APPENDICE A

THE OMOSS METHOD (OPTIMIZATION METHOD BASED ON SINGLETON SPECIES)

Daniele Lostia
RIASSUNTO
L’obiettivo di proteggere tutte le specie presenti in un determinato territorio, intervenendo nel minor numero possibile di quadranti può essere espresso matematicamente attraverso una serie di formule che nel loro complesso sono note come ILP (Integer Linear Programming o Programmazione lineare a numeri interi), e sono risolvibili attraverso appositi algoritmi disponibili anche sui PC.
Gli algoritmi che risolvono un problema di ILP sono però piuttosto pesanti in termini di tempi di calcolo; inoltre tali tempi di calcolo crescono esponenzialmente al crescere del numero di quadranti, rendendo ben presto impossibile, nella pratica, ottenere una soluzione ottimale, cioè un insieme di quadranti tale che ogni quadrante dell’insieme non possa essere eliminato senza perdere alcuna delle specie inizialmente presenti nel territorio considerato.
Di conseguenza si impone la necessità di trovare un metodo rigoroso per semplificare il problema originario portandolo a dimensioni piccole, e gestibili agevolmente dai programmi di calcolo ILP.
Il metodo OMOSS (Optimization Method Based On Singleton Species) che viene qui presentato, si basa essenzialmente su una caratteristica ben nota della distribuzione delle specie nei quadranti: un numero piccolo di specie è presente in molti quadranti, mentre un gran numero di specie è presente in un piccolo numero di quadranti. OMOSS sfrutta tale caratteristica per ottenere fin dalle prime iterazioni una drastica riduzione delle dimensioni del problema originale, riducendolo a dimensioni molto piccole o, in taluni casi, trovando direttamente una soluzione ottimale (cioè senza che sia necessario applicare algoritmi di ILP).
In particolare il processo di semplificazione (riduzione del numero di quadranti e di specie, e contestuale individuazione di quadranti che con certezza fanno o eliminazione di quadranti che con certezza non fanno parte della soluzione ottimale) si basa sull’applicazione ripetuta di uno o più dei 3 criteri seguenti:
1. Un quadrante che contiene una (o più) specie singleton deve necessariamente far parte della soluzione ottimale, in quanto in caso contrario la (o le) specie singleton non sarebbero presenti nell’insieme di quadranti ottimale. Questo criterio comporta l’individuazione di tutti i quadranti che contengono almeno una specie singleton; tali quadranti, necessariamente presenti nella soluzione ottimale, offriranno copertura a tutte le altre specie presenti negli stessi. Di conseguenza una volta applicato questo criterio il problema originario si ridurrà ai soli quadranti residui ed alle sole specie residue.
2. Un quadrante che, a seguito dell’applicazione del criterio 1, contenga una sola specie presente insieme ad altre in altri quadranti, può essere trascurato in quanto la specie in questione può essere coperta in maniera più efficace dagli altri quadranti in cui essa è presente. Anche l’applicazione di questo criterio produce una riduzione, sia pur piccola, del numero di quadranti, ma soprattutto può aprire la strada ad una nuova applicazione del criterio 1 in quanto la specie presente nel quadrante trascurato può essere diventata singleton nell’insieme di quadranti residui.
3. Se una specie è presente da sola in più quadranti (cioè non è presente mai assieme ad altre specie) tutti i quadranti che si trovano in tale condizione meno uno (che può essere scelto a caso) possono essere trascurati. Anche in questo caso si ha una riduzione del numero di quadranti a cui applicare eventualmente l’ILP.
Quando nessuno dei 3 criteri risulta ulteriormente applicabile, non resta che applicare l’ILP all’insieme di quadranti e specie residui. Tale insieme risulta generalmente molto piccolo e, talvolta, vuoto.
Per avere un’idea dell’efficacia del criterio 1, basti pensare che applicando una sola volta tale criterio ad un insieme di 786 specie in 540 quadranti si sono ottenuti:
1. 76 quadri (contenenti specie singleton) che davano copertura a 734 specie
2. Le residue 64 specie erano distribuite in soli 140 quadranti
3. Ben 324 quadranti sono stati eliminati in quanto contenenti solo specie già presenti nei 76 quadranti del punto 1
4. Pertanto la riduzione del numero delle specie è stata del 92% e quella del numero dei quadranti è stata del 74%
Poiché il numero minimo di quadranti necessario ad avere copertura di tutte le 786 specie è risultato pari a 107, i 76 quadranti individuati con la sola prima applicazione del criterio 1 costituiscono già il 71% dei quadranti della soluzione ottimale; il restante 29% andrà ricercato (attraverso la ripetuta applicazione dei 3 criteri) tra i 140 quadranti residui.
Come ulteriore risultato, il metodo OMOSS consente di costruire tutte le possibili soluzioni ottimali (che ovviamente avranno tutte lo stesso numero di quadranti, oltre che di specie). Il numero delle soluzioni ottimali può anche essere molto alto, e comunque la disponibilità di molte soluzioni che rispettano l’obiettivo originario (copertura di tutte le specie col minor numero di quadranti) consente di applicare ulteriori criteri di selezione/ottimizzazione. Si può, ad esempio, cercare la soluzione ottimale in cui la ricchezza media dei (107) quadranti è la più alta (massima ricchezza), oppure quella in cui è presente il minor numero di specie singleton (minor criticità) o quella in cui i quadranti sono più connessi tra di loro (massima connessione).
INTRODUCTION
Optimization task is identifying the minimum number of quadrants needed to obtain the coverage of all taxa, and, as we will see below, we will rely strongly on the concept of "singleton" taxa.

