USING N-ALKANES TO ESTIMATE DIET COMPOSITION OF HERBIVORES: A NOVEL MATHEMATICAL APPROACH

PAULO BARCIA, MIGUEL N. BUGALHO, MANUEL L. CAMPAGNOLO,
J. ORESTES CERDEIRA

Abstract. N-alkanes are long chain saturated hydrocarbons occurring in plant cuticles that can be used as chemical markers for estimating the diet composition of herbivores. An important constraint of using N-alkanes to estimate diet composition with currently employed mathematical procedures is that the number of markers must be equal or larger than the number of diet components. This is a considerable limitation when dealing with free-ranging herbivores feeding on complex plant communities.

We present a novel approach for the estimation of diet composition using N-alkanes which applies equally to cases where the number of markers is lower, equal or greater than the number of plant species in the diet. The model uses linear programming to estimate the minimum and maximum proportions of each plant in the diet, and avoids the need for grouping species in order to reduce the number of estimated dietary components.

We illustrate the model with two data sets of N-alkane content of plants and faeces obtained from a sheep grazing experiment conducted in Australia and a red deer study in Portugal. Our results are consistent with previous studies on those data sets and provide additional information on the proportions of individual species in the diet. Results show that sheep included in the diet high proportions of white clover (from 0.25 to 0.37), and relatively high proportions of grasses (e.g. brome from 0.14 to 0.26), but tended to avoid Lotus spp (always less than 0.04 of the diet). For red deer we found high proportions of legumes (e.g. Trifolium angustifolium and Vicia sativa reaching maximum proportions of 0.42 and 0.30 of the diet, respectively) with grasses being less important and Cistus ladanifer, a browse, also having relevance (from 0.21 to 0.42 of the diet).

Key words and phrases. Diet composition, plant markers, ungulates, linear equations, linear programming, convex cones.

This work was partially supported by Fundação para a Ciência e Tecnologia (FCT), EU FEDER and POCTI Program, namely by Post-Doctoral grant SFRH 727/2000 and Project POCTI/AGR/63322/2004 (MNB), Center for Logic and Computation (MLC), and Matemática Aplicada IISA (JOC) .

1. Introduction

Knowing the diet composition of herbivores is important for understanding their foraging ecology and for mediating their effects on vegetation and ecosystems. Different methods can be used to estimate herbivore diets, although most of them have limitations and must be selected according to aims of study. N-alkanes have been used to estimate diet composition of domestic animals (Duncan *et al.* 1999, Salt *et al.* 1992, Hutchings *et al.* 2000) and, less frequently, of wild herbivores (Bugalho *et al.* 2001, Hulbert *et al.* 2001, Rao *et al.* 2003). N-alkanes are long chain saturated hydrocarbons widespread in plant cuticles which can be recovered, after correcting for differential digestibility, in herbivore faeces as they are chemically inert (Dove and Mayes 1996, Mayes and Dove 2000). Additionally, since plant species are characterized by different concentrations of N-alkanes, the chemical markers recovered in faeces can be used to identify and quantify the plants ingested by herbivores.

Dove and Mayes 1991 suggested to model the problem of estimating diet composition based on N-alkanes with a constrained system of linear equations. A number of different least-squares procedures, including EATWHAT, developed in CSIRO, Australia (Dove and Moore 1995), have been designed to determine diet composition using this model. The requirement that the maximum number of components that can be differentiated in the diet is limited by the number of available markers, is common to all those procedures. In the case where there are more plant components than N-alkanes, it is usual to group individual plant species, and express diet estimates in terms of the plant groups defined. Those groups can be defined either functionally (e.g. grasses, legumes, browse) or statistically (e.g. pooling of species with similar N-alkanes profiles). The groups however may not be suitable. For instance, the group of legumes is not adequate to estimate nitrogen intake since species within this group may have quite different nitrogen content (Mayes and Dove 2000). Although Dove and Mayes (2000) suggest that N-alkanes can be used to differentiate plant parts in the diet of herbivores in our work "dietary component" always refers to plant species (whole sample of plant leaf and stem). Indeed plant species accounts for more than 85% of the variation in N-alkane concentrations over plant parts (Dove *et al.* 1996).

Our aim is to use N-alkanes to estimate diet composition of free-range herbivores feeding on complex plant communities, such as rangelands or other natural systems, where the number of plant species is generally high, usually outnumering the number of reliable N-alkanes. In these situations, establishing coherent groups of plant

species or limiting the potentially selected species to a pre-determined small subset may not be reliable.

We describe a model in which diet composition is identified with (possible) infinite bounded combinations of plant species. The upper and lower bounds on the concentrations of each plant species in the diets can easily be determined by linear programming. This approach can be applied equally to cases where the number of markers is lower, equal or greater than the number of species, thus avoiding the need for grouping species.

We tested the model on two data sets consisting of records of the N-alkane profiles of different plant species and faeces of sheep and deer collected in experiments conducted in Australia and Portugal. Each faecal sample has a distinct N-alkane content. By using a whole set of samples (4 and 9 samples of faeces for sheep and deer, respectively) our model incorporates information on the instrinsic variability of the data.

2. Mathematical description

The problem of estimating diet composition based on N-alkanes with a constraint system of linear equations, as introduced by Dove and Mayes 1991, can be stated as follows.

Suppose d is the number of N-alkanes in plants and faeces. Each faecal sample can be considered a point on the d-Euclidean space, which describes the concentration of each of the d N-alkanes. Each plant species can be interpreted as a vector in the d-Euclidean space, defining the value of each N-alkane in that plant. To ilustrate this we use a fictitious example with 2 N-alkanes, 2 plant species, and 6 faecal samples (Figure 1).

