is to present the basic SHARP model – the model that is able to generate sustainable, healthy, affordable, reliable and preferable (acceptable) diets, with the focus on demonstrating the underlying Data Envelopment Analysis (DEA) framework that can be used to benchmark current healthier diets and to provide guidelines for improving less healthy diets in a way that is acceptable by the studied population. Acceptability considerations are taken into account by identifying for each diet in our sample an alternative healthier diet which is as similar as possible, in terms of included food items, to the original current diet. Although not covering all SHARP dimension, the initial model's flexibility allows for additional dimensions to be included, such as sustainability indicators and prices, which will be included in consequent iterations thereof.

TEASER FOR SOCIAL MEDIA

This paper shows how the existing methods from the Operations Research domain can be applied within the context of diet modelling, notably Data Envelopment Analysis (DEA). Furthermore, it provides a tool for the design of healthier and more sustainable diets, while assuring their acceptability.

Benchmarking healthier and acceptable diets using Data Envelopment Analysis

A. Kanellopoulos[a,*], J. C. Gerdessen[a], A. Ivancic[a], J. M. Geleijnse[b], J.M. Bloemhof-Ruwaard[a]

[a]*Operations Research and Logistics group, Wageningen University, Hollandseweg 1, 6706 KN Wageningen, The Netherlands*
[b]*Human Nutrition group, Wageningen University, Stippeneng 4, 6708 WE Wageningen, The Netherlands*

Abstract

Designing healthier diets is a complex process which can have substantial public health benefits. The intakes, but also the requirements of multiple important nutrients for different population groups should be taken into account. Moreover, the current dietary preferences of individuals should be considered to promote the acceptability of the diet. Diet models have been developed and used for designing such healthier and acceptable diets. The main objective of these models is to determine the optimal quantities of available food items that should be included in a diet to optimize a specific indicator (e.g. maximize a dietary quality index). Additional constraints are defined to improve the acceptability of the calculated diets. These constraints are either in the form of upper and lower limits to the intake of specific food-items or in the form of fixed combinations of food-items in meals. Defining such constraints explicitly is challenging and involves expert knowledge and a substantial degree of arbitrariness. To avoid defining such acceptability constraints we propose a DEA based diet model that

*A. Kanellopoulos; tel.: +31 317 485 644; e-mail: argyris.kanellopoulos@wur.nl

Preprint submitted to European Journal of Clinical Nutrition *March 30, 2018*

benchmarks existing complete diets of a certain population and in our case identifies healthier alternatives. However, the model's flexibility allows for additional dimensions to be included, such as sustainability indicators and prices. The method was applied successfully to benchmark alternative diets of a group of individuals in the Netherlands.

Keywords: diet model, efficiency, nutrition, DEA, optimization

---

1. Introduction

Rising incomes and urbanization leads to global dietary changes. Traditional diets are replaced by diets higher in refined sugars, fats, oils and meats. Such dietary changes can increase the incidence of chronic diseases (like type II diabetes, coronary heart disease and other chronic non-communicable diseases) (Tilman et al., 2014). High intakes of sodium and alcohol, and low intakes of fruit and vegetables, whole grains, and nuts and seeds ranked among the leading risk factors for early death and disability in European populations in 2015 (Forouzanfar, 2015). Furthermore, in Europe there is a growing similarity of diets, in which traditional diets of Northern and Mediterranean countries are converging towards a more Western diet, viewed by the increased share of fruit and vegetables in Northern countries and the increased share of animal-based products in Mediterranean countries. Excess caloric intake has been thought as a key factor in nutrition transition, which warrants the need for public health action to promote healthier food patterns (Mertens et al., 2018). Replacing the current diets with healthier alternatives can have substantial public health and environmental benefits (Green et al., 2015; Springmann et al., 2016).

Designing healthier diets is a rather complex process. The intake of multiple nutrients and the differences of nutrient requirements between various groups of the population must be taken into account. Moreover, the dietary preferences of individuals should be also considered to design a healthier diet which is also acceptable by the general population.

Diet models have been developed to deal with the complexity of designing such alternative, healthier or more sustainable diets (Buttriss et al., 2014; Gerdessen and de Vries, 2015; Ribal et al., 2016). The main objective of current diet models is to determine the optimal quantities of available food items that should be included in a diet of an individual to optimize specific criteria (e.g. minimize the cost or CO_2 emissions or improve healthiness) taking into account nutritional and acceptability constraints (Buttriss et al., 2014; Gerdessen and de Vries, 2015; Ribal et al., 2016).

In these models, nutritional constraints are used to assess deviations of nutrient or food intakes of the calculated diet from specific dietary guidelines and recommendations of nutrient intakes [1]. Dietary quality indices are used to aggregate deviations between food or nutrient intakes of the calculated diet and the corresponding recommendations (Arvaniti and Panagiotakos, 2008; Alkerwi, 2014; Mertens et al., 2017)

Combining quantities of different food items to optimize a specific objective accounting only for nutritional constraints can result to optimal diets that are far from what people actually eat. Probably, such diets will not be accepted by individuals and consequently they can be highly irrelevant. For that reason, acceptability constraints are used in diet models to improve the acceptance of optimized diets. Acceptability constraints have either the form of upper and lower limits to the intake of specific food-items or the form

[1]Nutrient or food item recommendations of nutrient intakes and dietary guidelines are determined by experts for different groups of consumers

of fixed combinations of food-items in meals (Gerdessen and de Vries, 2015; van Dooren et al., 2015). Defining acceptability constraints can be a very challenging process that often involves expert knowledge and a substantial degree of subjectiveness and arbitrariness.

To avoid defining explicitly arbitrary acceptability constraints instead of trying to compose new diets from available food items we could try to look within existing diets of a certain population and identify the most healthy ones. These current healthier diets can then be 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 (by the models) diets that are healthier but also acceptable alternatives for the population under study.

A commonly used benchmarking technique in OR literature is Data Envelopment Analysis (Cooper et al., 2007). DEA has been used in many fields like banking, health care, agricultural economics, transportation and education (Liu et al., 2013; Zhou et al., 2018). In general, the aim of DEA is to identify Decision Making Units (DMUs) 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 can be seen as a DMU that can be evaluated based on the intake of multiple less-is-better (unhealthy) nutrients and more-is-better (healthy) nutrients.

The objective of this research is to present the basic SHARP model (SUSFANS, 2016), with the focus on demonstrating the underlying DEA-based framework that can be used to benchmark current healthier diets and to provide guidelines for improving less healthy diets in a way that is acceptable by the studied population. Acceptability considerations are taken into account by identifying for each diet in our sample an alternative healthier diet which is as similar as possible, in terms of included food items, to the original current diet. Although not covering all SHARP dimension, the initial model's flexibility allows for additional dimensions to be included, such as sustainability indicators and prices, which will be included in consequent iterations thereof.