The objective can be summarized, in mathematical terms, in the following way:

Given a table of presence of \( M \) species in \( N \) disjoint quadrants of equal size, find the minimum number of quadrants that must be taken into account so that each species is present at least in one quadrant.

More analytically,

Given a matrix with a row for each species and a column for each quadrant:

\[
\begin{bmatrix}
C_{11} & \cdots & C_{1j} & \cdots & C_{1M} \\
\vdots & \ddots & \vdots & \ddots & \vdots \\
C_{i1} & \cdots & C_{ij} & \cdots & C_{iM} \\
\vdots & \ddots & \vdots & \ddots & \vdots \\
C_{N1} & \cdots & C_{Nj} & \cdots & C_{NM}
\end{bmatrix}
\]

that we will call "coefficient matrix"

where:
- \( N \) is the number of quadrants
- \( M \) is the number of species
- \( C_{ij} \) is a coefficient whose value is:
  - 1 if species \( j \) is present in quadrant \( i \)
  - 0 otherwise

Find the vector \( \mathbf{X} = [X_1, \ldots, X_N] \) minimizing (objective function)

where:
- \( X_i \) is a variable that can assume the values:
  - 1 if the quadrant is part of the solution
  - 0 otherwise

subject to the constraint system:

\[
\sum_{i=1}^{N} C_{ij} X_i \geq 1 \quad \text{for each } j \in \{0,1,\ldots,M\}
\]

(which means that each species must be present at least once in the optimal solution)

This type of problem is classified in literature as ILP (Integer Linear Programming) and there are various algorithms able to find a solution, but all impact with calculation times that grow very rapidly with size (especially number of quadrants in the grid), and quickly becomes impossible to find a solution in a reasonable time. It is therefore necessary to find a way to simplify the problem, to reduce it to manageable size. The OMB OSS method deals primarily with this simplification and confines the use of an ILP algorithm to the final phase, when it is no longer possible to simplify.

To obtain a simplification method, we start from the observation that the diffusion of the species in the quadrants in which a territory is divided follows a typical trend that is shown by the following graph (Fig. 5.1).

The graph, which refers to the actual data of the distribution of 786 rare species in Lazio, makes clear that a small number of species is present in many quadrants, while a large number of species is present in a small number of quadrants. For example, the graph shows that only one species (the most common) is present in 304 quadrants, while 138 species (the rarest) are present in a single quadrant. See note 1 in this regard.

For sake of brevity, we will say that a species is "covered" by a quadrant if it is present in the quadrant; by extension we will say that it is "covered" by a set of quadrants if it is covered by at least one of the quadrants of the set. In particular, we will be interested in species that are covered by a single quadrant ("singleton"), which we will call in the following "single-quadrant species".

We now introduce some "simplification criteria", which aim to reduce the size of the problem to the point where a "hard core" problem will remain, to be solved with ILP.
Criterion 1
To better illustrate OMBoss method, we will take an example. Suppose the following table (Tab. 5.1) represents the presence of 26 species (A ... Z) in 14 quadrants (a ... n); in the last row and in the last column are shown the column and row totals respectively.
We observe that species X is present only in quadrant f; therefore, quadrant f will certainly be part of the optimal solution. Elsewhere, species X would not be “covered” by the optimal solution.
It is also clear that, if the quadrant f is certainly part of the optimal solution, this fact implies that all the species present in f (ie B, C, I, L and Q) are certainly covered in the optimal solution.
We have highlighted quadrant f and all species contained therein, to remember that they are already assigned to the optimal solution. From now on our problem can therefore be limited to that of minimizing the number of quadrants needed to cover all the “other” species with the “other” quadrants; we will therefore work on the following species-quadrants table (Tab. 5.2).
The table is obtained from the previous one by deleting row f and all the columns containing a “1” in row f.
We now point out (Tab. 5.3) all single-quadrant species still present, (U, W, X, Y, Z), all quadrants that contain them (a, g, h, n) and therefore necessarily belong to the optimal solution and all species covered by these quadrants (A, F, K, N, P). The elements highlighted are all those that will necessarily be added, in the optimal solution, to those previously identified.
It is easy to verify that at this point the species-quadrants table (with only quadrants and species that have not yet been assigned to the optimal solution) is reduced to the following table (Tab. 5.4).
In this table single-quadrant species are no longer present. It should also be noted that at this point the quadrant e is empty; this derives from the fact that all the species it contains...
The OMB OSS (Optimization Method Based On Singleton Species) method

tains are covered by quadrants so far included in the optimal solution; therefore from now on quadrant e will no longer be taken into account.