Diets are *linear combinations*, i.e., weighted sums of the vectors representing the plant species. The weight or *coefficient* of each vector defines the quantity of the corresponding species in the diet. Each coefficient value, divided by the sum of the coefficients, gives the proportion of each plant species in the diet.

If $P_1, P_2, \ldots, P_p$ are the d-vectors representing p plant species, each vector indicating the concentrations of N-alkanes on that species, the linear combination with coefficients $c_1, c_2, \ldots, c_p$ is $c_1P_1 + c_2P_2 + \cdots + c_pP_p$.

Clearly, linear combinations with negative coefficients are "nonsensical diets". Hence, we must restrict our attention to linear combinations having non-negative coefficients. The set of all non-negative linear combinations of vectors is called the

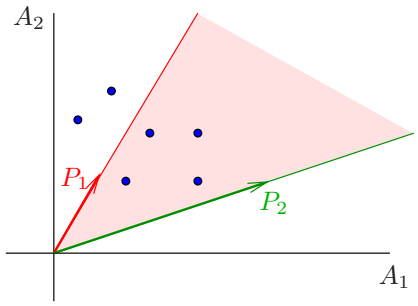

FIGURE 1. Example representing 2 N-alkanes (axes A_1 and A_2), 2 plant species (vectors P_1 and P_2), and 6 faecal samples (the 6 dots in the figure). The shaded area is the 2-dimensional cone generated by vectors P_1 and P_2.

cone generated by those vectors. In Figure 1 the cone generated by vectors P_1 and P_2 is the shaded region. We use C to denote the cone generated by the vectors corresponding to species, i.e.,

$$C = \{c_1P_1 + c_2P_2 + \cdots + c_pP_p, \text{with } c_1, c_2, \ldots, c_p \geq 0\}.$$

The estimation of diets from faecal samples $F_1, F_2, \ldots, F_q$ takes one sample F at a time and seeks whether either

(i) F belongs to the cone C, or
(ii) F is a point outside of C.

This is determined by solving the system of linear equations

$$Ax = F, \tag{S}$$

where A is the $d \times p$ matrix whose columns are the vectors $P_1, P_2, \ldots, P_p$ representing the plant species. If all components of the solution $\hat{x}$ are non-negative then (i) occurs, and $\hat{x}$ specifies the composition of the corresponding species in the faeces.

If there is a negative component, we are faced with situation (ii), indicating that only "nonsensical diets" can explain faeces F. This may result from errors in measuring the concentrations of N-alkanes, or from the existence of plant species which are included in the diet but not sampled in the field. The usual procedure in this case (e.g. Mayes *et al.* 1994, Salt *et al.* 1994, Dove and Moore 1995) is to find a non-negative vector of coefficients which "adequately approximates" $\hat{x}$. A reasonable approach consists of

(iii) identifying a point F' in the cone C considered to be "similar" to F, and
(iv) using F' instead of F, proceed as in (i) solving the system of equations (S).

Usually the projection of F onto C, i.e., the point in C which is *closest* to F, is selected to be F'.

In the example of Figure 1 there are two points in case (ii). Each of these points will be replaced by its projection onto C, that will lie in the *ray* of the cone defined by vector P_1. This implies that the herbage mixtures that will be derived from these two points consist only of P_1.

The estimated diet of the animal is finally computed from the solutions obtained with faecal samples $F_1, F_2, \ldots, F_q$. Often, the diet is defined to be the average of the q solutions. Note that this diet is in fact the solution of the linear system of equations (S) with F equal to the average point of the projections of the faecal samples $F_1, F_2, \ldots, F_q$ onto C. (The projection onto C of a point in C coincides with the point.)

We implicitly assumed that the number of N-alkanes is equal to the number of plant species, i.e., that A is a square (and non-singular) matrix. If there are more alkanes than species the situation does not change substantially, apart from a larger percentage of faecal samples F in condition (ii) that may now be expected. Figure 2 is an example of 3 N-alkanes, 2 plant species, and 6 faecal samples. Only one point is in the cone.

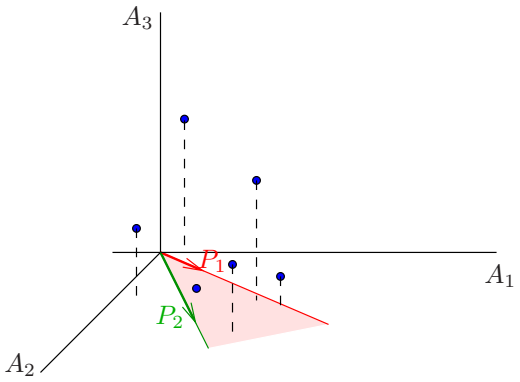

FIGURE 2. Example representing 3 N-alkanes (axes A_1, A_2 and A_3), 2 plant species (vectors P_1 and P_2), and 6 faecal samples (the 6 dots in the figure). The shaded area is the 2-dimensional cone generated by vectors P_1 and P_2.

Consider now that there are more plant species than N-alkanes. Figure 3 is an example of 2 N-alkanes, 3 plant species, and 6 faecal samples.