First, we present the state of art in diet modelling and we discuss the benefits of benchmarking current diets instead of using existing diet models to compose new diets from available food-items. Second, we explain DEA and we illustrate how it can be applied to benchmark healthier diets based on a simplified example. Then we apply DEA to benchmark a set of healthier diets in the Netherlands using the NQplus database (van Lee et al., 2016).

2. 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 redesigning healthier diets. Finally, we use this simple example to propose a mathematical programming model that enables to account for acceptability considerations in diet modelling.

2.1. Diet models for designing healthier diets

Existing diet models are often Mathematical Programming (MP) models that aim to compose an alternative optimal diet for a specific individual or group of individuals (Mertens et al., 2017). A schematic representation of an MP model that focus on optimizing diets is presented in Figure 1a.

-Figure 1-

The main components of MP diet models are: (i) the decision variables ($q_i \quad \forall i \in (1 \ldots n)$) which are the quantities of available food items 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 optimized expressed as a function of the decision variables[2], (iii) the nutrition constraints that are used to calculate deviations between nutrient (and food) intakes of the calculated diet and the dietary guidelines, and (iv) the acceptability constraints which aim to improve acceptance of the diet by imposing restrictions on food-item quantities following current consumption patterns. Without acceptability constraints the calculated diet comprise of food items that do not necessarily form a preferable diet.

Defining acceptability constraints is often based on expert knowledge and information on current meals. 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 defining explicitly acceptability constraints we propose a benchmarking method that focuses on creating an alternative, healthier than the current, diet by combining benchmarked current diets of other individuals. Our proposed method is summarized in (Figure 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 have been chosen together in meals which form the diet of other individuals. The benchmarked diets that

[2]for example the objective function could be the maximization of the value of a certain dietary quality index or the minimization of CO_2 emissions

are used to compose all diets of the population under study are identified based on nutritional benchmarking and DEA (Cooper et al., 2007).

2.2. DEA with an illustrative example

This section illustrates the basic concepts of DEA based on a simple two-dimensional example. DEA aims to compare Decision Making Units (DMUs) based on their capacity to convert multiple inputs into multiple outputs (Cooper et al., 2007). In the specific case of a diet problem the DMUs are individual diets, outputs are the more-is-better nutrients like vitamins, fibre, and protein while inputs are the less-is-better nutrients like saturated fat and sodium. 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 i.e. dietary fiber (as the more-is-better nutrient) and sodium (as the less-is-better nutrient). Figure 2 involves 6 individual diets with the same caloric intakes labelled with letters (A, B, C, D, E and F) and their performance with respect to dietary fiber and sodium content (e.g. diet C contains 60 units of dietary fiber and 20 units of sodium).

-Figure 2-

DEA aims to compare each diet with all other diets in the sample and identify those that are efficient i.e. 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 Figure 2 it is visible that diets B,C and D are DEA efficient because for their intake of sodium there are no linear combination of all other diets with higher or the same intake of protein. The line segments BC and CD are called the DEA efficient frontier. The DEA model assigns to 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 i.e. output-oriented DEA model (OO-DEA) or decrease the amount of less-is-better nutrients for the same level of more-is-better nutrients i.e. input-oriented DEA model (IO-DEA). All inefficient diets (i.e. diets A, E, and F) receive an efficiency score of less than one. 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 Figure 2, diet A is not efficient because it can be replaced by diet B which has the same sodium content but higher dietary fiber content. Diet E is inefficient because it can be replaced by diet D which has a lower sodium content for the same dietary fiber content. Finally, diet F is inefficient because by combining diet B and C we create a new diet on the efficient frontier i.e. diet F^i that has the same dietary fiber level for substantially less sodium (input oriented 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 fiber for the same level of fat (output oriented DEA model).

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 the efficient diet we solve a sequence of Linear Programming (LP) models IO-

DEA and OO-DEA of Appendix A.

2.3. "Acceptability" considerations

The IO-DEA and the OO-DEA models identify a subset of diets that are efficient. For each of the inefficient diet a new diet is calculated as a linear combination of its peers. The new "efficient" and healthier alternative is identified by minimizing the distance to the frontier. For example, for diet F a new diet F^i is calculated that is healthier (less sodium for the same dietary fiber level). However, the new healthier diet (F^i) might include food items completely different than the current diet.

Actually, all linear combinations of the efficient diets within the shaded area F^iFF^oC are healthier alternatives of the current diet F because they contain less sodium and more dietary fiber than F. One of these alternative diets is the most similar to the current diet in terms of quantities of food-items. Such alternative healthier diet will be accepted easier by the specific individual which will have to make less changes in her current dietary choices.

For this reason we propose that each inefficient diet is replaced by the healthier combination of efficient frontier that minimizes the deviation between the food-item (or food-group) intakes of the healthier alternative and the current diet. To do this in a multi-nutrient, multi-food-item context we use model MINDV of Appendix B.

3. Case study: designing healthy diets in the Netherlands

To demonstrate how DEA can be used to benchmark healthier and acceptable diets we used information from the NQplus dataset for the Netherlands (van Lee et al., 2016). The dataset comprises of the average nutrient and food-item intake of 1735 individuals (men and women from 20 to 76 years of age). The data was collected through a food frequency questionnaire.

To make diets comparable we defined 6 groups of individuals: (F1) females 20-40 years of age, (F2) females 41-50 years of age, (F3) females > 50 years of age, (M1) males 20-40 years of age, (M2) males 41-50 years of age, (M3) males > 50 years of age. The number of observations per consumer group are presented in Figure 3.

-Figure 3-

It should be noted that defining groups on key demographic features may or may not correspond to segmentation in consumer markets by choices in the market place (revealed preferences) or consumer motives (stated preferences). Alternative ways of grouping consumers will be explored in further research.

The nutrient and food intakes of the diets are standardized to a diet of 2000 kcal. For example to normalize the nutrient and food intakes of a diet of 3000 kcal we multiplied average nutrient and food intakes of this diet with a factor of $\frac{2}{3}$.

We assumed that the nutrients used for calculating the NRD 9.3 (Drewnowski, 2009) are the most important and sufficient nutrients for evaluating the healthiness of a diet[3]. We accounted for 9 more-is-better nutrients (i.e. protein, fibre, calcium, iron, magnesium, potassium and vitamins A, D and E) and 3 less-is-better nutrients (i.e. sodium, saturated fat and added sugars). For all the more-is-better nutrients (except for fiber) we assume that more

[3]The approach would remain the same for any other set of nutrients

is beneficial only below a certain maximum level. Intakes higher than these maximum levels were not considered to be beneficial. To account for this while benchmarking diets, we used a cut-off value for each of the more-is better nutrient (except from fiber). If a certain nutrient intake of a diet exceeds the cut-off value we set the nutrient intake of this diet equal to the cut-off value. The cut-off value of each nutrient was set equal to the intake that is expected to be adequate to meet the requirements of at least 97.5% of all relevant subgroups (excluding pregnant or lactating women) in western populations[4]. The distributions of nutrient intake per consumer group and the cut-off values of selected nutrients are presented in Figure 4.