Summing up, at present:
- Quadrants a, f, g, h, n and all species they cover are already part of the optimal solution
- Quadrant e is not part of the optimal solution, as it unnecessarily increases the number of quadrants, without adding species
- Quadrants b, c, d, i, j, k, l, m and residual species contained in them are still to be optimized

It should also be noted that, in the example, quadrant h contains more than one single-quadrant species.

We can now formulate the

Criterion 1:
If a species s is present only in the quadrant q:
- quadrant q (and all species contained in it) are assigned to the optimal solution (definitive assignment).
- quadrant q (and all species contained in it) can be from now on ignored, eliminating from the presence table row q and all columns containing a 1 in row q (simplification). The further optimization will therefore only concern residual quadrants and residual species, i.e. those quadrants for which it is not yet possible to establish whether they are part of the optimal solution or not and all residual species that they cover.

The simplification criterion can be applied repeatedly (or in a single step), until there are no single-quadrant species anymore.

Criterion 2
Let's go back to the sub-problem that has remained to be tackled; after Step 1 the table of presence that remains to be optimized is the following (10 species in 8 quadrants) (Tab. 5.5).

As already mentioned, on this matrix it is no longer possible to apply Criterion 1, because, as can be easily observed, it no longer contains any single-quadrant species.

However, we can observe that there are some quadrants (i, j, k, l, m) that contain only one species each; we will call these quadrants "mono-species" and "pluri-species" all others (b, c, d).

In particular, quadrants i and j (mono-species) contain species D, which is also present in the pluri-species quadrant b; this

Tab. 5.3 - Example table from tab. 3 without more "singleton species".

Tab. 5.4 - Example table from tab. 3 without more "singleton species".

Tab. 5.5 - Example table after applying Criterion 1 (no more singleton species).
means that quadrants i and j can be eliminated without this leading to the impossibility of achieving the optimal solution of the original problem. In fact, if optimal solution (i.e. solution with minimum number of quadrants and total coverage of species) contains quadrant i then is optimal also the solution that can be obtained by replacing quadrant i with the quadrant b. It should also be noted that from the point of view of biodiversity conservation, the inclusion of quadrant b instead of quadrant i in the optimal solution is also preferable, as it involves a “strengthening” of coverage for species J, M, O, R, S, T (which are not present in i).

Similar considerations lead to eliminate quadrant m. Therefore we can limit ourselves, in the following, to optimize the following table of presence (Tab. 5.6).

At this point we can formulate the

Criterion 2:
If a species s is present both in single-species quadrants and in multi-species quadrants:
- mono-species quadrants in which s is present are excluded from the optimal solution (definitive elimination).
- columns corresponding to the mono-species quadrants in which s is present can be eliminated from the presence matrix (simplification).

Criterion 3
Let us now go back to the last table of presence (Tab. 6), to observe that the mono-species quadrants k and l contain both same species E, which however is not present in any pluri-species quadrant. This means that both quadrants can not be eliminated, otherwise the species E would no longer be covered, and that consequently one (and only one) of them must necessarily be present in the optimal solution. If we choose to keep quadrant k and eliminate quadrant l, we already know that an optimal alternative solution can be obtained simply by replacing quadrant l with quadrant k. In this hypothesis, quadrant k will be part of the optimal solution and quadrant l will be replaceable with quadrant k in the optimal solution.

Both quadrant l and the quadrant k (and the species contained therein) can therefore be eliminated from the table, obtaining the following situation (Tab. 5.7).

With:
- quadrants a, f, g, h, n (with all the species they cover) already part of the optimal solution.
- quadrant l not in the optimal solution, but alternative to quadrant k.
- quadrants e, i, j, m not in the optimal solution.
- quadrants b, c, d (with the residual species contained in them) still to be optimized.

At this point we can formulate the

Criterion 3:
If a species s is present in one or more single-species quadrants but not in multi-species quadrants:
- any one and only one of the mono-species quadrants in which it appears must necessarily be present in the optimal solution (definitive assignment).
- the other single-species quadrants can be eliminated but must be kept in mind to generate optimal alternative solutions (temporary elimination).
- the columns corresponding to the quadrants assigned or eliminated, and all the species included in them can be removed from the presence table (simplification).

