

# Adaptation d'une méthode d'aide à la décision pour évaluer les variations de la qualité de l'information selon les périodes étudiées.

---

## INTRODUCTION

Dans le contexte d'incertitude décrit précédemment inhérent à la base de données, la nécessité de produire une image plus claire et globale de l'information collectée s'est ainsi rapidement fait sentir, dans le but de visualiser les variations dans le nombre de données et dans les niveaux de qualité et de fiabilité de l'information selon les périodes renseignées. Cette étape du travail était nécessaire non seulement pour le programme PaléoSyr, mais aussi pour l'aboutissement de cette recherche : comme nous l'avons précisé dans le chapitre précédent, notre démarche est double et se situe à la fois au niveau du traitement de la donnée archéologique « pour elle-même », et dans une mobilisation de ce traitement au service d'une étude des structures et des dynamiques spatiales des peuplements passés. Cependant, notre démarche –et la demande– étant avant tout méthodologiques et non thématiques, le choix d'une période et d'une zone géographique d'étude sur la multitude de possibilités offertes par la base de données n'allait pas de soi, et auraient été, s'ils n'avaient été nourris par l'étape à venir (c'est-à-dire la partie II de cette thèse), totalement arbitraires. Notre parti-pris fut ainsi d'aller du général au particulier, en nous intéressant tout d'abord à la base de données dans son intégralité, puis de resserrer au fur et à mesure de cette exploration, l'éventail des possibles qui s'offraient à nous, non selon des critères thématiques mais selon la plus ou moins grande adéquation des données à notre problématique « géographique et archéologique » (proposer une méthodologie d'analyse exploratoire des structures et des dynamiques spatiales des peuplements passés en contexte incertain), et des diverses échelles spatiales et temporelles qu'elles impliquent.

Ainsi confrontée à notre propre positionnement et à la masse d'informations de la base de données, c'est-à-dire, au besoin de description pour l'orientation d'un choix, l'idée de l'analyse multicritères, méthode d'aide à la décision, s'est rapidement imposée à nous : en orientant celle-ci vers une qualification de l'hétérogénéité de la qualité de l'information, celle-ci nous permettrait à la fois de proposer une image synthétique de cet aspect de la base de données, tout en mettant à

notre disposition les indications nécessaires au choix que nous avons à réaliser. Ce chapitre présente ainsi la mise en place et les aboutissements de cette première étape de l'étude.

## **1. LA QUALITE DE L'INFORMATION : UN CONCEPT AUX MULTIPLES FACETTES**

Définir ce que l'on entend par qualité de l'information peut s'avérer fort difficile étant donnée la diversité de dimensions que prend ce concept. De façon générale, nous souscrivons à la définition de Goodchild (2006): «Quality [...] is a measure of the difference between the data and the reality that they represent, and becomes poorer as the data and the corresponding reality diverge. Thus, if data are of poor quality, and tell us little about the geographic world, then they have little value».

Ainsi, dans le cas de données archéologiques et spatiales, évaluer la qualité de l'information est nécessaire afin de pouvoir mesurer notre «légitimité» à formuler certaines hypothèses et interpréter les résultats des analyses effectuées avec ces mêmes données au sujet des structures et des dynamiques spatiales passées. Evaluer la qualité des données nous permet de fixer les limites du cadre interprétatif : objectiver le cadre interprétatif et calibrer les questionnements en fonction des données disponibles. Les poids attribués aux données permettent d'estimer le risque qu'on prend en réalisant telle ou telle interprétation.

Cette procédure s'avère cependant complexe. Naumann et Rolker (2000) en expliquent les raisons: «1) IQ criteria are often of subjective nature and can therefore not be assessed automatically 2) Information sources usually are autonomous and often not publish useful (and possibly compromising quality metadata). Additionally, many sources take measures to hinder information quality assessment. 3) The enormous amount of data to be assessed impedes assessment of the entire formation set. Thus sampling techniques are often necessary which decrease the accuracy of the assessed scores. 4) Information from autonomous sources is subject to sometimes surprising changes in content and quality». Bovee et al. (2003) ajoutent: « Without clearly defined attributes and their relationships, we are not just unable to assess information quality; we may be unaware of the problem. We need to understand the attributes of information quality and to have a broadly applicable, meaningful way to combine evaluations of them into a single quality measure».

Si l'on se réfère à la description des données effectuée dans le chapitre 1, on constate que la qualité de notre information se décline en deux catégories majeures : la précision de l'information, et la fiabilité de l'information (Tableau 9).

**Tableau 9 - Précision et fiabilité de l'information de la base de données PaléoSyr**

| <b>Précision de l'information</b>                          | <b>Fiabilité de l'information</b>                  |
|--|--|
| Morphologie de l'UNAR                                      | Fiabilité de la datation de l'UNAR                 |
| Morphologie du site archéologique auquel appartient l'UNAR | Fiabilité de l'attribution chronologique de l'UNAR |
| Niveau de datation de l'UNAR                               |  |

L'enjeu de cette étape du raisonnement est ainsi de répondre aux objectifs énoncés ci-dessus, en tenant compte des diverses caractéristiques inhérentes à la base de données et aux données elles-mêmes, qui complexifient la manœuvre. Les questions qui se posent sont alors :

- Comment évaluer la qualité de l'information en prenant en compte toutes les variables permettant de la décrire ?
- Comment distinguer la fiabilité et la précision de l'information dans cette évaluation ?

Dans cette optique, nous pensons qu'adopter une logique multivaluée peut répondre à nos objectifs et aux caractéristiques de notre base de données. Une analyse multicritère a donc été testée.

## **2. LE CHOIX D'UNE UNITE TEMPORELLE DE REFERENCE : LE SIECLE**

Dans la mesure où il s'agit ici de décrire la qualité de l'information associée aux données de la base PaléoSyr, la question du mode de représentation du temps s'est rapidement posée.

Nous avons opté pour une déconstruction de la chronologie par périodes, et d'une discrétisation de celle-ci par tranches de 100 ans.

Il convient de préciser que ce choix ne relève pas d'un parti-pris thématique fort, mais de la nécessité de choisir une unité de référence continue, nous permettant de visualiser l'information de manière précise mais synthétique. Ce choix relève d'une volonté d'expérimentation en amont de notre étude, et pourra être remis en question par les résultats des premières analyses, et tout au long de l'étude. Nous nous inscrivons donc dans cette pensée de Daniel S. Milo (1992): « Si l'on admet que l'aspect expérimental du siècle réside dans son arbitraire, il faut assumer cet arbitraire ! En d'autres mots, on ne peut fort mal tenir une expérimentation, puis la délaissier car non-rentable. En d'autres mots encore, l'expérimentation séculaire n'a pas produit les résultats escomptés parce qu'on ne l'a pas comprise, donc on l'a mal gérée – du coup, on était incapable d'évaluer les résultats qu'on a quand même produits ».