-Figure 4-

DEA was used to benchmark diets for each of the 6 consumer group i.e. we only compared diets that belong to the same consumer group. We used both the input-oriented (IO-DEA) and the output-oriented (OO-DEA) model. The MINDV model was used to combine DEA-efficient (benchmarked) 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 NEVO classification) intakes of the current diet.

The IO-DEA and the OO-DEA were used to benchmark the efficient diets and rank current diets based on an efficiency score. The IO-DEA model focus on calculating alternative diets that minimize less-is-better nutrients for the same or even lower levels of more-is-better nutrients. The OO-DEA model focuses on calculating alternative diets that maximize more-is-better

[4]This is based on combined information from EFSA (2018) and IOM (2018)

nutrients for the same or even lower levels of more-is-better nutrients. The MINDV focuses on combining benchmarked diets (identified by the IO-DEA and the OO-DEA DEA models) and design alternative diets that are as close as possible with respect of food-group-item intakes to the current diets. To identify interesting benchmarked diets we calculated the frequency that each benchmarked diet is used as peer using the following equation (1):

$$FRQ_k = 100 \; \frac{\sum\limits_{l|l \neq k} \lambda^*_{l,k}}{\sum\limits_{l,q} \lambda^*_{l,q}} \quad \% \tag{1}$$

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.

4. Results

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 Figure 5.

-Figure 5-

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 6 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 use of cut-off values during benchmarking with DEA. The calculated alternative diets have an intake level of vitamin A higher than the cut-off vale but lower than the original intake of the current diet.

As expected the MINDV model results in the smallest improvements since it focuses on minimizing deviation between food group items of the current and calculated diets instead of optimizing the level of nutrients in the calculated 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).

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 (Figure 6). Benchmarked diets with higher frequency scores (calculated according to eq. 1) can provide guidelines for defining nutritional and policy goals.

-Figure 6-

Benchmarked 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 gr) of the observed diets, and the diets calculated with the IO-DEA, the OO-DEA and the MINDV models are presented in Figure 7.

-Figure 7-

In general, compared to the current diets, the healthier diets calculated with the three models contain higher intakes of fruits and vegetables. Healthier diets also contain lower intakes of bread; nuts and snacks; fats and oils; soups, pastry and cakes; and sugar/sweets. Decrease of intakes of bread and the nuts, seed and snacks groups is not in line with existing dietary guidelines that recommentd 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. Figure 8 and Figure 9 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.

Calculated Healthier diets of younger males contains higher intakes of potatoes, milk and dairy. While healthier diets of females include lower intakes of potatoes and higher intakes of eggs.

As expected the most similar to the current diets in terms of food-group-item intakes are the diets calculated with the MINDV model.

The calculated diets can be detailed to food item level intake. For example

-Figure 8- -Figure 9-

5. 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 alternative diets are composed by combining existing diets. 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.

We demonstrated how acceptability concerns can be taken into account. For each of the current diets we identified a healthier alternative which is also as close as possible to average food-item intakes of the current diet. Similarly, other combinations of the benchmarked diets can be used to identify healthier alternative diets that instead of maximizing acceptance they maximize environmental performance or minimize total cost.

The DEA-based models exploit efficiently existing empirical data sets which improve representation of the current actual dietary choices of the studied population. However, the model 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 straight forward. 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.

An important underlying assumption of DEA models is that data should be non-negative (which means also 0's are not allowed). 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 (which is not very probable) data transformations can be applied (Sarkis, 2007; Cook and Seiford, 2009).

Another implicit assumption is that the diets compared are representative of the diet of the population but also of the average daily intakes of the individual.

Furthermore, we implicitly assume that linear combinations of diets will result in new diets that are transferable to other individuals in the population that likely have not been observed consuming neither the original, nor the new diets. Although an assumption, we still claim that it is more likely to end up with an acceptable diet if one takes linear combinations of current, already existing diets within a specific subpopulation, than generating a completely new diet that can consist of any possible subset of food items present in the food item data set. Indeed, such an approach is possible, however highly unlikely to result in a diet that will be general enough to adhere to all constraints that reflect preferences of different subpopulations. Stated differently, these preferences would have to be defined for all subpopulations, which might be highly impractical.

In this paper we define subpopulations (groups) on the basis of key demographic features (age, gender) which, as stated before, may or may not correspond to e.g. segmentation in consumer markets based on their revealed preferences, or any other segmentation that would accurately model similarities between consumers. Alternative ways of grouping consumers will be explored in further research, notably the ones that take a "bottom-up" approach, in the sense of building consumer relations based on their purchasing behavior, which would result in a more flexible (probabilistic) groupings, that would serve as a basis for more accurate predictions of their preferences.

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 combination of existing healthier diets that optimize 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. EU) level.

Acknowledgement

This research has received funding from the European Union Horizon 2020 research and innovation program SUSFANS (EU grant agreement no. 633692) and the Top Institute of Food and Nutrition project SHARP-BASIC.

Bibliography

Alkerwi, A., 2014. Diet quality concept. Nutrition 30, 613 – 618.

Arvaniti, F., Panagiotakos, D. B., 2008. Healthy indexes in public health practice and research: A review. Critical Reviews in Food Science and Nutrition 48, 317–327.

Buttriss, J. L., Briend, A., Darmon, N., Ferguson, E. L., Maillot, M., Lluch, A., 2014. Diet modelling: How it can inform the development of dietary recommendations and public health policy. Nutrition Bulletin 39, 115–125.

Cook, W. D., Seiford, L. M., 2009. Data envelopment analysis (dea) thirty years on. European Journal of Operational Research 192, 1–17.

Cooper, W. W., Seiford, L. M., Kaoru, T., 2007. Data Envelopment Analysis. Springer.

Drewnowski, A., 2009. Defining nutrient density: Development and validation of the nutrient rich foods index. Journal of the American College of Nutrition 28, 421S–426S.

EFSA, 2018. Dietary reference values and dietary guidelines. URL `https://www.efsa.europa.eu/en/topics/topic/dietary-reference-values-and-dietary-guidelines`

Forouzanfar, M. H., 2015. Global, regional, and national comparative risk assessment of 79 behavioural, environmental and occupational, and metabolic risks or clusters of risks, 1990-2015: a systematic analysis for the global burden of disease study 2015. The Lancet.

Gerdessen, J., de Vries, J., 2015. Diet models with linear goal programming: impact of achievement functions. European Journal of Clinical Nutrition 69, 12721278.