Criteria 2 and 3 can be applied (in one step) until the residual presence table no longer contains a single-species quadrant. After application of criteria 2 and 3, maybe it is again possible to apply Criterion 1 (Tab. 5.8).
The OMB OSS (Optimization Method Based On Singleton Species) method

**Species**

<table>
<thead>
<tr>
<th>QUADRANTS</th>
<th>SPECIES</th>
<th>Tot</th>
</tr>
</thead>
<tbody>
<tr>
<td>a</td>
<td>1 1 1 1 1 1 1 1</td>
<td>8</td>
</tr>
<tr>
<td>b</td>
<td>1 1 1 1 1 1 1 1 1</td>
<td>13</td>
</tr>
<tr>
<td>c</td>
<td>1 1 1 1 1 1 1</td>
<td>5</td>
</tr>
<tr>
<td>f</td>
<td>1 1 1 1</td>
<td>4</td>
</tr>
<tr>
<td>g</td>
<td>1 1 1 1 1 1</td>
<td>6</td>
</tr>
<tr>
<td>h</td>
<td>1 1 1 1 1</td>
<td>6</td>
</tr>
<tr>
<td>k</td>
<td>1 1 1</td>
<td>2</td>
</tr>
<tr>
<td>n</td>
<td>1 1 1</td>
<td>3</td>
</tr>
<tr>
<td>Tot</td>
<td>3 2 2 1 1 2 1 2 3 2 2 1 3 6</td>
<td>46</td>
</tr>
</tbody>
</table>

Tab. 5.9 - First optimal solution after applying the Criteria 1, 2, 3.

This is the case in our example, in which species D and H are present in quadrants b and c respectively. These quadrants contain all the species still present (including all those covered by quadrant d). Consequently, application of Criterion 1 leads to the following situation:

- quadrants a, b, c, f, g, h, k, n (with all the species they cover) are already part of the optimal solution.
- quadrant l is not part of the optimal solution, but is alternative to the quadrant k.
- quadrants d, e, i, j, m are not part of the optimal solution.
- There is nothing more to optimize.

At this point we can say that we have identified not one but two optimal solutions (a, b, c, f, g, h, k, n) and (a, b, c, f, g, h, l, n) both made up of 8 quadrants, and moreover without having to resort to ILP. Extracting from the original matrix only the lines relating to quadrants present in each of the two optimal solutions, it is easy to verify that both solutions identified completely cover the 26 species: first optimal solution (Tab. 5.9); second optimal solution (Tab. 5.10).

### Summarizing

**Application of Criterion 1 involves:**

1a) Identification of quadrants and related species that are part of the optimal solution.
1b) Reduction (often drastic, due to the high initial presence of singleton species) of the number of species to be considered in the following steps.

**Application of Criterion 2 involves:**

2a) Elimination of quadrants that certainly are not part of the optimal solution.
2b) No reduction in number of species.

**Application of Criterion 3 involves:**

3a) Identification of quadrants and related species that are part of the optimal solution.
3b) Identification of possible alternative quadrants, useful for generating further optimal solutions.
3c) Identification of species to be ignored in the following steps because covered by quadrants identified as belonging to optimal solution in previous point 3a.

### Criterion 4 (The “hard core”)

But what if, after repeatedly reducing the size of the problem, none of the 3 simplification criteria is still applicable? In the following example (Tab. 5.11).
There are neither single-quadrant species nor single-species quadrants. The only way is to use one of the appropriate algorithms (for example Branch & Bound) to solve the following ILP problem:

\[
\text{Minimize: } X_a + X_b + X_c + X_d \quad \text{(objective function representing the number of quadrants)}
\]

\[
\text{Subject to constraints: } 
\begin{align*}
X_a + X_b &\geq 1 \quad \text{(Species A must be present at least once)} \\
X_a + X_c &\geq 1 \quad \text{(Species B must be present at least once)} \\
X_a + X_b + X_c + X_d &\geq 1 \quad \text{(The species C must be present at least once)} \\
X_c + X_d &\geq 1 \quad \text{(The species D must be present at least once)}
\end{align*}
\]

With:
- \(X_a, X_b, X_c\) and \(X_d\) all Boolean variables

We will not explain how complex algorithms work to solve ILP-type problems, but, given the small size of the problem brought as an example, we can solve it with a simple reasoning.

First of all, let’s observe that no quadrant alone solves the problem, because no quadrant contains all the species. There are 6 couples of quadrants (\(ab, ac, ad, bc, bd\) and \(cd\); 3 couples \((ab, bc, cd)\) do not cover all species, while the other 3 couples \((ac, ad, bd)\) cover them all. Therefore 2 is the minimum number of quadrants needed to cover all species, and there are 3 and only 3 pairs of quadrants that meet all the constraints.

At this point we can formulate the

**Criterion 4**

If we arrive at a situation in which there are neither single-quadrant species nor mono-species quadrants, the residual problem (“hard core”) must be solved using an ILP algorithm.

It should be noted at this point that the application of the ILP algorithm provides only one solution of the “hard core”. But we will see later that there is a way to get all possible solutions by repeatedly applying the ILP to a slightly modified “hard core” problem.

Thus, Criterion 1 can apply to all single-quadrant species of a given matrix in a single step. We will call this passage “odd step” (because we do it in step 1, in step 3, in step 5, etc.). Similarly, Criteria 2 and 3 are applicable to several single-species quadrants in a single step. We will call this step “even step” (as we do it in step 2, in step 4, in step 6, etc.)
The OMB OSS (Optimization Method Based On Singleton Species) method

Definitely:

The whole process can be represented by a sequence of “odd steps” alternating with “even steps” (until neither of the two steps produces further simplifications) followed, if needed, by one or more “ILP steps”.