The situation has now changed considerably. Now any non-negative linear combination of vectors P_1 and P_3 of Figure 3 defines a new vector P contained in the

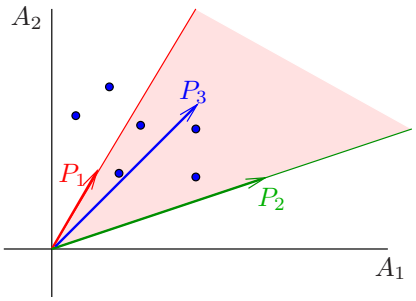

FIGURE 3. Example representing 2 N-alkanes (axes A_1 and A_2), 3 plant species (vectors P_1, P_2 and P_3), and 6 faecal samples (the 6 dots in the figure). The shaded area is the 2-dimensional cone generated by vectors P_1, P_2 and P_3.

cone generated by P_1 and P_3. Any point F inside the cone generated by P and P_2 is a non-negative linear combination of P (and consequently of P_1 and P_3) and P_2. Since there may be be infinitely many vectors P defined from P_1 and P_3 such that F is inside the cone generated by P and P_2, system (S) may have an infinite number of non-negative solutions. In other words, the concentration of N-alkanes in faecal samples may result from infinitely many different plant mixtures.

Whenever the number of species exceeds the number of N-alkanes, the usual approach is to group species into relevant categories (e.g. Rao *et al.* 2003, Hulbert *et al.* 2001). This is settled defining a partition of the set of vectors $P_1, P_2, \ldots, P_p$ into d (the number of N-alkanes) subsets, and identifying each subset with a certain vector. Once this is done, the procedure described for case $d = p$ is used to estimate the diet from the faecal samples $F_1, F_2, \ldots, F_q$. The estimated diet consists of the non-negative coefficients of the linear combination of the vectors identified with the subsets of the partition.

However, there is no satisfactory criteria to decide how species should be grouped (in other words, how to define a *suitable* d-partition of the set of vectors $P_1, P_2, \ldots, P_p$), and how to weight species within each group (in other words, how to select a vector which *adequately* represents a whole subset of vectors). The choices made with respect to those questions will affect the final result and will ultimately determine the estimated diet.

We suggest an alternative approach to estimate diet composition which avoids these difficulties and allows arbitrary numbers of plant species and N-alkanes.

3. The model

Let $P_1, P_2, \ldots, P_p$ be the d-vectors representing the p potentially selected plant species, where d is the number of N-alkanes used, and let C be the cone generated by these vectors. If P is one of the vectors $P_1, P_2, \ldots, P_p$, let C_P be the cone generated by the remaining vectors. We denote by F_i, $i = 1, 2, \ldots, q$, the projection onto C of the point on the d-Euclidean space representing each of the q faecal samples, and define $\mathcal{F} = \{F_1, F_2, \ldots, F_q\}$ as the set of "corrected" faecal samples.

We define *feasible diet* as every non-negative solution of the linear system $Ax = F$, for any point F in $\mathcal{F}$, where A is the $d \times p$ matrix with columns $P_1, P_2, \ldots, P_p$. There may be infinitely many feasible diets. Nevertheless, the set of diets is bounded by specified upper and lower bounds on the values of each component. Linear programming is a suitable tool to address questions about dietary composition such as:

- which are the maximum and minimum proportions of each species in the diet?
- which are the maximum and minimum proportions of each functional group (e.g grasses, legumes, browse) in the diet?
- which is the minimum proportion of a particular plant species in those diets that satisfy some nutritional requirement (e.g levels of nitrogen intake or tannin contents)?

In particular, the knowledge of $M_P^i = \max\{x_P$, subject to $Ax = F_i$ and $x \geq 0\}$ and $m_P^i = \min\{x_P$, subject to $Ax = F_i$ and $x \geq 0\}$, for $i = 1, 2, \ldots, q$, allow to answer questions regarding presence or absence of plant species in the diet. For instance:

- which species are never eaten (absent species)?
- which species are always eaten (mandatory species)?
- which species are identifiable in some faecal samples but not in all (conditional mandatory species)?

We cannot confirm if species that do not fall in any of the previous categories (optional species) are eaten or not. Contrary to the absent species, whose corresponding concentrations of N-alkanes are not recovered in any faeces, there is evidence of the chemical markers of an optional plant species in faeces. However, the concentration of N-alkanes of an optional species is identical to those derived from certain combinations of other potentially selected plants. To confirm if an

optional species is actually eaten may require further field work to collect evidence of the utilization of that species.

3.1. **Absent species.** *Absent species* are the plants not included in any diet. Species P is absent if the corresponding coefficient in every non-negative solution of $Ax = F$, for every F in $\mathcal{F}$, is equal to zero. This can be determined by checking whether $M_P^i = 0$, for $i = 1, 2, \ldots, q$.

In Figure 4 species P_4 is the unique absent plant.

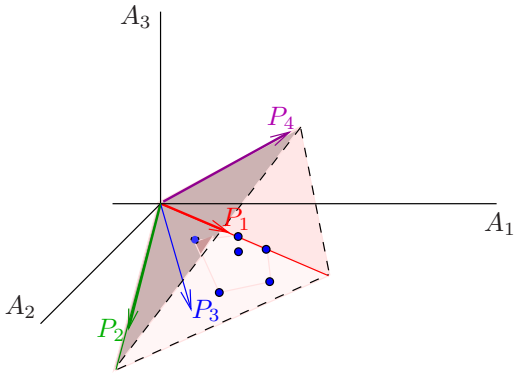

FIGURE 4. Example representing 3 N-alkanes (axes A_1, A_2 and A_3), 4 plant species (vectors P_1, P_2, P_3 and $P4$), and 6 faecal samples (the 6 dots in the figure). The shaded area is the 3-dimensional cone generated by vectors P_1, P_2, P_3 and P_4.

Once the absent species have been identified they can be removed from further consideration. Figure 5 represents the example in Figure 4 after the absent plant species P_4 have been removed. From now on we assume that there are no absent species.