Green, R., Milner, J., Dangour, A. D., Haines, A., Chalabi, Z., Markandya, A., Spadaro, J., Wilkinson, P., 2015. The potential to reduce greenhouse gas emissions in the uk through healthy and realistic dietary change. Climatic Change 129, 253–265.

IOM, 2018. Nutrient recommendations: Dietary reference intakes (dri). URL http://web.archive.org/web/20080207010024/http://www.808multimedia.com/winnt/kernel.htm

Liu, J. S., Lu, L. Y., Lu, W.-M., Lin, B. J., 2013. A survey of dea applications. Omega 41 (5), 893 – 902.

Mertens, E., Kuijsten, A., Dofková, M., Mistura, L., D'Addezio, L., Turrini, A., Dubuisson, C., Favret, S., Havard, S., Trolle, E., van't Veer, P., Geleijnse, J. M., Mar 2018. Geographic and socioeconomic diversity of food and nutrient intakes: a comparison of four european countries. European Journal of Nutrition.

Mertens, E., van't Veer, P., Hiddink, G. J., Steijns, J. M., Kuijsten, A., 2017. Operationalising the health aspects of sustainable diets: a review. Public Health Nutrition 20, 739–757.

Ribal, J., Fenollosa, M. L., García-Segovia, P., Clemente, G., Escobar, N., Sanjuán, N., 2016. Designing healthy, climate friendly and affordable school lunches. The International Journal of Life Cycle Assessment 21, 631–645.

Sarkis, J., 2007. Preparing Your Data for DEA. Springer US, Boston, MA, pp. 305–320.

Springmann, M., Godfray, H. C. J., Rayner, M., Scarborough, P., 2016. Analysis and valuation of the health and climate change cobenefits of dietary change. Proceedings of the National Academy of Sciences.

SUSFANS, 2016. Wp 7: Modelling sharp diets for eu-consumers.
URL `http://www.susfans.org/wp-7-modelling-sharp-diets-eu-consumers`

Tilman, D., Yang, Y., Chen, Y., Zhu, J., 2014. Global diets link environmental sustainability and human health. Nature 515, 518 – 522.

van Dooren, C., Tyszler, M., Kramer, G. F. H., Aiking, H., Sep. 2015. Combining Low Price, Low Climate Impact and High Nutritional Value in One Shopping Basket through Diet Optimization by Linear Programming. Sustainability 7 (9), 12837–12855.

van Lee, L., Geelen, A., Hooft van Huysduynen, E. J. C., de Vries, J. H. M., van 't Veer, P., Feskens, E. J. M., 2016. Associations between company at dinner and daily diet quality in dutch men and women from the nqplus study. European Journal Of Clinical Nutrition 70, 1368.

Zhou, H., Yang, Y., Chen, Y., Zhu, J., 2018. Data envelopment analysis application in sustainability: The origins, development and future directions. European Journal of Operational Research 264, 1 – 16.

Figures and Tables

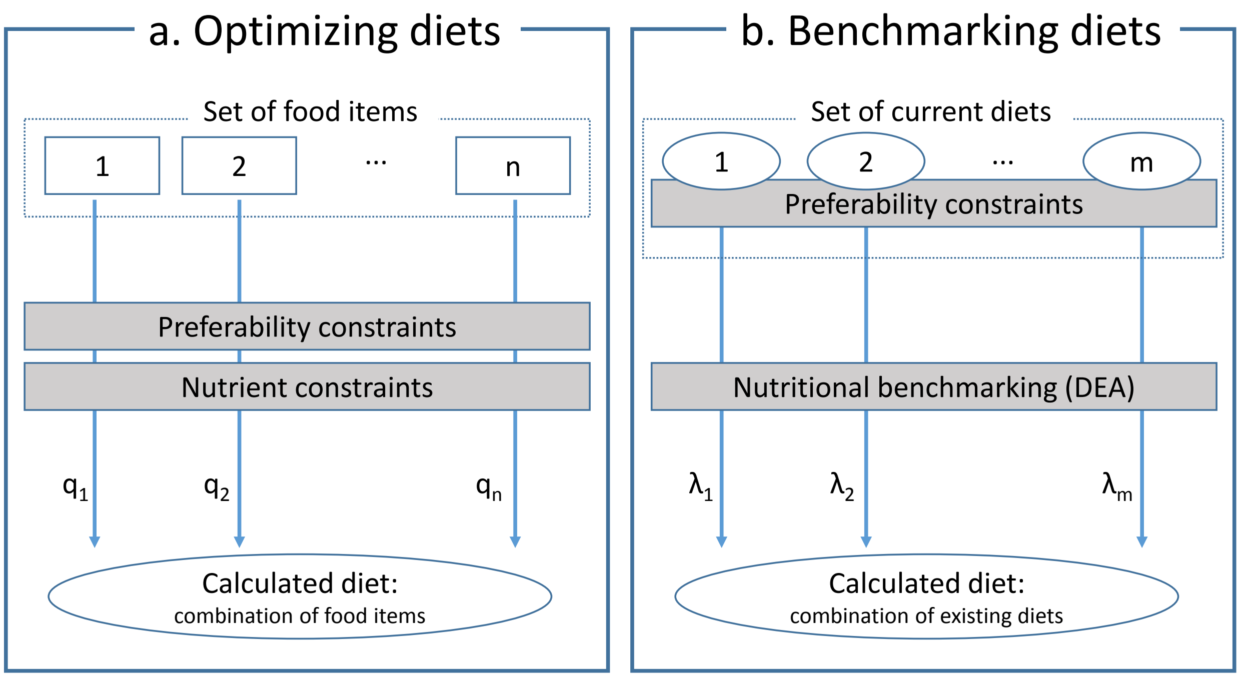

Figure 1: Schematic representation of the differences between current diet models, that focus on optimizing diets by deciding on optimal intakes of available food items, and the proposed benchmarking approach, that focuses on identifying efficient current diets and combine them to healthier alternatives.