En outre, 100 ans étant la durée minimale de certaines périodes de la chronologie –dont il est, nous l'avons souligné dans la première partie, extrêmement complexe de se défaire-, nous avons souhaité conserver cette unité de temps : ni trop large, pour ne pas regrouper plusieurs périodes, perdant ainsi la finesse de la datation ; ni trop petite, sachant qu'une datation plus fine que l'échelle du siècle n'a pas été possible pour la plupart des objets archéologiques étudiés –adopter une résolution de 50 ans ne ferait que créer une information extrêmement artificielle et alourdirait notre tâche.

### **3. ADAPTATION D'UNE METHODE D'AIDE A LA DECISION : L'ANALYSE MULTICRITERE**

#### **3.1 Etat de l'art et présentation générale de la méthode**

En utilisant l'analyse multicritère (AMC), nous pouvons attribuer à chaque période un niveau de qualité sous forme d'un « score », afin d'identifier celles qui sont le mieux pourvues en termes de qualité et de quantité d'information. Le but est ainsi de définir et d'adapter ce cadre conceptuel et méthodologique aux problématiques et aux bases de données archéologiques.

Diverses écoles ont émergé dans l'AMC depuis les années 1970. Parmi elles, Thomas L. Saaty, professeur à l'Université de Pittsburgh, est le créateur et le premier théoricien de l'« Analytic Hierarchy Process » qui est à la base de la méthode utilisée dans cette étude. Son objectif était de créer une approche qui permettrait de répondre au besoin de réaliser des comparaisons dans le cas où des échelles standard ne seraient pas disponibles pour mesurer les phénomènes de manière absolue, et de dériver des mesures de ces comparaisons (Saaty, 1990 ; 1993 ; 2008).

L'AMC est une méthode d'évaluation multidimensionnelle nous permettant de résoudre des problèmes de décision complexes. Cette approche se fonde sur le principe qu'une solution à un problème de décision peut être trouvée en le divisant en plusieurs petites parts plus facilement appréhendables, et en les intégrant de manière logique (Verhagen, 2006). L'intérêt de la méthode est double : «It provides an overall view of the complex relationships inherent in the situation; and helps the decision maker assess whether the issues in each level are of the same order of magnitude, so he can compare such homogeneous elements accurately» (Saaty 1990).

En contexte d'aide à la décision, l'une des tâches les plus complexes est de sélectionner les éléments importants à considérer dans la prise de décision. L'AMC nous fournit un moyen de classer ces éléments, une fois sélectionnés, dans une structure hiérarchique, allant d'un but général jusqu'à plusieurs critères, sous-critères et alternatives en niveaux successifs (Saaty, 1990). Cette méthode a été appliquée dans divers champs de recherche tels que le changement

climatique, la finance ou les risques, domaines où des buts multiples et conflictuels compliquent souvent la tâche des décideurs, afin que ceux-ci puissent bénéficier de résultats systématiques et concrets leur permettant d'adopter les politiques adaptées à la situation. Ainsi, l'objectif de l'AMC dans ce contexte n'est pas de calculer quelle est la « bonne » décision, mais de contribuer à améliorer la connaissance dans des situations impliquant des risques, plusieurs critères, et des intérêts multiples (Bell et al., 2003). Cette méthode a été rarement appliquée en archéologie, et lorsqu'elle l'a été, elle avait une fonction prédictive. Il s'agissait alors d'évaluer le potentiel archéologique d'une zone, les limites des zones de prospection et les procédures de planification (Verhagen, 2006, 2007 ; Krist, 2006). En outre, les archéologues ont exploité le potentiel de l'analyse multicritères afin de déterminer les surfaces-coûts dans l'exploration des paysages à échelle régionale, afin de construire et valider des modèles des développements sociaux passés, et prendre en compte les liens sociaux entre les groupes d'individus (Howey 2007).

Notre problème de décision est donc ici très différent des applications communes de l'AMC: nous souhaitons classer les diverses périodes identifiées dans la base de données en fonction de la qualité de l'information permettant de les décrire, par l'élaboration d'un certain nombre de critères pouvant permettre d'évaluer l'information archéologique. En d'autres termes, il s'agit d'identifier les siècles pour lesquels l'information disponible maximise la précision et la fiabilité des données, et minimise donc l'incertitude. Il convient de préciser que les critères, alternatives et points de vue déterminés dans cette étude et décrits plus loin ont été identifiés en fonction d'informations propres à notre base de données et à nos objectifs. La solidité et la flexibilité de la méthode assurent toutefois sa reproductibilité à diverses problématiques et bases de données.

## **3.2 Déroulement de la procédure d'analyse multicritère**

### *3.2.1 Etape 1: définir les objectifs*

Il s'agit tout d'abord de définir un ou plusieurs état(s) final(aux) souhaité(s). Le nôtre est de classer les périodes de la chronologie de notre base de données en fonction de la qualité de l'information qui les caractérisent (précision et fiabilité de l'information).

### *3.2.2 Etape 2: identifier les alternatives*

Les alternatives sont les diverses options que l'on souhaite classer, dans notre cas, les périodes de la chronologie. Ce qui implique de choisir un mode de division du temps pour cette étude, sachant que celles-ci n'ont pas toutes la même durée, et s'échelonnent en plusieurs sous-périodes (cf. chapitre 1).

Nous avons choisi ici de raisonner en dates calendaires plutôt qu'en périodes. Nous avons ainsi choisi de diviser la chronologie, et donc la base de données en intervalles de 100 ans, le siècle étant le niveau de précision -ou la durée- maximum atteint dans la base de données. Chaque intervalle de 100 ans est caractérisé par les données et leurs caractéristiques qui y correspondent, et considéré comme une entité à part entière.

Ainsi, nos alternatives, ou en d'autres termes, les éléments que l'on souhaite hiérarchiser, sont chacun des intervalles de 100 ans présents entre le début du Néolithique (-9600 av JC) jusqu'à aujourd'hui, ce qui nous donne un total de 116 alternatives.

Les alternatives sont notées  $i = 1, \dots, I$ .