The following figure (Fig. 5.2) schematically illustrates the simplification/optimization process followed by the OMB OSS method.

To conclude, we note that:

- During the simplification steps (“even” and “odd”), to each quadrant that is identified as being part of the optimal solution (necessarily or as an alternative to others) is associated a species (or more than one) that represents “the reason” why the quadrant has been selected. This aspect is important because it allows to understand also from a botanical point of view the solution that is gradually built.
- More in particular:
  1. to each quadrant identified with Criterion 1 as part of the optimal solution are associated all the mono-quadrant species contained in it: the quadrant is essential, in the optimal solution, for the coverage of these species; the number of single-quadrant species in each quadrant identified by Criterion 1 is a good indicator of the “importance” of the quadrant for the purposes of bioconservation and therefore can be used as prioritization criterion;
  2. to each quadrant identified with Criterion 3 as part of the optimal solution is associated the single species it contains; also in this case the quadrant is essential in the optimal solution for covering that species.

SEARCH FOR ALL OPTIMAL SOLUTIONS

We have already seen that application of Criterion 3 can lead to identify alternatives to the single optimal solution that is obtained at the end of the process described.

Finding all the best solutions can be useful for many purposes; we mention some of them by way of example:

1. Evaluate for each quadrant it is “irreplaceable” in the set of optimal solutions; for example, an “irreplaceable index” can be calculated for each quadrant as the ratio between the number of solutions in which that quadrant is present and the total number of solutions. It will therefore be noticed that some quadrants are 100% irreplaceable and some have lower irreplaceable values because they can be replaced by other quadrants.
2. Compare the existing florist protection areas with those that come out from optimizations.
3. Perform a second-level optimization by choosing an optimal solution (among all optimal solutions) on the basis of a different criterion; for example select the optimal solution in which certain species are more covered, or the one in which the average number of species per quadrant is higher, etc. The topic will be taken over later.
4. Identify among the best solutions those in which the quadrants are less scattered. Also this will be discussed later.
5. Select the optimal solution, based on the type of territory it includes.
6. Provide for each optimal solution a set of indicators suitable to direct a “reasoned” choice, without losing the optimality. It therefore makes sense to try to identify all the possible optimal solutions.

We already know that the application of criterion 3 makes it possible to identify which quadrants can be replaced by which other ones. But if it is necessary to solve the “hard core” with an ILP algorithm, which notoriously provides only one optimal solution.

1. how do you check if there are other (always optimal) solutions and eventually identify them all?
2. and how do these solutions combine with those identified in the simplification steps?

The answer to the first question can be obtained simply by introducing in the ILP problem a new constraint that excludes the solution just found. In the previous example, if we assume that the first application of the ILP algorithm has supplied the \( a+c \) solution, it will be sufficient to add the following constraint to the previously seen constraints: \( X_a + X_c \leq 1 \) (which requires that quadrants \( a \) and \( c \) cannot be chosen together).

And run the ILP algorithm again to find a new solution. So:

To find all the optimal solutions of the “hard core” it is sufficient that every time an optimal solution is identified a new problem of ILP identical to the previous one is solved, except for the introduction of a further constraint that excludes the solution just found. This constraint will take the form:

\[
\sum_{i \in S} X_i \leq Q-1
\]

where \( S \) is the set of quadrants indexes present in the last optimal solution found, and \( Q \) is the total number of quadrants present in the optimal solution. The process ends when one of the following two conditions occurs:

1) The ILP algorithm no longer finds any solution.
2) The solution found has a number of quadrants higher than \( Q \).
The constraint introduced, at every step, excludes from the set of "optimal" solutions only the optimal solution found as last. The repeated application of the ILP algorithm, with a number of constraints that increases by 1 at each iteration, will lead to identify, one at a time, other optimal solutions. Finally, we note that if the process is not stopped when condition 2) occurs, solutions with Q + 1 quadrants would be obtained, then solutions with Q + 2 quadrants, and so on, until condition 1 would occur. To answer the second question, let's start observing that while the application of Criterion 1 allows to identify only "quadrants that cannot be replaced by any other", Criterion 3 leads to identify alternatives of the type "the quadrant i can be replaced by the quadrant j, the quadrant k, etc." and the ILP step identifies alternatives of the type "the set of Q quadrants S1 can be replaced by the set S2 (also of Q quadrants) or the set S3, etc.".