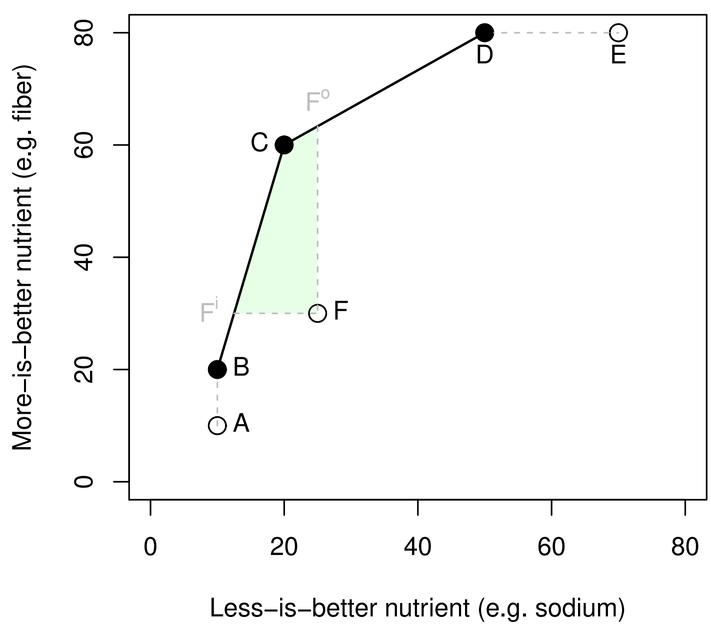

Figure 2: A two dimensional illustrative example of Data Envelopment Analysis (DEA) for benchmarking diets.

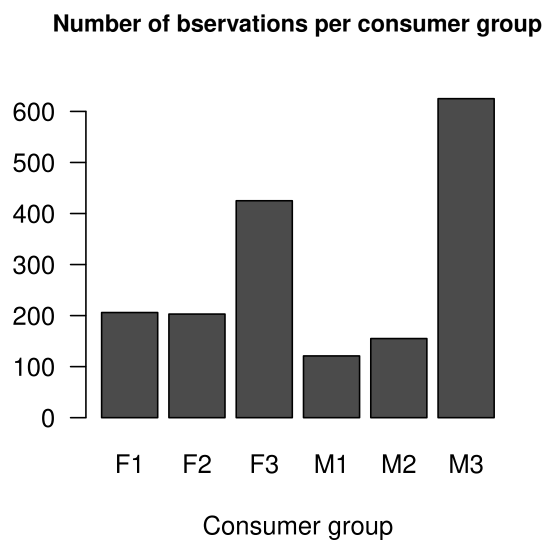

Figure 3: Number of observations per consumer group: F1(20-40 yrs), F2(41-50 yrs), F3($>$ 51 yrs), M1(20-40 yrs), M2(41-50 yrs), M3($>$ 51 yrs).

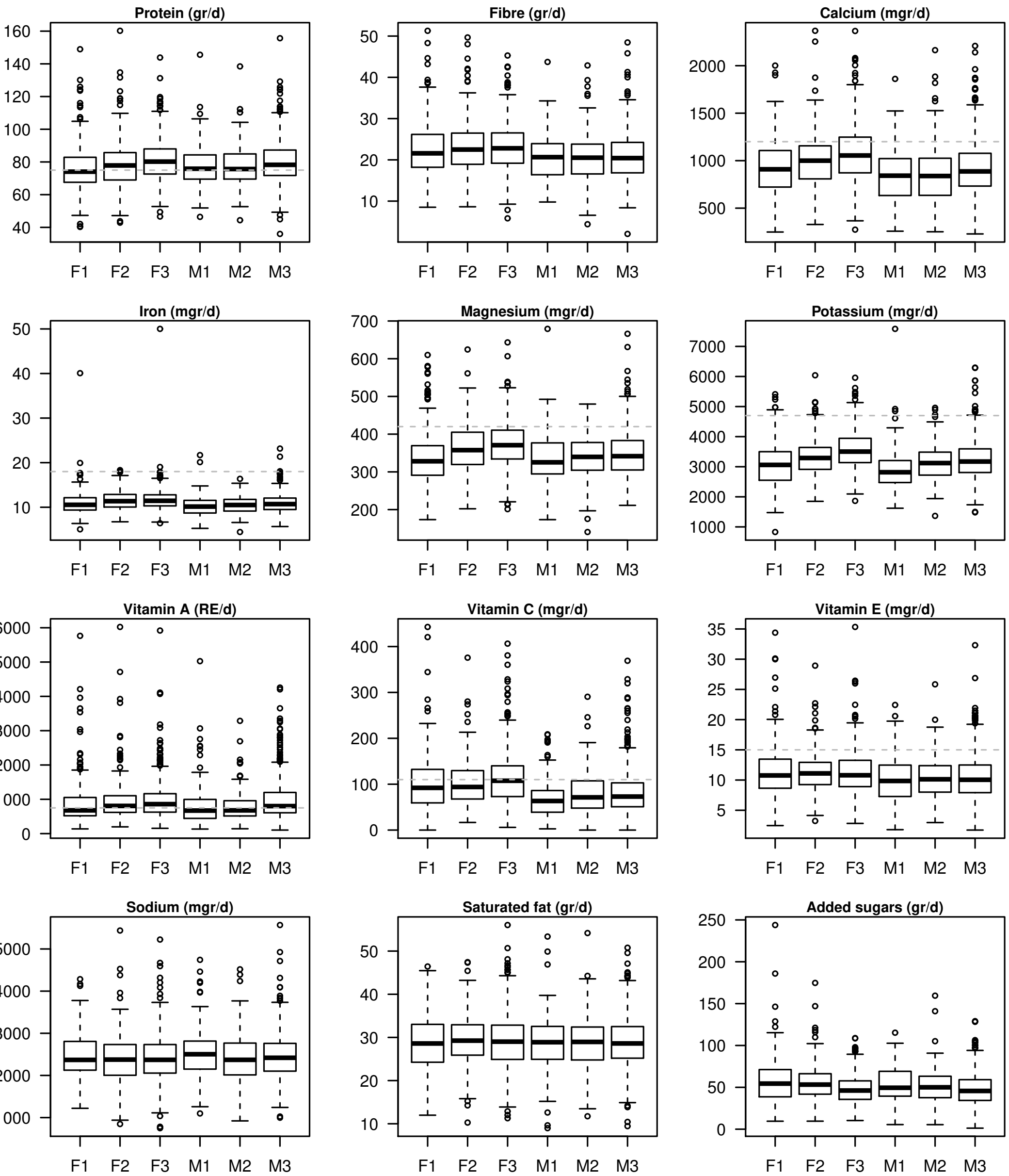

Figure 4: Distribution of Nutrient intakes of 9 more-is-better nutrients (protein, fibre, calcium, iron, magnessium, potassium,vitamine A, vitamine C and vitamin E) and 3 less-is-better nutrients (sodium, saturated fat, total sugars) per consumer group. The cut-off value of more-is-better nutrients (except from fibre) are represented as a horizontal dotted line.