### *3.2.3 Etape 3: définir et quantifier les critères*

L'objectif de cette étape est de définir un ensemble de critères concis à partir desquels l'information va être évaluée (Eppler, Wittig, 2000). Les critères sont donc les diverses caractéristiques de la qualité de l'information que l'on juge pertinentes (en théorie et en fonction de la base de données) pour analyser les alternatives. Ce sont les aspects en fonction desquels chaque alternative sera examinée et comparée aux autres. On ne raisonne pas ici en nombres absolus, mais en nombres relatifs. Saaty (1993) souligne cet aspect de la méthode : «Measurements on a scale are simply indicators. They must be interpreted in some human value system to make sense».

Il convient à présent de présenter les cinq critères que nous avons choisis pour leur aptitude, selon nous, à caractériser la qualité de l'information de la base de données, et donc hiérarchiser chacune des alternatives. Des échelles de mesure ont été définies afin d'évaluer la correspondance entre chaque critère et chaque alternative. Ces calculs ont été appliqués tout d'abord à chaque UNAR de la base de données, afin de situer chacune d'entre elles sur les échelles de valeurs définies pour chaque critère. Les valeurs de chaque UNAR pour chaque critère ont ensuite été additionnées en fonction de l'alternative –l'intervalle de 100 ans- à laquelle elles appartiennent. Les explications qui suivent montrent la procédure d'évaluation de chaque UNAR pour chaque critère défini.

- *Critère 1* : L'information dont on dispose à propos de la morphologie de l'UNAR, c'est-à-dire, notre niveau de connaissance sur le type d'UNAR en question.

Si la fonction de l'UNAR n'est pas déterminée dans la base de données, (c'est-à-dire : si l'UNAR est répertoriée comme une « occupation indéfinie » dans le Tableau 5), un pourcentage de

connaissance est défini en se basant sur la morphologie du site auquel l'UNAR est associée (Tableau 10).

**Tableau 10 - Probabilités (en %) pour que chaque type d'UNAR soit représenté dans un site en fonction de la morphologie de ce dernier (Braemer, Fusco)**

| Morpho1                    | Morpho2  | Bâtiment | Funéraire | Agro-pastoral | Hydraulique | Autre | Indéfini |
|----------------------------|--|----------|-----------|---------------|-------------|-------|----------|
| 1. Indice de site          | 1.01. Concentration matériel lithique          | 75%      | 0         | 0             | 0           | 0     | 25%      |
| 1. Indice de site          | 1.02. Concentration matériel céramique         | 33%      | 33%       | 33%           | 0           | 0     |          |
| 1. Indice de site          | 1.03. Groupe d'inscriptions (2 et plus)        | 50%      | 50%       | 0             | 0           | 0     | 0        |
| 1. Indice de site          | 1.04. Eléments architecturaux (2 et plus)      | 75%      | 25%       | 0             | 0           | 0     | 0        |
| 1. Indice de site          | 1.05. Différence de couleur de sol             | 0        | 0         | 0             | 0           | 0     | 100%     |
| 1. Indice de site          | 1.06 Faible relief                             | 0        | 0         | 0             | 0           | 0     | 100%     |
| 2. Terre dominante         | 2.01. Tell haut ( + de 5 m de haut)            | 80%      | 20%       | 0             | 0           | 0     | 0        |
| 2. Terre dominante         | 2.02. Tell + ville basse                       | 80%      | 20%       | 0             | 0           | 0     | 0        |
| 2. Terre dominante         | 2.03. Tell bas (moins de 5m de haut)           | 80%      | 20%       | 0             | 0           | 0     | 0        |
| 2. Terre dominante         | 2.04. Ensemble de buttes basses                | 80%      | 20%       | 0             | 0           | 0     | 0        |
| 2. Terre dominante         | 2.07. Tell + site au pied                      | 80%      | 20%       | 0             | 0           | 0     | 0        |
| 2. Terre dominante         | 2.06. Tell                                     | 80%      | 20%       | 0             | 0           | 0     | 0        |
| 2. Terre dominante         | 2.05. Tell à structure multiples et distinctes | 80%      | 20%       | 0             | 0           | 0     | 0        |
| 3. Pierre dominante        | 3.07. Tell + site au pied                      | 80%      | 20%       | 0             | 0           | 0     | 0        |
| 3. Pierre dominante        | 3.01. Tell haut ( + de 5 m de haut)            | 80%      | 20%       | 0             | 0           | 0     | 0        |
| 3. Pierre dominante        | 3.05. Kôm, rujum, tas de cailloux              | 0        | 75%       | 25%           | 0           | 0     | 0        |
| 3. Pierre dominante        | 3.03. Murs conservés hors sol                  | 80%      | 20%       | 0             | 0           | 0     | 0        |
| 3. Pierre dominante        | 3.04. Murs arasés                              | 80%      | 20%       | 0             | 0           | 0     | 0        |
| 3. Pierre dominante        | 3.02. Tell bas                                 | 80%      | 20%       | 0             | 0           | 0     | 0        |
| 3. Pierre dominante        | 3.06. Tell                                     | 80%      | 20%       | 0             | 0           | 0     | 0        |
| 4. Matériaux mixtes        | 4.01. Tell haut ( + de 5 m de haut)            | 80%      | 20%       | 0             | 0           | 0     | 0        |
| 4. Matériaux mixtes        | 4.02. Tell bas (moins de 5m de haut)           | 80%      | 20%       | 0             | 0           | 0     | 0        |
| 4. Matériaux mixtes        | 4.03. Murs visibles, arasés, hors sol          | 80%      | 20%       | 0             | 0           | 0     | 0        |
| 4. Matériaux mixtes        | 4.04. Ensemble de buttes basses                | 80%      | 20%       | 0             | 0           | 0     | 0        |
| 5. Cavité /excavation      | 5.01. Abri(s) sous roche                       | 80%      | 20%       | 0             | 0           | 0     | 0        |
| 5. Cavité /excavation      | 5.02. Grotte                                   | 80%      | 20%       | 0             | 0           | 0     | 0        |
| 5. Cavité /excavation      | 5.03. Excavation                               | 80%      | 20%       | 0             | 0           | 0     | 0        |
| 5. Cavité /excavation      | 5.04. Nécropole isolée (tombe à fosse)         | 0        | 100%      | 0             | 0           | 0     | 0        |
| 6. Hydraulique hors site   | 6.00. Hydraulique hors site                    | 0        | 0         | 0             | 100%        | 0     | 0        |
| 7. Voies et communications | 7.00. Voies et communications                  | 0        | 0         | 0             | 0           | 100%  | 0        |
| 8. Trame parcellaire       | 8.01. Trame parcellaire                        | 0        | 0         | 100%          | 0           | 0     | 0        |
| 8. Trame parcellaire       | 8.02. Terrasses                                | 0        | 0         | 100%          | 0           | 0     | 0        |

Par exemple, si le site archéologique est répertorié comme étant un Tell, nous savons que l'UNAR en présence pourrait avoir eu, avec une probabilité égale, une fonction d'habitat ou

funéraire. Une valeur faible sera donc attribuée à cette UNAR dans l'échelle de mesure du critère, le niveau d'incertitude étant ici élevé.