3.2. **Mandatory species.** *Mandatory species* are the plants which are present in every diet. Species P is mandatory if the corresponding coefficient in every non-negative solution of $Ax = F$, for every F in $\mathcal{F}$, is greater than zero. From a geometrical point of view, P is mandatory if there is no F_i in the cone C_P. To decide if species P is mandatory corresponds to check if $m_P^i > 0$, for $i = 1, 2, \ldots, q$.

In Figure 5 species P_1 is the unique plant which is mandatory.

3.3. **Conditional mandatory species.** *Conditional mandatory species* are plants which are obligatory for some diets, but not for every diet. Species P is conditional mandatory if for some, but not all, F in $\mathcal{F}$ there is no non-negative solution of $Ax = F$ with the coefficient corresponding to P equal to zero. Equivalently, species

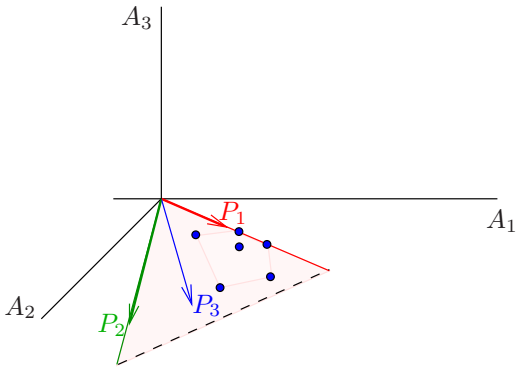

FIGURE 5. Example representing 3 N-alkanes (axes A_1, A_2 and A_3), 3 plant species (vectors P_1, P_2 and P_3), and 6 faecal samples (the 6 dots in the figure), obtained removing P_4 from the example in Figure 5. The shaded area is the 2-dimensional cone generated by vectors P_1, P_2 and P_3.

P is conditional mandatory if there are points of $\mathcal{F}$ which are in C_P and others which are not. Testing if species P is conditional mandatory can be achieved by checking whether, for some but not every $i = 1, 2, \ldots, q$, $m_P^i > 0$.

If plant P_2 (P_3) were not present in Figure 5, P_3 (P_2) would be an example of one such species.

3.4. **Optional species.** *Optional species* are the plants which are not in any of the previous cases. Species P is optional if for every F in $\mathcal{F}$ there are non-negative solutions of $Ax = F$ for which the coefficient P is equal to zero, but for at least one right hand side F a non-negative solution exists for which the coefficient of P is greater than zero. Species P is optional if (P is not absent and) $\mathcal{F}$ is contained in C_P. If species P is optional then $m_P^i = 0$, for every $i = 1, 2, \ldots, q$, and $M_P^i > 0$, for at least one i.

In Figure 5 P_2 and P_3 are both optional species.

Optional species are the only plants for which we cannot conclude if they were eaten or not. Hence it may be interesting to identify a minimum size subset of plants, with no optional species, that generate a cone that still contains $\mathcal{F}$. This can be achieved solving a mixed integer linear programming problem (with pq continuous variables and p binary variables, and $p(q+1)$ constraints). It may also be relevant to examine if a given group of optional species (e.g. legumes, grasses) are eaten. If, excluding from the input every plant of the group, the cone generated

by the remaining species no longer contains $\mathcal{F}$, we can conclude that at least one of the plants is eaten.

We tested this model on two data sets where these questions were highlighted in estimating the proportions of each plant species in the diets.

4. The data sets

The two data sets consist of records of the N-alkane profiles of plant species and faeces of sheep and deer collected in experiments conducted in Australia and Portugal, respectively.

4.1. **The Australian sheep data set.** Australian sheep data set came from a grazing experiment conducted at Ginninderra Experimental Station (35°12' S, 149°12' E) in New South Wales, Australia. This experiment aimed to assess the effects of grazing management on the productivity and persistence of Lotus species (Ayres and Blumenthal 2000) and consists of the N-alkane profiles of four sown cultivars of Lotus (L. corniculatus cv. Goldie, L. corniculatus cv. Prostrate, L. pedunculatus cv. Grasslands Maku and L. pedunculatus cv. Sharnae), two sown grasses (Phalaris aquatica and Austrodanthonia richardsonii), and volunteer grasses (Vulpia myuros, Bromus catharticus and Festuca arundinacea) and legumes (Trifolium repens, T. striatum and T. glomeratum).

The faecal data consists of the N-alkane composition of faecal samples collected from four different sheep that grazed the study area during five days. After a preliminary period of two days to allow passage of faeces from previously grazed material, four rectal grab samples of faeces were obtained from each sheep over three consecutive days. Faecal samples were bulked within sheep across days, resulting in one bulk faeces sample per sheep. These samples were stored at -18° C before sample preparation and analysis as described in Kelman *et al.* 2003.

During the grazing period, samples of whole plants of the sward species were harvested from random sites within the sward, using electric clippers. For each species, the harvested material was bulked until there was sufficient amount for chemical analysis and then stored at -18° C. Faecal and plant samples were freeze dried and ground prior to the extraction of cuticular wax N-alkanes by direct saponification, followed by solvent separation and purification through silica gel. Purified N-alkanes were quantified by gas chromotography relative to docosane (C22) and tetratriacontane (C34) N-alkanes as internal standards. Faecal N-alkane recovery corrections were based on those reported by Dove and Oliván 1998.

Proportions of plant species available at the sward, before the grazing trial, were 0.31 and 0.07 of the legumes Lotus spp and Trifolium spp, respectively, and 0.27, 0.17 of the grasses Phalaris aquatica and Vulpia myuros, respectively. Remaining proportion (0.18) of the sward was composed by the grasses A. richardsonii, B. catharticus and Festuca arundinacea(Kelman et al. 2003).