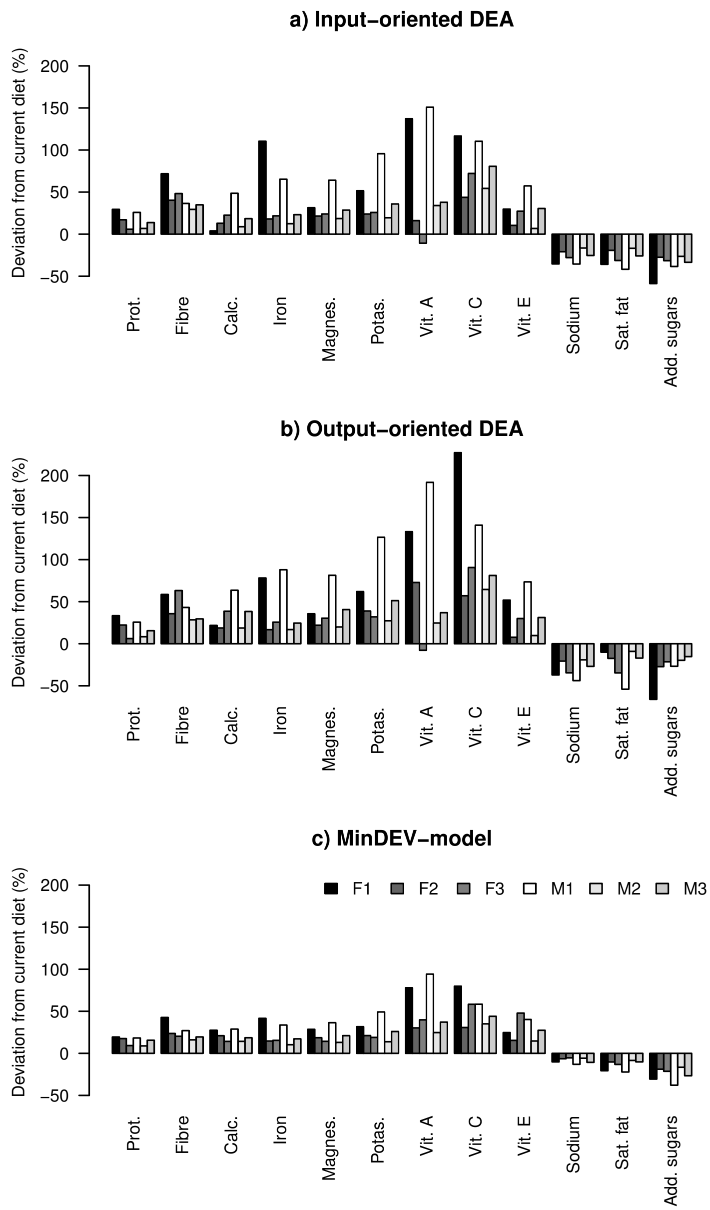

Figure 5: Difference between efficient and observed nutrient intakes as calculated with the (a) IO-DEA , (b) OO-DEA and the (c) MINDV models.

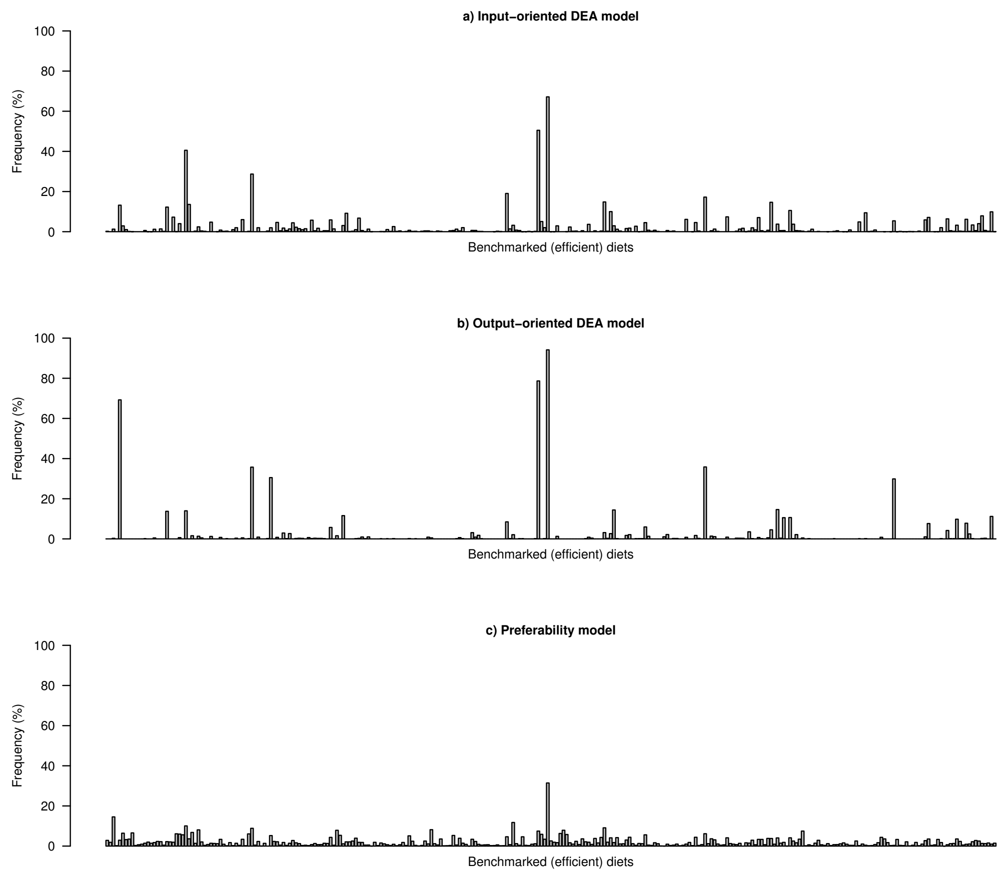

Figure 6: The frequency of an efficient diet as a peer of the inefficient diets calculated with the IO-DEA, the OO-DEA and the MINDV models.