Le Tableau 10, croisant les différents types de morphologie des sites archéologiques et les probabilités pour que chaque type d'UNAR y soit représenté a donc été créé, en se basant sur la connaissance de terrain des archéologues du projet.

A partir de ce tableau, des indices ont été créés afin d'attribuer à chaque UNAR la valeur correspondant à sa représentation dans le critère 1 (Tableau 11). Chaque indice représente ainsi notre niveau de connaissance sur la morphologie de chaque UNAR.

**Tableau 11 - Indice attribué à chaque UNAR en fonction du degré de connaissance disponible sur sa morphologie**

| <b>Pourcentage de connaissance sur la morphologie de l'UNAR (cf. Tableau 10)</b> | <b>Indice attribué à l'UNAR</b> |
|--|---------------------------------|
| 100% d'une catégorie (sauf indéfini)   | 20                              |
| 80% d'une catégorie, 20% d'une autre   | 10                              |
| 75% d'une catégorie, 25% d'une autre   | 8                               |
| 50% pour deux catégories   | 6                               |
| 25% d'une catégorie, 75% d'indéfini  | 5                               |
| 33% pour trois catégories  | 4                               |
| 100% d'indéfini  | 0                               |

Cette approche, malgré ses failles (impossibilité de détecter les exceptions, indices et pourcentages attribués de manière empirique) permet toutefois de dépasser une approche binaire connaissance/absence de connaissance, en distinguant les UNAR qui pourraient avoir n'importe quelle fonction, et celles pour lesquelles les possibilités sont plus restreintes (et donc, l'incertitude plus faible). Il va de soi que les pourcentages attribués ainsi que les indices peuvent être discutés : le but n'est pas ici d'obtenir une information certaine, mais de proposer une démarche expérimentale, un des chemins possibles pour prendre en compte l'hétérogénéité de l'information.

La Figure 16 illustre la manière dont chaque UNAR, puis chaque alternative est quantifiée pour le critère 1.

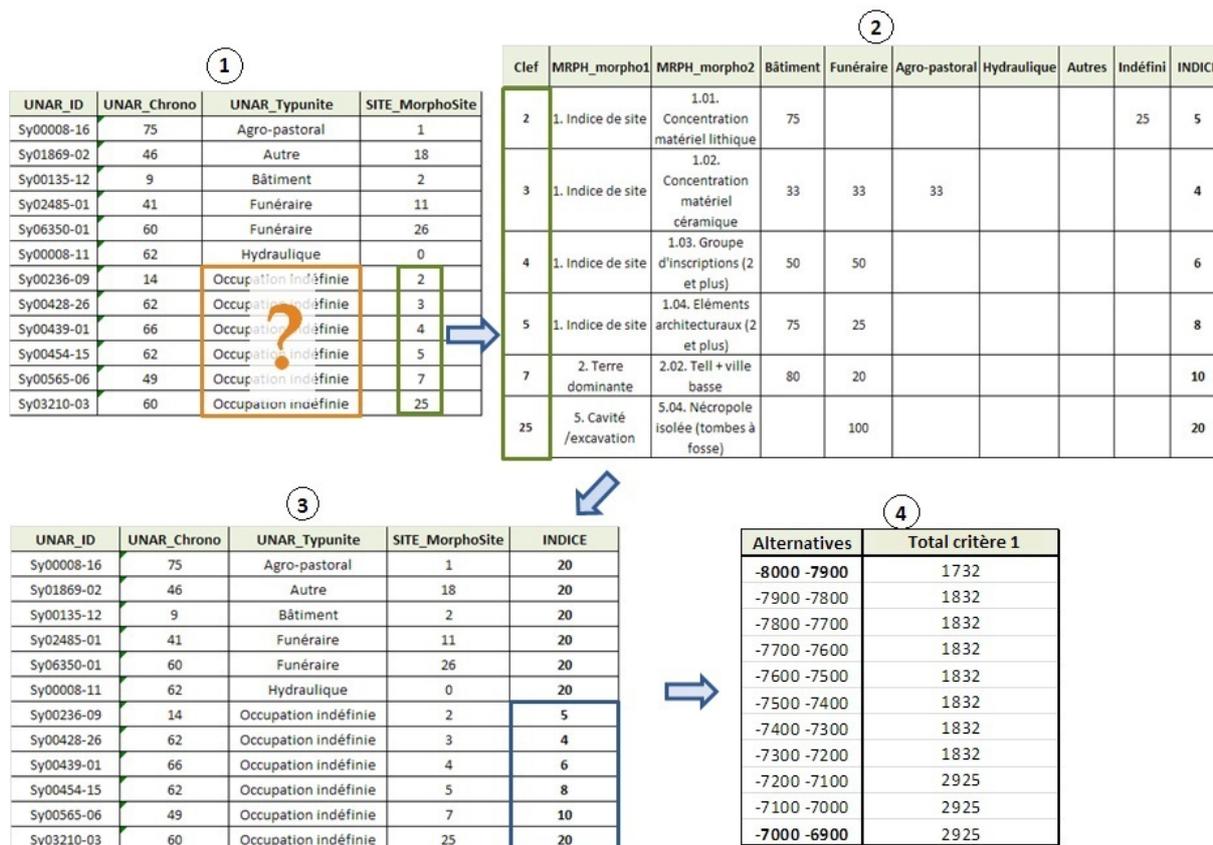


Figure 16 - Critère n°1: des UNAR aux alternatives.

Le passage de l'étape 3 à l'étape 4 sur la Figure 16 (c'est-à-dire, le passage des UNAR et de leurs indices de connaissance à chaque alternative) est décrit sur le Tableau 12.