4.2. **The Portuguese deer data set.** This data set was collected in a study area located in Vila Viçosa (38°47' N, 7°25' W) in Southern Portugal as part of a general program on the foraging behaviour of red deer (Cervus elaphus) in a Mediterranean environment. Samples of leaf and stem of cork oak (Quercus suber), holm oak (Quercus rotundifolia), gum cistus (Cistus ladanifer) blackberry (Rubus ulmisolius), were collected in the beginning of summer (June).

About 30 g dry weight of a sample of leaf and stem (diameter not greater than 0.5 cm) of each browse species was collected from 20 different individuals and pooled within species. Samples were kept in sealed polythene bags, tagged, and stored at -15° C in a freezer until they were analysed.

Samples of species of the herbage layer, consisting of about 30 g of dry weight, were also collected in the study area where deer had been seen feeding previously. These samples, consisting of approximately 30 g dry weight of plant leaf and stem, were collected from the following species: Vulpia bromoides, Phalaris brachystachys, Holcus lanatus, Gaudinia fragilis, Agrostis porreti, Briza maxima, Avena barbata, Lolium multiflorum and Bromus hordeaceus (grasses); Trifolium arvensis, Trifolium angustifolium, Vicia sativa, Trifolium campestris, Trifolium subterraneum and Ornithopus compressus (legumes); Chamameleum mixtum and Coleostepus micorris (Asteraceae) Spergularia purpurea and Echium plantagineum (other species).

Nine samples of fresh faeces of red deer were also collected. Each sample consisted of the full group of faecal pellets found at a particular location. These samples were kept in sealed polythene bags, tagged, and stored at -15° C until they were oven-dried at 60° C for 48 hours.

After milling through a 1 mm size mesh, plant and faecal samples were analysed for N-alkane composition by gas chromatography using the method of Mayes *et al.* 1986 with the modification described by Salt *et al.* 1992. N-alkanes C22 and C34 were used as internal standards.

Proportions of plant species available at the sward were 0.50, 0.25, 0.22 and 0.02 of grasses, legumes, Asteraceae and "other species", respectively (Bugalho and Milne, 2003).

5. Computational results

The computational results were obtained using Microsoft Excel with the LINDO Systems What'sBest! solver for the optimization problems (a solver tolerance of 10^{-6} was given).

In both data sets all the faecal samples were slightly out of the cone generated by the plants. This could result from laboratory errors in measuring the concentrations of the markers. Hence projections onto the cone had to be performed. This was achieved, using quadratic programming, by finding the point in the cone which is closest (the Euclidean distance was used) to each faecal sample. The "projected faeces" obtained this way differed at most 2% on each coordinate from the original ones, which is a value below the usually admitted analytical error of 3 to 4% when measuring N-alkane concentrations (H. Dove, personal communication).

For the purpose of classification of the plant species according to Section 3 we assumed that values per unit less than 0.01 are equal to zero, to cope with round-off errors of the solver. The contributions of each plant species in the diets are expressed as proportions.

5.1. **The Australian sheep data set.** For this example we used 9 N-alkanes (C25 to C33) and 12 plant species. There were 2 mandatory legumes (T.repens and T.striatus) and 3 mandatory grasses (Vulpia myuros, Bromus catharticus and Festuca arundinacea) in the diet of sheep. Among mandatory legumes T.repens had the highest proportion in the diet (proportions varying between 25% and 34%) while T.striatus (5% and 10%) was consumed at comparatively lower proportions. Mandatory grasses V. myuros (9% to 23% of the diet), B. catharticus (14% to 27%) and F.arundinacea (11% to 27%) occurred with approximately similar proportions in the diet (Table 1).

There were 4 conditional mandatory legumes (T. glomeratum, L. pedunculatus Sharnae, L. pedunculatus Maku) and one conditional mandatory grass (F. arundinacea) in the diet. When occurring, cluster clover, L. pedunculatus Sharnae and L. pedunculatus Maku reached 3%, 4% and 2 %, respectively, of the diet, but were not detectable in remaining samples. The grass P. aquatica reached 5% in faecal F_1 but was absent from remaining samples (Table 1).

	status	F_1		F_2		F_3		F_4	
		Min	Max	Min	Max	Min	Max	Min	Max
L. corniculatus Prostrate	absent	0	0	0	0	0	0	0	0
L. corniculatus Goldie	absent	0	0	0	0	0	0	0	0
L. pedunculatus Sharnae	cond. mand.	0.04518	0.04518	0	0	0.01374	0.01374	0.01812	0.01812
L. pedunculatus Maku	cond. mand.	0	0	0	0	0.01972	0.01972	0	0
A. richardsonii	absent	0	0	0	0	0	0	0	0
P. aquatica	cond. mand.	0.04810	0.04810	0	0	0	0	0	0
V. myuros	mandatory	0.22702	0.22702	0.09375	0.09375	0.12632	0.12632	0.12293	0.12293
Bromus catharticus	mandatory	0.25366	0.25366	0.26611	0.26611	0.13777	0.13777	0.21798	0.21798
F. arundinacea	mandatory	0.11879	0.11879	0.18392	0.18392	0.23013	0.23013	0.23105	0.23105
T. striatus	mandatory	0.05146	0.05146	0.08413	0.08413	0.09908	0.09908	0.04962	0.04962
T. glomeratum	cond. mand.	0	0	0	0	0	0	0.03343	0.03343
T. repens	mandatory	0.24747	0.24747	0.36785	0.36785	0.37324	0.37324	0.32686	0.32686

TABLE 1. Results for the Australian sheep data set. Classification of plant species, and minimum and maximum proportions of each plant species in the diet of sheep. (Values less than 0.01 appear as 0.) F_i represents each faecal sample projection.