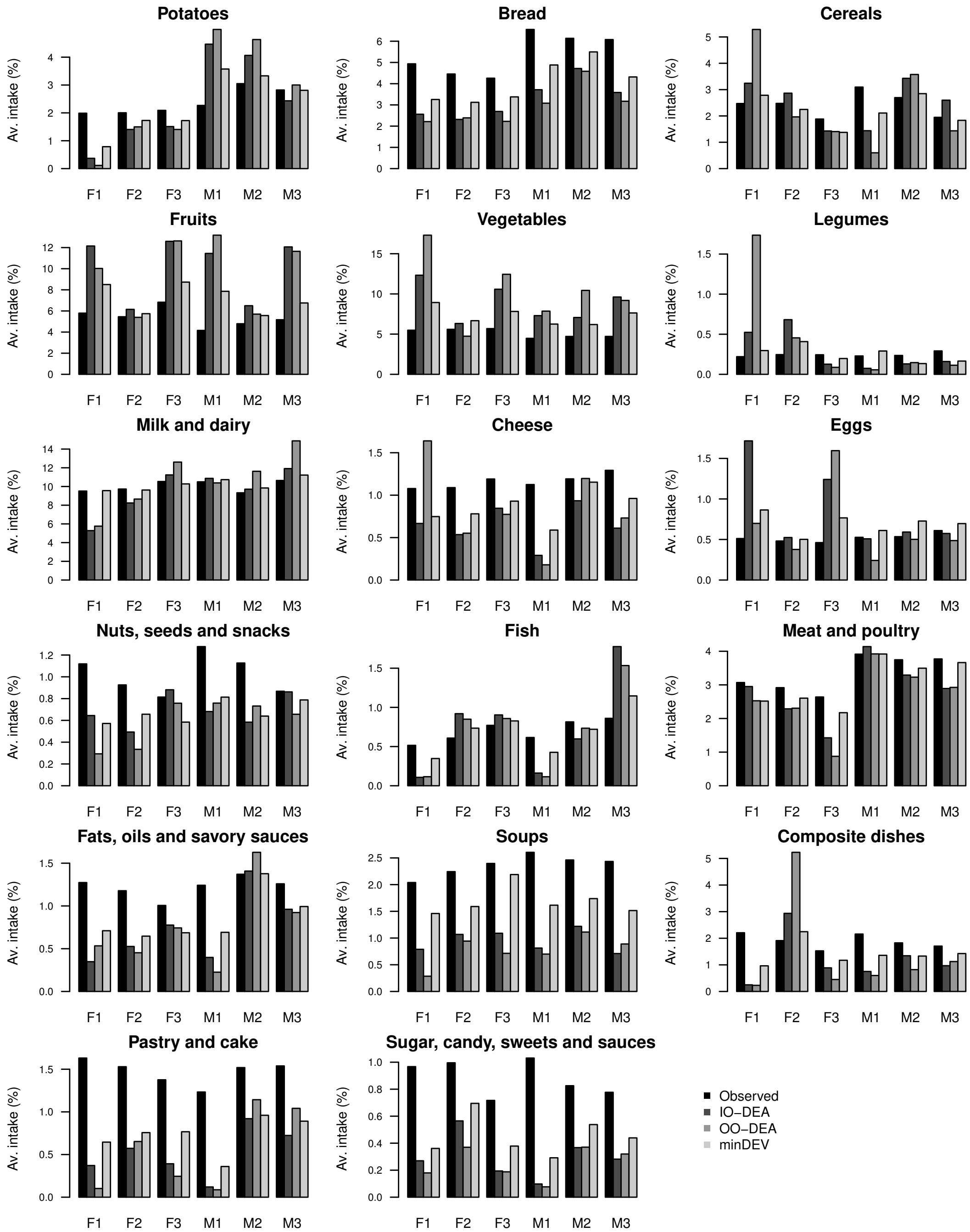

Figure 7: Average daily intakes of important food groups (expressed in % of total weight of the diet) in the current and the calculated diets with the IO-DEA, the OO-DEA and the MINDV models.

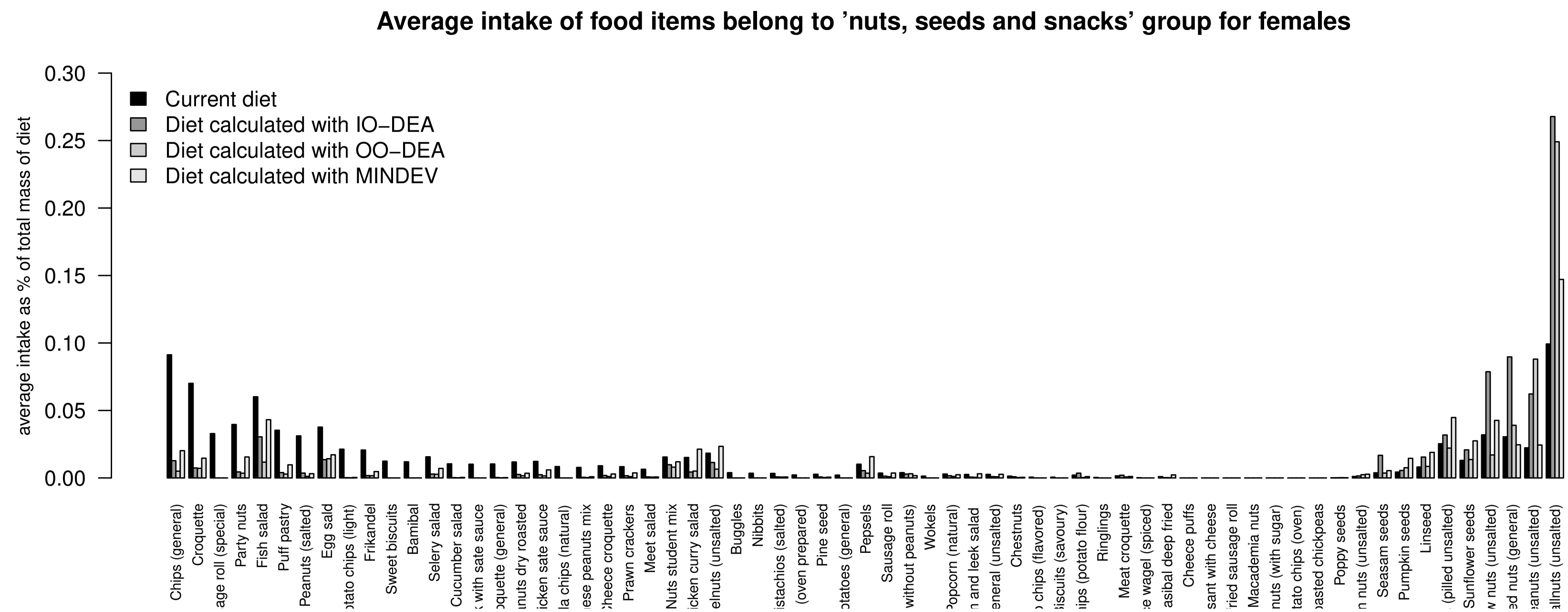

Figure 8: Detailed average intake of food items that belong to the nuts, seeds and snack group for female consumer of the dataset.