Tableau 12 - Calcul de la valeur représentant le critère n°1 pour chaque alternative

| Alternative   | Indices de connaissance                                  |   |    |   |   |    |    | Total critère 1                           |
|---------------|--|---|----|---|---|----|----|---|
|               | 0  | 4 | 5  | 6 | 8 | 10 | 20 |   |
| -15000 -14900 | 0  | 0 | 5  | 0 | 0 | 16 | 17 | $(5*5) + (16*10) + (17*20) = 525$         |
| ...           | Nombre d'UNAR de l'alternative représentées par l'indice |   |    |   |   |    |    | ...                                       |
| -7100 -7000   | 0  | 1 | 27 | 0 | 7 | 79 | 97 | $(1*4) + (27*5) + (7*8) + (79*10) = 2925$ |

- *Critère 2* : La fiabilité de la datation de l'UNAR. Ce critère permet d'évaluer la fiabilité de la datation des UNAR (et des alternatives). Un indice a été attribué à chaque niveau de fiabilité des UNAR (Tableau 13).

**Tableau 13 - Indices de fiabilité de la datation des UNAR**

| Fiabilité de la datation de l'UNAR | Indice |
|------------------------------------|--------|
| Fiable                             | 3      |
| Incertain                          | 2      |
| Douteux                            | 1      |
| Non connue                         | 0      |

Le total des indices de fiabilité de la datation des UNAR pour chaque alternative a été calculé comme pour le critère n°1 (Tableau 14).

**Tableau 14 - Calcul de la valeur représentant le critère n°2 pour chaque alternative**

| Alternative   | Fiabilité de la datation                                 |   |    |     | Total critère 2          |
|---------------|--|---|----|-----|--------------------------|
|               | 0  | 1 | 2  | 3   |                          |
| -15000 -14900 | 1  | 0 | 20 | 24  | $(20*2) + (24*3) = 112$  |
| ...           | Nombre d'UNAR de l'alternative représentées par l'indice |   |    |     | ...                      |
| -7100 -7000   | 4  | 0 | 84 | 126 | $(84*2) + (126*3) = 546$ |

- *Critère 3* : la fiabilité de l'attribution chronologique de l'UNAR : Ce critère permet d'évaluer la fiabilité de l'attribution chronologique des UNAR (et des alternatives). Un indice a été attribué à chaque niveau de fiabilité des UNAR (Tableau 15).

**Tableau 15 - Indices de fiabilité de la datation des UNAR**

| Fiabilité de l'attribution chronologique | Indice |
|--|--------|
| Fiable                                   | 9      |
| Incertain                                | 6      |
| Douteux                                  | 3      |
| Non connue                               | 0      |

Le total des indices de fiabilité de l'attribution chronologique des UNAR pour chaque alternative a été calculé comme pour le critère n° 2 (cf. Tableau 14).

- *Critère 4* : Le nombre d'UNAR « indéterminées » dans l'alternative. Ce critère permet de classer les siècles (alternatives) en fonction du nombre d'UNAR dites « indéterminées » (cf. tableau 5) dans leur attribution à une période en particulier, et celles qui sont datées avec plus de précision.

On calcule ici un ratio en divisant le nombre d'UNAR « déterminées » par le nombre d'UNAR « indéterminées » pour chaque tranche de 100 ans. En outre, une contrainte supplémentaire a été introduite dans l'élaboration de ce critère : la chronologie générale de la base de données étant déclinée en 4 niveaux de précision sous forme de sous-périodes (cf. tableau 6), nous avons considéré dans notre calcul que si l'indétermination de l'UNAR est survenue au troisième niveau de datation, l'incertitude sera moins élevée que si elle survient au second. Ainsi, le but étant de détecter les alternatives les plus fiables et précises, ici plus le ratio calculé est élevé, plus le score de l'alternative sera bon, et plus celle-ci sera privilégiée. La procédure de calcul est décrite sur le Tableau 16.

Tableau 16 - Calcul des valeurs du critère 4 pour chaque alternative

| Alternatives  | Niveau de datation 2 |                    | Niveau de datation 3 |                    | Niveau de datation 4 |                    |
|---------------|----------------------|--------------------|----------------------|--------------------|----------------------|--------------------|
|               | UNAR déterminées     | UNAR indéterminées | UNAR déterminées     | UNAR indéterminées | UNAR déterminées     | UNAR indéterminées |
| -15000 -14900 | 0                    | 2                  | 35                   | 8                  | 50                   | 2                  |
| -7100 -7000   | 0                    | 73                 | 26                   | 131                | 0                    | 0                  |



| Alternatives  | Indices des niveaux de datation |                |             | INDICE                     |
|---------------|---------------------------------|----------------|-------------|----------------------------|
|               | 1                               | 3              | 6           |                            |
| Alternatives  | Ratio N2                        | Ratio N3       | Ratio N4    |                            |
| -15000 -14900 | 0                               | $35/8 = 4,3$   | $50/2 = 25$ | $(4,3*3) + (25*6) = 227,4$ |
| -7100 -7000   | 0                               | $26/131 = 0,2$ | 0,000       | $0,2*3 = 0,6$              |

- *Critère 5* : Le niveau de datation maximum atteint pour l'UNAR. Ce critère reprend une information déjà impliquée dans la construction du critère 4, mais l'intègre cette fois-ci non plus comme un facteur de pondération du critère mais comme un critère déterminant à part entière. Le niveau atteint dans la datation de l'UNAR est ainsi considéré comme un niveau de précision. Aucun calcul n'a été nécessaire pour l'élaboration des valeurs de ce critère, qui sont tout simplement les niveaux de datation 2,3 et 4.

Cinq critères ont ainsi été élaborés, et les valeurs qu'ils prennent pour chaque alternative, et qui permettront de hiérarchiser celles-ci ont été calculées. L'information créée est synthétisée dans le Tableau 17.

Tableau 17 – Synthèse : signification et procédure de création de chaque critère

|                  | Définition   | Méthode de calcul   | Echelle  |
|------------------|--|---|--|
| <b>Critère 1</b> | Information sur la morphologie de l'UNAR.          | <ul style="list-style-type: none"> <li>- Si UNAR = «Agro-pastoral», «funéraire», «bâtiment», «hydraulique», «autre» → 100% de connaissance;</li> <li>- Si UNAR = «indéterminée» → on se réfère à la morphologie du site auquel elle appartient</li> </ul> | De 0 à 20  |
| <b>Critère 2</b> | Fiabilité de la datation de l'UNAR.                | Information déjà disponible dans la base de données.  | <b>0</b> = Non renseigné<br><b>1</b> = Douteux;<br><b>2</b> = Incertain;<br><b>3</b> = Fiable.   |
| <b>Critère 3</b> | Fiabilité de l'attribution chronologique de l'UNAR | Information déjà disponible dans la base de données.  | <b>0</b> = Non renseigné;<br><b>3</b> = Douteux ;<br><b>6</b> = Incertain;<br><b>9</b> = Fiable. |
| <b>Critère 4</b> | Nombre d'UNAR indéterminées                        | (nombre d'UNAR déterminées / nombre d'UNAR indéterminées) x niveau de datation maximum atteint  | Plus le résultat est élevé, plus l'incertitude est faible.                                       |
| <b>Critère 5</b> | Niveau de datation maximum atteint                 | Information déjà disponible dans la base de données.  | De 2 à 4.  |

Les critères sont notés  $j = 1, \dots, J$ .