We can therefore conceptually divide the quadrants of each optimal solution into 3 sets:

1. The set of "certainly irreplaceable" quadrants; derive from the application of Criterion 1, and, by construction, this set contains quadrants that are certainly present in all the optimal solutions (100% irreplaceable);

2. The set of "certainly replaceable" quadrants; they are the result of the application of Criterion 3; note that each quadrant of this set is part of a group of interchangeable quadrants. Combining in every possible way a quadrant for each group, you will get a number of alternatives equal to the product of the cardinality of the individual subsets. For example, if the application of Criterion 3 leads to identify a quadrant with 2 alternatives, a second quadrant with 1 alternative and a third quadrant with 4 alternatives, the cardinality of the 3 groups will be 3, 2 and 5 respectively; as a result, 3 x 2 x 5 = 30 different combinations will be created. Moreover each quadrant of the first subset will be present in 1/3 of the 30 combinations (i.e. it will be irreplaceable at 33%), each quadrant of the second subset will be present in half of the 30 combinations (irreplaceable at 50%) and each quadrant of the third subset will be present in 1/5 of the 30 combinations (irreplaceable at 20%);

3. The set of quadrants deriving from the application of the ILP; this set will consist of as many groups of alternative quadrants as the ILP iterations have been; the cardinality of these alternative sets is constant. Unlike the quadrant groups of point 2 (which must be combined with each other since in each group a single quadrant is alternative to all the others), in this case every set of quadrants, generated by ILP, is an alternative to the other sets generated by ILP. Therefore, substitutability refers to sets of quadrants (equally numerous) and not to single quadrants. The irreplaceability of each quadrant present in an ILP solution can be calculated (as for all quadrants) by comparing the number of its occurrences in the various optimal solutions to the total number of optimal solutions.

At this point:

To build all the optimal solutions it is enough to combine in all possible ways:
- all the quadrants of point 1.
- each of the combinations of point 2.
- each of the sets of point 3.

Ultimately, will be obtained a number of optimal solutions equal to the product of the number of combinations generated in point 2 and the number of optimal solutions generated by ILP (point 3).

The way in which the best solutions are built (see last box) makes easy to find a solution that optimizes a second criterion for a wide range of problems: those in which the contribution of each quadrant to the optimization criterion does not depend on the contribution of the other quadrants. To give an example, the richness (number of species) of a quadrant does not depend (from a mathematical point of view) by that of the other quadrants, in the sense that to measure it, it is sufficient to measure the richness of each single quadrant independently of the others. In this situation, it will be possible to find the solution that maximizes the "average richness" with ease, i.e. without necessarily having to generate all the best possible solutions (which could be millions).

**CHOICE OF THE SOLUTION**

It is clear at this point that we need to find a criterion for choosing a solution among all those found. We propose below three possible criteria (without pretending to be exhaustive).

**Maximization of “average richness”**
(i.e. find the solution with greatest number of species per quadrant)

Since our optimal solutions all have the same number of quadrants, we can easily reason on the "total richness" (i.e. the sum of the richness of the quadrants that compose it) rather than on the "average richness". This means that we will maximize the overall species presence. It is easy to identify, among the optimal solutions identified, the one that has the greatest total richness. This solution can be easily identified by taking:
- All irreplaceable quadrants (as obtained in odd steps)
- The quadrant with greater richness, for each set of inter-
changeable quadrants (as obtained in even steps)
- The set of quadrants, among those generated by the ILP step iteration, which has the highest total richness (also easily calculable).

Maximization of the “average connection” (i.e. find the solution with less dispersed quadrants)
We can define that two quadrants are “connected” if they have a common side (other definitions are possible, of course). With the definition adopted, each quadrant belonging to an optimal solution can be connected to a maximum of 4 quadrants. Since each connection is seen twice (one for each of the two connected quadrants) we can assign to each quadrant a connection value equal to the number of sides in common with other quadrants divided by two (therefore each connection will be worth 0.5). Given this definition, an indicator of the degree of “connection” of a set of quadrants will simply be the average of the connection values of the quadrants that are part of it. Also here, since all the solutions have the same number of quadrants, the “maximum number of total connections” can be used as the criterion of choice. Note that the average number of connections per quadrant varies from 0 to a limit of 2.

Identifying the solution with maximum number of connections is however much more complex than that with maximum total richness, as it involves the calculation of the number of connections for each quadrant in each of the alternative solutions (observe that the number of connections of a given quadrant cannot be calculated without knowing which other quadrant of the optimal solution is adjacent to given one).

Minimize “average criticality” (i.e. find the solution with lower mean of singleton species per quadrant)
It should be noted that every optimal solution consists of quadrants which all contain at least one singleton species. If, in fact, a solution contains some quadrants without singleton species, all species contained in these quadrants would, by definition, also be present elsewhere, and ultimately the quadrants without singleton could be eliminated. Consequently, one can choose, among alternative solutions, the one in which the average singleton number (or, which is the same, the total number of singleton) is minimal.

Again, obtaining the solution with the minimum number of singleton is rather complex, because it involves the singleton species identification for each quadrant in each of the alternative solutions.