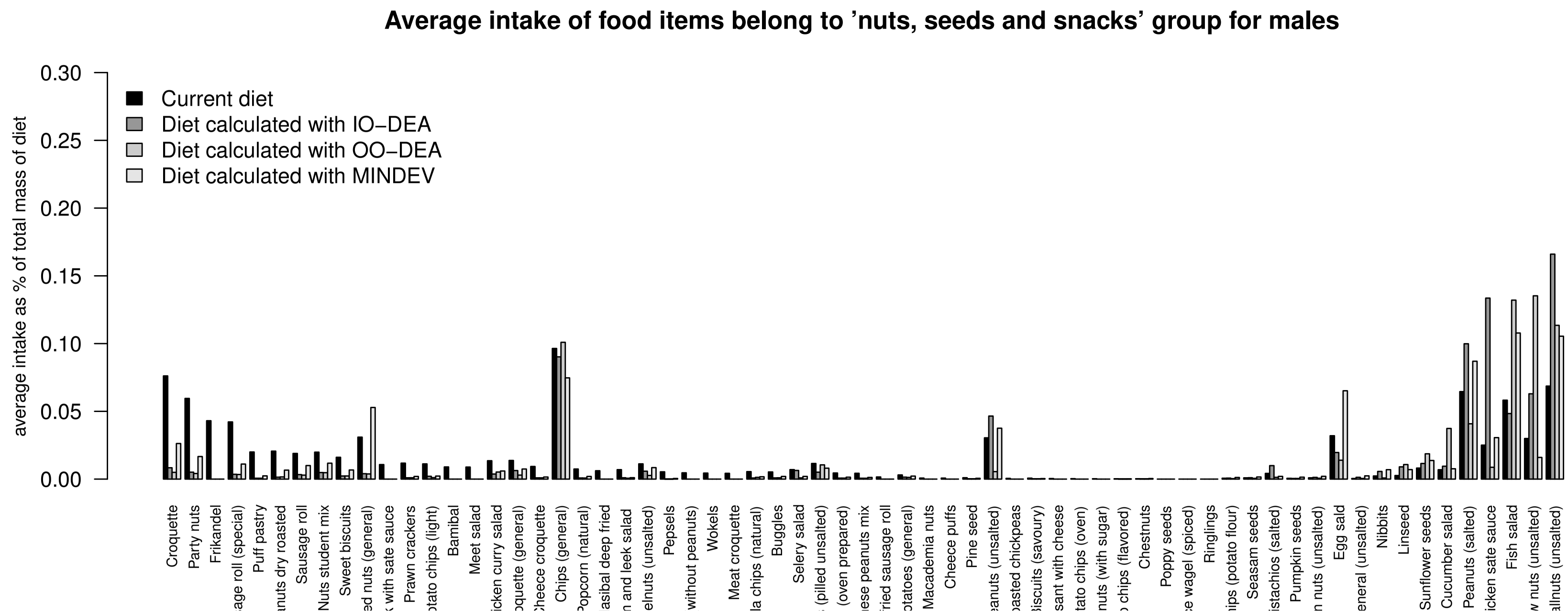

Figure 9: Detailed average intake of food items that belong to the nuts, seeds and snack group for male consumer of the dataset.

Appendix A. Description of basic DEA models

In this paper we use both the input-oriented and the output oriented CCR models (Cooper et al., 2007). The models are used multiple times (i.e. as many times as the number of diets in the sample) and they compare each diet with all other diets in the sample. The mathematical formulation of the input-oriented DEA model (IO-DEA) model is presented below:

$$\begin{aligned}
min \quad & \{\theta - \epsilon(\sum_i s_i^{in} + \sum_j s_j^{ot})\} \\
\text{s.t.} \quad & \sum_k y_{jk}\lambda_k - \quad y_j^0 - s_j^{ot} = 0, \quad \forall j \\
& \sum_k x_{ik}\lambda_k - \theta x_i^0 + s_i^{in} = 0, \quad \forall i \qquad \text{(IO-DEA)} \\
& \sum_k \lambda_k = 1 \\
& \lambda_k, s_j^{ot}, s_i^{in} \geq 0, \ \forall i, j, k,
\end{aligned}$$

Where $\theta \in \{0 \dots 1\}$ is the efficiency score of the evaluated diet calculated with the input-oriented DEA model (efficient diets get the value of 1), λ_k is a decision variable and the weight of diet k in the efficient alternative of the evaluated diet, s_j^{ot} is a slack decision variable for output j, s_i^{in} is a slack decision variable for input i, ϵ is a marginal (i.e. very small) positive number, x_{ik} is the content of less-is-better nutrient i in diet k, y_{jk} is the content of more-is-better nutrient j in diet k, x_i^0 is the content of less-is-better nutrient i in the evaluated diet, and y_j^0 is the content of more-is-better nutrient j in the evaluated diet.

The mathematical formulation of the output-oriented DEA model (OO-DEA) model is presented below:

$$
\begin{aligned}
max \quad & \{\phi + \epsilon(\sum_i s_i^{in} + \sum_j s_j^{ot})\} \\
\text{s.t.} \quad & \sum_k y_{jk}\lambda_k - \phi y_j^0 - s_j^{ot} = 0, \quad \forall j \\
& \sum_k x_{ik}\lambda_k - \ x_i^0 + s_i^{in} = 0, \quad \forall i \\
& \sum_k \lambda_k = 1 \\
& \lambda_k, s_j^{ot}, s_i^{in} \geq 0, \ \forall i, j, k,
\end{aligned}
\tag{OO-DEA}
$$

Where ϕ is the efficiency score of the evaluated diet calculated with the output-oriented DEA model. The higher the value of phi the higher the efficiency of the evaluated diet. To normalize the efficiency scores of the OO-DEA to values from 0 (i.e. lowest efficiency) to 1 (i.e. highest efficiency) we report efficiency scores as ϕ^{-1}.

For computational efficiency and to avoid determining an appropriate marginal positive parameter ϵ both the IO-DEA and the OO-DEA models were solved in two stages following Cooper et al. (2007).

Appendix B. Preferability extentions

To identify healthier diets that are as close as possible to current food-item intakes we use model MINDV. This model minimizes total absolute deviation between the food item intake between the new calculated diet (a point of the efficient frontier) and the food-item intakes of the current diet. The model makes sure that the new diet contains at least the intake of more-is-better nutrients of the current diet and at most the intake of less-is-better nutrients of the current diet. Similar to the basic IO-DEA and OO-DEA model this model is also used multiple times. Diets that have been identified efficient based on the basic DEA models do not have to be evaluated (DEA-efficient diets are benchmarks and they do not have to change the current food-item intakes.)

$$
\begin{aligned}
min \quad & \{\sum_i (d_f^+ + d_f^-)\} \\
\text{s.t.} \quad & \sum_k FI_{f,k}\lambda_k + d_f^- - d_f^+ = FC_f^0, \quad \forall f \\
& \sum_k y_{jk}\lambda_k - y_j^0 \geq 0, \qquad \forall j \\
& \sum_k x_{ik}\lambda_k - x_i^0 \leq 0, \qquad \forall i \\
& \sum_k \lambda_k = 1 \\
& \lambda_k, d_f^+, d_f^- \geq 0, \ \forall k, f,
\end{aligned}
\tag{MINDV}
$$

Where d_f^+ and d_f^- are the positive and negative deviation of the calculated food-item intake from the food-item intake of the current diet, $FI_{f,k}$ is the intake of food-item f of diet k, FC_f^0 is the intake of food-item f of the evaluated diet.