#### 3.2.4 Etape 4: définir les points de vue et les poids des critères

L'AMC a pour particularité de permettre la définition de divers sous-objectifs, appelés les points de vue. Ces éléments permettent d'implémenter diverses règles de décision, en accordant des poids différents aux divers critères en fonction du point de vue que l'on souhaite examiner : «Broadly speaking, the purpose of criterion (...) weights is to express the importance of each criterion relative to other criteria» (Malczewski 1999).

Nous avons identifié deux points de vue, en fonction de la dimension de la qualité d'information que l'on souhaite mettre en relief: le point de vue de la précision, et celui de la fiabilité de l'information.

Les points de vue sont notés  $v = 1, \dots, V$

Nous avons à présent déterminé un ensemble d'alternatives, de critères, de points de vue et leurs poids respectifs. Les données peuvent alors être organisées sous la forme de deux matrices de décision :

- La matrice d'impacts (Tableau 18) notée E possède I lignes et J colonnes ; les éléments e (i,j) représentent les valeurs attribuées à l'alternative i évaluée en fonction du critère j.

$$E = \begin{pmatrix} e(1,1) & \cdots & e(1,J) \\ \vdots & & \\ e(I,1) & \cdots & e(I,J) \end{pmatrix}$$

Tableau 18 - Extrait de la matrice d'impacts

| Alternatives | Critère 1 | Critère 2 | Critère 3 | Critère 4 | Critère 5 |
|--------------|-----------|-----------|-----------|-----------|-----------|
| -2200 -2100  | 16893     | 2652      | 6349      | 1,21      | 4         |
| -2100 -2000  | 16893     | 2652      | 6349      | 1,21      | 4         |
| -2000 -1900  | 8748      | 1462      | 3149      | 2,43      | 3         |
| -1900 -1800  | 10538     | 1749      | 3777      | 2,43      | 3         |

- La matrice des points de vue (Tableau 19) notée W (est composée de j lignes et v colonnes. Les éléments p (j, v) représentent le poids attribué au critère j du point de vue v. Le poids est intuitivement déterminé par l'utilisateur, en fonction de sa connaissance de la situation décrite.

$$W = \begin{pmatrix} p(1,1) & \cdots & p(1,V) \\ \vdots & & \vdots \\ p(J,1) & \cdots & p(J,V) \end{pmatrix}$$

Tableau 19 - Matrice des points de vue

|           | Fiabilité | Précision |
|-----------|-----------|-----------|
| Critère 1 | 5         | 5         |
| Critère 2 | 8         | 8         |
| Critère 3 | 10        | 8         |
| Critère 4 | 5         | 10        |
| Critère 5 | 5         | 10        |

La Figure 17 récapitule les premières étapes réalisées dans l'AMC.

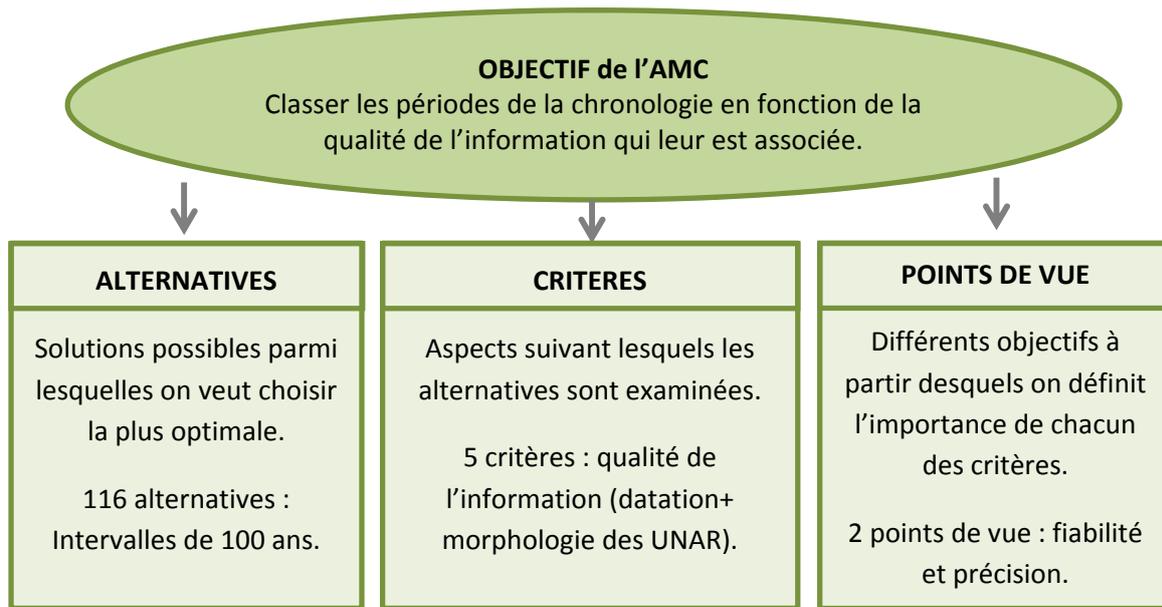


Figure 17 - Récapitulatif des premières étapes de l'AMC

### 3.2.5 Etape 5: calcul et hiérarchisation des scores de chaque alternative

Le but de l'AMC est d'évaluer et classer chacune des alternatives  $i$  par chacun des points de vue  $v$ , afin d'obtenir un score final  $s(i, v)$ . Notre méthode de calcul implique que les valeurs les plus élevées représentent les alternatives optimales, et les critères les plus importants.

Les AMC ici présentées ont été réalisées à l'aide de MULTI, un programme développé par le Politecnico di Milano (département Architecture et Aménagement).

Les éléments de la matrice d'impacts ne sont souvent pas directement comparables s'ils ont été obtenus, comme c'est le cas ici, par des calculs différents impliquant des échelles différentes. Les valeurs assignées aux alternatives  $i$  en fonction d'un critère  $j$  n'ont qu'une valeur intrinsèque en relation avec ce critère en particulier. La normalisation des valeurs est donc nécessaire afin de rendre les données comparables entre elles.