Both varieties of L. corniculatus (legumes) and A. richardsonii (a grass) were absent from the diet. No species were classified as optional (Table 1).

The minimum and maximum proportions of each plant species corresponding to each F_i in Table 1 almost coincide. Actually the small differences between these values are certainly due to round off computational errors. Note that since 3 species are absent from the diet, the 9 remaining intervenient plant species equal the number of N-alkanes. Consequently (S) becomes a square linear system with a unique solution.

5.2. **The Portuguese deer data set.** In this example we used 9 N-alkanes (C25 to C33) and 17 plant species. There were 3 mandatory legumes (Trifolium arvensis, T. subterraneum and Vicia sativa) a mandatory Asteraceae (Chamameleum mixtu) and a mandatory browse species (Cistus ladanifer) in the diet of red deer. The mandatory legumes T. arvensis, V. sativa and T. subterraneum had proportions of 3% to 22%, 5% to 30% and 2% to 9%, respectively. The only mandatory browse species, C. ladanifer, varied between 21% and 42% whilst C. mixtum (a Asteraceae) varied between 4% and 17% of the diet (Table 2).

There were 6 conditional mandatory species: T. angustifolium, Ornithopus compressus (legumes), Lolium multiflorum (grass), Coleostepus mycorris (Asteraceae), Quercus suber (browse) and Echium plantagineum (an herb). As compared with other conditional mandatory species T. angustifolium occurred in most of the samples and reached relatively high proportions in some of them (e.g. faecal samples F_2 and F_9, with 42% and 31%, respectively). Remaining mandatory species were detectable in a lower number of samples but some still reaching relatively high values.

	status	F_1		F_2		F_3		F_4		F_5		F_6		F_7		F_8		F_9	
		Min	Max	Min	Max	Min	Max	Min	Max	Min	Max	Min	Max	Min	Max	Min	Max	Min	Max
V. bromoides	absent	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
S. purpurea	absent	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
H. lanatus	absent	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
A. barbata	absent	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
L. multiflorum	cond. mand.	0.03895	0.03926	0	0	0	0	0	0	0	0	0	0	0	0	0.02044	0.02082	0	0
E. plantaginum	cond. mand.	0.06156	0.06184	0	0	0	0	0	0	0	0	0	0	0.19162	0.19175	0.0928	0.09314	0	0
C. mixtu	mandatory	0.0444	0.0445	0.12092	0.12097	0.02143	0.02148	0.0772	0.07723	0.11132	0.11136	0.17312	0.17317	0.07724	0.07728	0.12361	0.12374	0.05581	0.05585
C. mycorris	cond. mand.	0.09946	0.0997	0	0	0.04612	0.0463	0	0	0	0	0	0	0	0	0.12212	0.1224	0	0
T. arvensis	mandatory	0.15865	0.15873	0.04913	0.04916	0.21216	0.21218	0.2183	0.21835	0.17349	0.17354	0.1769	0.17695	0.13009	0.13013	0.1533	0.15339	0.03098	0.03103
T. angustifolium	cond. mand.	0	0	0.42273	0.42289	0	0	0.19022	0.19035	0.2634	0.26354	0.11155	0.11169	0.21499	0.21513	0	0	0.31929	0.31955
V. sativa	mandatory	0.30179	0.30179	0.05837	0.05841	0.14611	0.14613	0.04807	0.0481	0.12263	0.12267	0.14662	0.14667	0.11466	0.1147	0.12501	0.12504	0.21039	0.21049
T. subterraneum	mandatory	0.08733	0.08735	0.06268	0.06269	0.03966	0.03967	0.04723	0.04723	0.08467	0.08467	0.04897	0.04898	0.02464	0.02465	0.07467	0.07469	0.01737	0.01737
O. compressus	cond. mand.	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.02615	0.02618
Q. suber	cond. mand.	0	0	0.07815	0.07818	0.19483	0.19489	0	0	0	0	0	0	0	0	0	0	0	0
Q. rotundifolia	absent	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
C. ladanifer	mandatory	0.20716	0.20717	0.20771	0.20771	0.33947	0.33949	0.41873	0.41871	0.24422	0.24421	0.34254	0.34253	0.24646	0.24646	0.28719	0.2872	0.33954	0.33951
R. ulmifolius	absent	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

TABLE 2. Results for the portuguese deer data set. Classification of plant species, and minimum and maximum proportions of each plant species in the diet of deer. (Values less than 0.01 appear as 0.) F_i represents each faecal sample projection.

This is the case of Q. suber and E. plantagineum that occurred with proportions of

19% in faecal samples F_3, and F_8, respectively, or C. mycorris with 12% in faecal samples F_9 (Table 2). O. compressus and L. multiflorum have always very low proportions and may have been misclassified as conditional mandatory.

Vulpia bromoides, Holcus lanatus, Avena barbata (grasses), R. ulmifolius, Quercus rotundifolia (browse) and Spergularia purpurea (herb) were absent from the diet of red deer. There were no species classified as optional (Table 2).

As for the case of the Australian sheep data set, the minimum and maximum proportions of each plant species corresponding to each F_i in Table 2 are similar. From the original set of 17 plant species, 6 are absent and 2 other (O. compressus and L. multiflorum) have very low contributions, suggesting that only 9 plant species (as many as the number of N-alkanes) are effective in solving the linear system (S). This explains the observed similarity of the minimum and maximum values.