DISCUSSION
Some points of what has been exposed so far require further study and may give rise to further research:
1. In the method we followed, the minimization of the number of quadrants (or, if desired, of the surface) to be protected has maximum priority; other criteria have a lower level of priority and can be taken into account when choosing an optimal solution among all the optimal. We think this, however, is not a major limitation in cases where it is used to support the decision to identify new areas to be protected, as in these cases the costs of protection depend significantly on the greater or lesser extent of the surface to be protected.
2. From a strictly mathematical point of view, we have not really generated all the best possible solutions; in defining Criterion 2, in fact, we have not taken into account that some of the mono-species quadrants that we eliminate could be an alternative to one of the multi-species quadrants that contain the species s. This is not correct from a mathematical point of view (the criterion for identifying alternatives should be greatly complicated), but it is, we believe, completely acceptable from the point of view of bioconservation.
3. We assumed that all quadrants have the same surface; for this reason we did not introduce any weight diversification between the various quadrants (i.e. the coefficients in the objective function are all equal to 1). But what would happen if we want to weigh the various quadrants (for example, using the cost that would have to be incurred to protect the flora)? The proposed method can possibly be modified to take this need into account. See note 2 in this regard.
4. We are not sure that the simplification criteria illustrated will always lead to significant reductions in the dimensions of the original problem in all real situations; however, in all the tests we have performed on our data (on subsets of quadrants) the trend of the distribution of the Number of Frames by Number of Species is always similar to that reported in Fig. 1 (See also Note 1). As a result of this, simplification has always been possible and has reduced the size dramatically, allowing to:
   a. get (on a PC with a software for nothing optimized) all the solutions in a matter of seconds;
   b. get, (in many cases), optimal solutions without having to perform the ILP step;
   c. apply ILP, when necessary, to problems that are always very small.
5. Consider the extreme (and only theoretical) case of a single quadrant in which all species are present; it is...
obvious that such a solution would not be acceptable from a floristic protection point of view, since the rarity of species (which would be all singleton) would seriously endanger its continuity over time. The method proposed, however, could be modified as to ensure that each species in the optimal solution is present at least in a predefined number of quadrants. See note 3 below.

6. The results of the OMBOSS method make it possible to obtain, in various ways, a prioritization of the quadrants belonging to the optimal solution; as an example you can order these quadrants based on:
   a. Species richness of each quadrant;
   b. Level of irreplaceability of each quadrant;
   c. Number of singleton species of each quadrant; this value can be taken as criticality level of the quadrant, because neglecting a quadrant with a high number of singleton would cause more damage than neglecting a quadrant with only one singleton species.

7. Each of the prioritizations referred to in point 6 then allows to associate to each quadrant a level of complementarity, measurable as the number of new species that the quadrant adds to the set of species contained in the set of quadrants that precede it in prioritization.

**SUMMARY OF RESULTS**

We provide below a table (Tab. 12) summarizing the results of the steps of the OMBOSS method as applied to protection of rare species of Lazio (Tab. 5.12).

**Step 1:** The table indicates that 9771 observations of 786 species, distributed in 540 quadrants, were taken into consideration. Of the 786 species, 138 (17%) were single-quadrant and concentrated in 76 (14%) quadrants (100% irreplaceable). The number of species covered by these quadrants is 734 (92%).

**Step 2:** After elimination of the 76 quadrants and the 734 species identified in previous step, 140 quadrants are still to be optimized (with a reduction of 74.1% compared to the initial 540), with 64 species (a reduction of 92% compared to initial 786) and 228 observations in total (-97%). The application of criteria 2 and 3 led to identification of 88 single-species quadrants, among which 7 selected for the optimal solution and 81 (88-7) discarded or set aside to give rise to 648 alternatives. In particular, the 648 alternatives come out from the following table (Tab. 5.13).

The 7 groups are those that contribute to the optimal solution (choosing 1 single quadrant for each group). The total number of alternative combinations to which the 7 groups give rise is 3x3x2x2x2x3x3 = 648. The irreplaceable column is calculated as 1 / number of quadrants of the group.

**Steps 3 to 10:** They lead to successive reductions in size of the residual problem and in 2 additional alternative solutions (1 group of 2 quadrants in step 6), until.

---

**Table 5.12 - Summary of the OMBOSS method application and results for 786 species regarded as subjected to conservation.**

**Table 5.13 - Groups of alternative solutions identified in step 2.**
**Step 11:** There are still 2 quadrants and 2 species without multiple single-quadrant species or single-species quadrants. The application of the ILP step leads to identify 2 alternatives for the last quadrant.

Ultimately: to cover all 786 rare species you need to select 107 quadrants (20%), and there are $648 \times 2 \times 2 = 2592$ combinations of 107 quadrants, each of which is optimal.

The map in Fig. 5.3 shows the quadrant distribution and the step order.