Dans la formule qui suit,  $e'(i, j)$  renvoie à l'élément normalisé obtenu à partir des données initiales  $e(i, j)$ . Nous avons choisi la méthode de normalisation zéro-max (proposée dans le programme MULTI), afin de ne pas perdre d'importantes informations sur les valeurs de la meilleure et de la moins bonne alternative pour chaque critère. Dans cette méthode, on divise les éléments de la colonne  $j$  par le plus élevé d'entre eux ( $e_{\max}(j)$ ):

$$e'(i, j) = \frac{e(i, j)}{e_{\max}(j)}$$

Par cette méthode, nous obtenons des valeurs normalisées par les valeurs de  $e(i, j)$  comprises entre  $e_{\min}(j)/e_{\max}(j)$  et 1, 1 demeurant la valeur maximale. Ainsi, les rapports entre  $e(i, j)$  demeurent inaltérés.

Une fois les données normalisées, nous avons déterminé la valeur des alternatives pour chaque point de vue, et de la sorte, un classement final. Le programme MULTI nous offre plusieurs méthodes de calcul possibles, que nous avons choisi de tester afin d'évaluer leur impact sur le résultat final.

- La somme pondérée: il s'agit de la méthode la plus simple à effectuer. Elle implique les deux matrices  $S = E * W$ , où les scores  $s(i, v)$  sont calculés en multipliant les deux matrices ligne par colonne :

$$s(i,v) = \sum_j e(i, j) * p(j, v)$$

- L'analyse de la concordance et de la discordance : cette méthode de calcul confronte les alternatives avec toutes les autres prises individuellement. Pour chaque couple d'alternatives confrontées, la méthode analyse combien de critères prévalent pour l'alternative étudiée plutôt que pour les autres (analyse de la concordance) et calcule l'effet global de séparation, c'est-à-dire le point où chaque critère distingue l'alternative étudiée des autres (analyse de la discordance).

#### **4. PRESENTATION ET INTERPRETATION DES RESULTATS**

##### **4.1 Application n°1 de l'AMC : à la base de données entière, depuis le début du Néolithique jusqu'à aujourd'hui**

Afin de tester la sensibilité des résultats aux paramètres de calcul, ainsi que la robustesse du classement final, nous avons appliqué tour à tour les deux méthodes expliquées ci-dessus.

La Figure 18 montre les résultats des quatre AMC réalisées, replacés sur la séquence chronologique de la base de données. Les courbes bleue et rouge représentent les scores de chaque alternative en fonction des deux points de vue identifiés (précision et fiabilité de la donnée). La courbe noire indique le nombre de données (le nombre d'UNAR) de chaque alternative, nous ayant permis de calculer l'AMC. Cette courbe nous rappelle de ne pas mésestimer le poids du nombre de données dans notre analyse : certaines alternatives comprennent en effet une centaine d'UNAR, tandis que d'autres en contiennent deux mille. Les fluctuations des résultats des AMC peuvent ainsi en être grandement influencées.

Afin de diminuer le poids du nombre de données dans l'analyse, deux des AMC ont été calculées en divisant chaque valeur de la matrice d'impacts par le nombre de données total qui la composent. Le Tableau 20 récapitule ainsi les méthodes de calcul et les données utilisées pour les quatre AMC réalisées.

**Tableau 20 - Données et méthodes utilisées pour chaque AMC**

|                          | <b>AMC 1</b>      | <b>AMC 2</b>                      | <b>AMC 3</b>                                   | <b>AMC 4</b>                                   |
|--------------------------|-------------------|-----------------------------------|--|--|
| <b>Données utilisées</b> | Valeur du critère | Valeur du critère / nombre d'UNAR | Valeur du critère                              | Valeur du critère / nombre d'UNAR              |
| <b>Méthode de calcul</b> | Somme pondérée    | Somme pondérée                    | Analyse de la concordance et de la discordance | Analyse de la concordance et de la discordance |

Chapitre 4 : Adaptation d'une méthode d'aide à la décision pour évaluer les variations de la qualité de l'information selon les périodes étudiées.