6. Discussion

The approach described in the present paper does not limit the number of dietary components to be identified in the diet of herbivores. This is particularly useful when using N-alkanes to estimate diet composition of animals feeding on complex plant communities. In rangelands or other natural systems, where herbivores have large numbers of plant species available for selection, the method described may provide relevant information to the analist by identifying plants which are certainly eaten and others which are excluded from the diet. Additionally, by giving information on minimum and maximum proportions of a particular plant species in the faeces, the method provides an approach to quantification of diet composition. Two real data sets allowed testing of the approach described.

6.1. **The Australian sheep data set.** Our model suggested strong selection of clovers by sheep and particularly of T.repens. From proportions of 0.07 in the at the sward (see methods). T.repens was found in the diet, for instance, at proportions of 0.37. High consumption of T.repens, in agreement with results shown by Kelman *et al.* 2003 for the same data set. Conversely, Lotus species were not mandatory and, when occurring in the diet, were consumed at very low proportions, particularly when compared with proportions available at the sward (0.31, see methods). A maximum of 0.04 of the diet was reached by L. pedunculatus Sharnae. Lotus species are characterized by the presence of tannins in their stem and leaves (Reed 1995, van Soest 2004). Tannins are anti-nutritional components of plant species which

may depress diet digestibility (van Soest 2004). As a consequence, plants rich in tannins, such as Lotus spp. are expected to be avoided by sheep. If ingested at low levels, however, tannins may form complexes with proteins, protecting them from microbial enzymes in the rumen and increasing protein digestibility in the lower digestive tract (Reed 1995, Bento 2004). Other potential positive benefits such as antioxidant, antibacterial and antihelmintic properties could account for the intake of low doses of tannins (Clausen *et al.* 2003, Iason 2005). Low proportions of Lotus spp. in the diet of sheep, as found here, were not unlikely. Indeed Verheyden-Tixier and Duncan 2000 found that roe deer (Capreolus capreolus), a wild ruminant, actively included a certain amount of tannins in their daily ration.

It is known that concentrations of condensed tannins are genetically determined and may vary widely among varieties of the same plant species. The present approach potentially allowed the differentiation of the two varieties of L. pedunculatus - Sharnae and Maku - in the diet of sheep which may be of potential interest. Grasses were included in the diet at proportions similar to their availability in the sward as could also be expected and also shown by Kelman *et al.* (2003).

6.2. **The Portuguese deer data set.** Our computational results suggest that deer consumed different species of legumes. Although we do not have data on availability of individual species, the group of legumes comprised 0.25 of the sward (Bugalho and Milne 2003) and for legumes such as Trifolium angustifolium or Vicia sativa proportions in the diet were in some cases well above 0.25, suggesting that deer were selecting legumes for their diets. A grazing exclusion experiment in the same study site has also shown that the proportion of legumes were significantly higher in areas where deer grazing had been excluded which again suggests selection of legumes by red deer (Bugalho *et al.* 2006). Legumes such as V. sativa and T. subterraneum, 2 cultivated and ameliorated species, were mandatory in red deer diet. This could be expected as cultivated species generally have higher nutritive value and are preferentially selected than wild varieties. With respect to browse, only a shrub (C. ladanifer) was mandatory in the diet, as estimated by our model, with Q. suber (conditional mandatory) found only in one faecal sample and Q. rotundifolia absent from the diet. In Mediterranean environments, such as that occurring in the Portuguese deer study area, browse species are predominantly consumed by ruminants between mid-summer to the end of summer, when most of the herbage layer is senescent and of low nutritive value (Seligman 1996, Bugalho and Milne 2003). Indeed present data was collected in the beginning of June, when

a proportion of green plant material was still available in the herbage layer and thus a very high consumption of browse should not be expected.

7. Conclusions

Previous approaches for estimating diet composition using N-alkanes require that the number of markers is greater or equal than the number of dietary components. However, this constraint is not realistic in natural or semi-natural systems where a high number of plant species is generally available for selection by the herbivores and usually exceeds the number of plant markers. While the method described in this paper needs further validation with additional data sets, it gives relevant information to the analyst by identifying plants which are certainly eaten and others which are excluded from the diet. Furthermore, since it provides minimum and maximum proportions of each particular plant species in the faeces, the method also allows us to quantify the diet composition.

Acknowledgements. We thank Hugh Dove and Walter Kelman for permission to use the Australian sheep data set. We also thank Kathleen Merrill for corrections on an early version of the manuscript.

References

Ayres, J.F., and Blumenthal, M.J. 2000. Lotus grazing management for weaner production. Final Report, Project DAN082. Meat and Livestock Australia. ISBN 0 7347 1126 3. New South Wales Agriculture, Australia.

Bento, H. 2004. The effects of condensed tannins in rumen and post-rumen digestion of nutrients. PhD Thesis. University of Aberdeen. Aberdeen.

Bugalho, M.N., and Milne J.A. 2003. The composition of the diet of red deer (Cervus elaphus) in a Mediterranean environment: a case of summer nutritional constraint? *Forest Ecology and Management.* **181**: 23-29.

Bugalho, M.N., Milne J.A., and Racey, P.A. 2001. The foraging ecology of red deer (Cervus elaphus) in a Mediterranean environment: is a larger body size advantageous? *Journal of Zoology (London)* **255**: 285-289.

Bugalho, M.N., Lecomte, X., Pile S., and Caldeira, M.C. 2006. Effects of deer on oak regeneration and species composition of a Mediterranean grassland. *Revista de Ciências Agrárias* (in press).

Clauss M., Lason K., Gehrke J., Lechner-Doll M., Fickel J., Grune T., and Streich W.J. 2003. Captive roe deer (Capreolus capreolus) select for low amounts of tannic acid but not quebracho: fluctuation of preferences and potential benefits. *Comparative Biochemistry and Physiology Part B* **136**: 369-382.