While we are aware that these results do not claim to be completely general, we believe we can draw some conclusions:

1. The OMBOSS method obtains from the first step drastic reductions to the size of the starting problem (in our case: reduction of the number of quadrants of about 74.1% and reduction of the number of species of about 92%).
2. That the contribution of the even steps is generally small (in terms of the number of quadrants identified), but such as to allow the next odd step to significantly increase the number of quadrants.
3. The reduction of the number of quadrants at the end of the simplification process (before the ILP step) is about 99.6% and that of number of species is about 99.7%; this means that without making excessive mistakes, the ILP step could also be avoided and a sub optimal solution consisting of all the quadrants chosen during the simplification plus all the quadrants (2) that were the object of the optimization with the ILP.
4. Contrary to what might be expected, the number of optimal solutions is decidedly high; this number of alternatives is useful because, as already noted, it allows to apply further selections with different criteria.
5. The calculation times for the application of the simplification criteria in the cases we tested was very short (some seconds).
6. Since the problem to be solved with ILP is very small (only 2 variables and 2 constraints) the calculation time for the application of the ILP and the identification of all the possible optimal solutions has also been very limited (some seconds).
7. 98 quadrants are irreplaceable, while the 2,592 alternative solutions are different from each other only for the 9 quadrants needed to reach 107. As a consequence it is to be expected that the variability of total richness, total connection and criticality total within the 2592 solutions.

**Fig. 5.3** - Quadrant distribution and step order in the Latium map for the matrix 540 quad. x 786 taxa. We can observe that already at the first step we find 71% of the quadrants of the optimal solution and 77.5% of the irreplaceable ones; 76 quadrants containing singletons are extracted at this step. The numbers correspond to the step in which the quadrant is included in the optimal solution.
is very limited. In fact it has been verified that:

a. the total richness varies from 3712 to 3808 (i.e. from 34.76 to 35.59 species present on average in each of the 107 quadrants versus 18.09 species present on average in the total 540 quadrants),

b. the total connection varies from 50 to 58 (i.e. from 0.47 to 0.54 average connections per quadrant versus 1.82 of the total 540 quadrants) and indicates that "on average" the quadrants are connected two by two.

c. the total criticality varies from 246 to 257 (i.e. from 2.30 to 2.40 singleton species on average per quadrant), compared to 138 (0.26 on average) in the whole of the 540 quadrants.

NOTES

Note 1: persistence of number of Quadrants by number of species
The trend reported in Fig. 1, which shows how many of quadrants in which species are present, is function of the species rarity; the tests we carried out show that it is particularly persistent; in fact it has always been detected against:

- Sub-assemblies of quadrants that contain all the “rare” species, i.e. species present in no more than a given number of quadrants (which is quite obvious, if we consider that this set contains “a priori” all the mono-quadrant species).

- Subset of quadrants that have richness above a given threshold.

- Subset of quadrants between given pairs of geographic coordinates.

- Subset of quadrants randomly extracted with preset probability (same for all).

- Subset of quadrants selected on the base of orographic homogeneity.

By way of example, we report the graphs relative to the trends detected (on a larger data set than that of the rare species only) by extracting all the quadrants with richness exceeding a threshold ranging from 100 to 600 species per quadrant (Fig. 5.4).

Contrary to what might be expected, it also emerges that as the minimum richness of the selected quadrants increases, the number of single-quadrant species also increases, from 227 (in the whole of all quadrants) to 305 (in quadrants with no less than 600 species), with obvious positive consequences on the effectiveness of simplifications.

Note 2: how to weigh quadrants differently
Apart from the obvious modification of the introduction of weights in the objective function (which would reflect on the ILP step), the simplification process would still be applicable, making small modifications to Criterion 2 and Criterion 3. In particular:

- Criterion 2 should eliminate a given mono-species quadrant only if it has a weight lower than that of all mono-species and multi-species quadrants covering the same species, and

- Criterion 3 should
  - maintain in the optimal solution only the quadrant with greatest weight
  - consider as alternatives those with weight equal to this, and
  - eliminate all those with a lower weight.

The rest of the method would remain unchanged. Obviously, since we have not carried out any tests in this sense, we do not have information on the effectiveness, in this scenario, of the proposed simplifications.

Note 3: how to impose a minimum on the presence of each species in the optimal solution
It is evident that the presence of single-quadrant species prevents to respect in full any objective of presence of each species greater than 1. If, for example, the constraint is that each species (to be adequately protected) must be present in at least 3 quadrants, it is not generally possible to find a solution, as we know that in general there are species present in 1 or 2 quadrants only. The constraint has to be a little less strong, and, since we cannot do otherwise, we have to also accept in the optimal solution quadrants containing species present only in 1 or 2 quadrants. To obtain this, a preliminary step is needed to identify these species and the related quadrants and eliminate them from the matrix. It has also to be taken into account that eliminating those quadrants has the consequence of eliminating some species that are neither mono-quadrant nor double-quadrant, and this has to be considered in the following steps. Simplification criteria and mathematical model have to be changed according to the goal. By example, in the mathematical model the known term of each constraint has to be modified at each step in order to take into account that many species are, maybe, covered 1, 2, 3 or more times by the quadrants already selected in the partial solution.
Fig. 5.4 - Number of Quadrants vs Species Richness (on a larger data set than that of the rare species only). Each graph shows only quadrants whose richness exceeds a threshold ranging from 100 to 600 species per quadrant.