Dove, H., and Mayes, R.W. 1991. The use of plant wax alkanes as marker substances in studies of the nutrition of herbivores: a Review. *Australian Journal of Agricultural Research* **42**: 913-952.

Dove, H., and Mayes, R.W. 1996. Plant wax components: A new approach to estimating intake and diet composition in herbivores. *Journal of Nutrition* **126**: 13-26.

Dove, H., Mayes, R.W. and Freer, M. 1996. Effect of species, plant part and plant age on the n-alkane concentrations in the cuticular wax of pasture plants. *Australian Journal of Agricultural Research* **47**: 1333-1347.

Dove, H., and Moore, A.D. 1995. Using a least-squares optimization procedure to estimate botanical composition based on the n-alkanes of plant cuticular wax. *Australian Journal of Agricultural Research* **46**: 1535-1544.

Dove, H., and Oliván, M. 1998. Using synthetic or bees-wax alkanes for estimating supplement intake in sheep. *Animal Production In Australia* **22**: 189-192.

Duncan, A.J., Mayes, R.W., Lamb, C.S., Young, S.A., and Castillo, I. 1999. The use of naturally occurring and artificially applied n-alkanes as markers for estimation of short-term diet composition and intake in sheep. *Journal of Agricultural Science, Cambridge* **132**: 233-246.

Hulbert, I.A., Iason, G.R., and Mayes, R.W. 2001. The flexibility of an intermediate feeder: dietary selection by mountain hares measured using n-alkane analysis. *Oecologia* **129**: 197-205.

Hutchings, M.R., Gordon, I.J., Robertson, E., Kyriazakis, I., and Jackson, F. 2000. Effects of parasitic status and level of feeding motivation on the diet selected by sheep grazing grass/clover swards. *Journal Agricultural Science* **135**: 65-75.

Iason G.R. 2005. The role of plant secondary metabolites in mammalian herbivory: ecological perspectives. *Proceedings of the Nutrition Society* **64**: 124-131.

Kelman, W.M., Bugalho, M.N., and Dove, H. 2003. Cuticular wax alkanes and alcohols used as markers to estimate diet composition of sheep (Ovis aries). *Biochemistry Systematics and Ecology* **31**: 919-927.

Mayes, R.W., Beresford, N.A., Lamb, C.S., Barnett, C.L., Howard, B.J., Jones, B-E.V., Eriksson, O., Hove, K., Pedersen, O., and Staines, B.W. 1994. Novel approaches to the estimation of intake and bioavailability of radiocaesium in ruminants grazing forested areas. *Science of the Total Environment* **157**: 289-300.

Mayes, R.W., and Dove, H. 2000. Measurement of dietary nutrient intake in free-ranging mammalian herbivores. *Nutrition Research Reviews* **13**: 107-138.

Mayes, R.W., Lamb, C.S., and Colgrove, P.M. 1986. The use of dosed and herbage n-alkanes as markers for the determination of herbage intake. *Journal of Agricultural Science, Cambridge* **107**: 161-170.

Rao, S.J., Iason, G.R., Hulbert, I.A.R., Mayes, R.W., and Racey, P.A. 2003. Estimating diet composition for mountain hares in newly established native woodland: development and application of plant-wax faecal markers. *Canadian Journal of Zoology* **81**: 1047-1056.

Reed, J.D. 1995. Nutritional toxicology of tannins and Related polyphenols in forage legumes. *Journal of Animal Science* **73**: 1516-1528.

Salt, C.A., Mayes, R.W., Colgrove, P.M., and Lamb, C.S. 1994. The effects of season and diet composition on the radiocaesium intake by sheep grazing heather moorland. *Journal of Applied Ecology* **31**: 125-36.

Salt, C.A., Mayes, R.W., and Elston, D.A. 1992. Effects of season, grazing intensity and diet composition on the radiocaesium intake by sheep on re-seeded hill pasture. *Journal of Applied Ecology* **29**: 378-387.

Tixier, H., Duncan, P., Scehovic, J., Yani, A., Gleizes, M., Lila, M. 1997. Food selection by European roe deer (Capreolus capreolus): effects of plant chemistry, and consequences for the nutritional value of their diets. *Journal of Zoology (London)* **242**: 229-245.

Seligman, N.G. 1996. Management of Mediterranean Grasslands. In The Ecology and Management of Grazing Systems. Edited by J. Hodgson and A.W. Illius. CAB International, Wallingford. pp: 359-392.

van Soest, P.J. 1994. Nutritional Ecology of the Ruminant 4th ed. O&B Books, Corvallis, OR.

Verheyden-Tixier, H., Duncan, P. 2000. Selection for small amounts of hydrolysable tannins by a concentrate-selecting mammalian herbivore. *Journal of Chemical Ecology* **26**: 351-358.

Fac. de Economia,
Universidade Nova de Lisboa,
Campus de Campolide,
1099-032 Lisboa, Portugal
E-mail address: barcia@fe.unl.pt

Centro de Ecologia Aplicada "Baeta Neves",
Inst. Sup. de Agronomia,
Universidade Técnica de Lisboa,
Tapada da Ajuda,
1349-017 Lisboa, Portugal
E-mail address: migbugalho@isa.utl.pt

Inst. Sup. de Agronomia
and SQIG-IT,
Universidade Técnica de Lisboa,
Tapada da Ajuda,
1349-017 Lisboa, Portugal.
E-mail address: mlc@isa.utl.pt

Dep. Matemática,
Inst. Sup. de Agronomia,
Universidade Técnica de Lisboa,
Tapada da Ajuda,
1349-017 Lisboa, Portugal.
E-mail address: orestes@isa.utl.pt