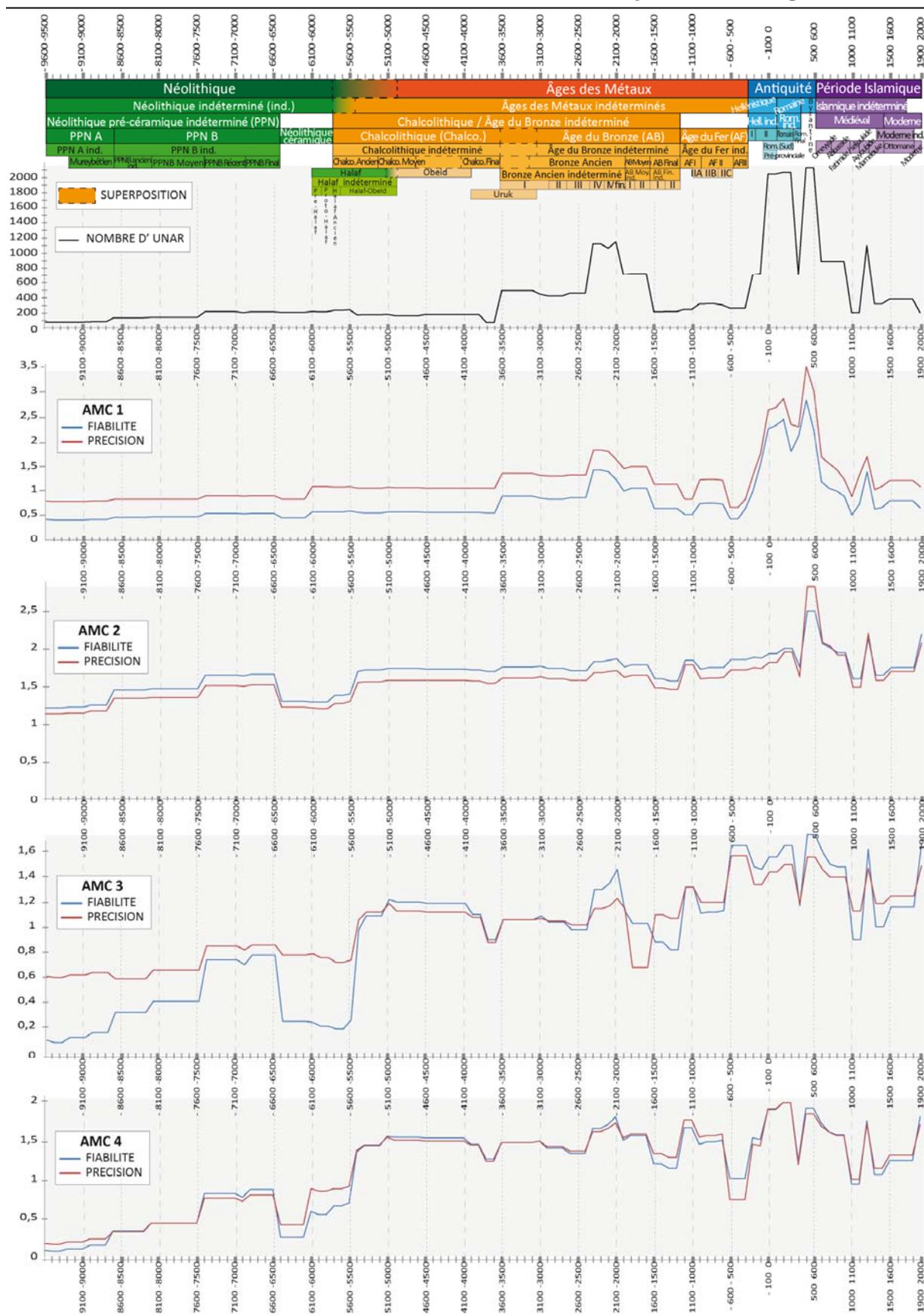


Figure 18 - Résultats de la première application de l'AMC: sur la totalité de la base de données

Les AMC 1 et 2 ont été calculées par la méthode de la somme pondérée, méthode de calcul la plus simple pour apprécier la tendance globale des variations de la qualité de l'information en fonction de chaque alternative. On constate ainsi que les périodes comprises entre le début du Néolithique et la fin du Chalcolithique sont décrites par des données à la fois peu nombreuses, et au niveau de qualité très bas, tant en précision qu'en fiabilité.

Les âges du Bronze ancien et moyen semblent mieux pourvus, tandis que les courbes retombent à partir de l'âge du Bronze récent jusqu'à la fin de l'âge du Fer. Un pic dans la qualité des données est atteint à l'Antiquité, notamment aux périodes Romaine et Byzantine, toutefois entrecoupées d'une légère baisse à la période Romaine tardive. Les courbes baissent ensuite à nouveau à la période Islamique (avec une légère remontée à la période Ayyoubide), la plus récente, et atteignent le même niveau de qualité que les données décrivant l'âge du Bronze.

Les deux points de vue n'introduisent pour le moment aucune différence dans l'analyse, les courbes sont en effet très proches l'une de l'autre. Il est cependant évident que les fluctuations de la qualité de l'information détectées dans la première AMC sont très proches des variations dans le nombre de données, qui ont pu conditionner les résultats.

En divisant les scores de chaque alternative par le nombre de données qu'elles comportent (AMC 2), nous observons de -9600 à -6400 un comportement identique des courbes à l'AMC 1, mais aux variations plus marquées. A partir de -3700, alors que le nombre de données augmente considérablement par rapport aux périodes antérieures, l'analyse tend à atténuer les variations des courbes, en particulier entre -1500 et 200 où la hausse constatée plus haut est ici gommée, et les niveaux de qualité ramenés au même niveau que ceux de la période Néolithique. Quelques inversions sont également notables : alors que l'AMC 1 indiquait une baisse dans la qualité des données à l'âge du Fer, l'AMC 2 montre une légère hausse à cette période – et qui sera confirmée dans les AMC suivantes. En outre, si la hausse observée à la période Romaine est moins accentuée que précédemment, le pic observé à la période Byzantine demeure.

La méthode de la somme pondérée a donc ici certainement une tendance à lisser les résultats, notamment lorsqu'on leur ôte l'influence du nombre de données : l'AMC 2 rééquilibre les résultats en remontant le niveau de qualité des périodes anciennes, et en réduisant les scores des périodes mieux pourvues. Cette courbe indique globalement une qualité de l'information assez constante dans la base de données au fil des périodes, où ne sont notables que des baisses au début et à la fin du Néolithique, et une hausse à la période Byzantine et, dans une moindre mesure, à la période Ayyoubide.

Cette remarque est toutefois fortement remise en cause dès lors qu'on change de méthode de calcul. La méthode de la concordance et de la discordance fait fluctuer les courbes intensément, même durant les périodes précédemment identifiées comme constantes. Les oscillations des courbes ne suivent en outre pas nécessairement les variations dans le nombre de données.

Dans l'AMC 3, la qualité de l'information tend à augmenter alors que l'on avance dans le temps. Trois paliers sont cependant notables dans cette augmentation : le premier du début du Néolithique jusqu'à -5400, le second de -5400 à -600, et le dernier de -600 à aujourd'hui. En outre, les courbes ne sont pas parfaitement stables à l'intérieur de ces trois paliers.

Le premier palier nous indique que la précision des données suit la même tendance que dans les deux précédentes MCA, alors que la courbe représentant la fiabilité se détache de celle-ci et indique des résultats bien plus bas, sauf entre -7500 et -6500.

Les deux courbes se rejoignent à nouveau à -5500 et augmentent alors, marquant le début du second palier. Cette augmentation ne semble pas liée à une croissance du nombre de données, qui s'observe plutôt au début de l'âge du Bronze. C'est en outre à cette période que survient un décalage dans la fiabilité et la précision des données, qui s'opposent significativement.

Le troisième palier est caractérisé par des fluctuations bien plus rapides, avec des alternances de pics et de creux tous les deux cents ans en moyenne. Les deux courbes se suivent cependant à nouveau. La forte baisse à la période Romaine tardive et le pic à la période Byzantine sont toujours visibles, mais moins marqués que sur les courbes précédentes.

La quatrième AMC confirme les tendances observées jusqu'à présent. Les courbes adoptent un comportement proche de l'AMC 3, mais les oppositions entre précision et fiabilité notées précédemment ne sont plus ici visibles, et les deux courbes se superposent presque parfaitement au fil des périodes. L'âge du Fer III est en outre ici marqué par une baisse dans les valeurs, s'opposant à la hausse constatée jusqu'à présent.

Dans un esprit de synthèse, diverses informations peuvent être extraites de l'observation et de la comparaison de ces courbes.

Chaque graphique apporte une dimension nouvelle à l'analyse, et offre au lecteur l'opportunité de mettre en lumière ses divers aspects. Chaque résultat ne doit donc pas être considéré individuellement, mais au regard de tous les autres. L'analyse exploratoire qui est réalisée ici n'a en outre aucunement pour objectif d'identifier un résultat « parfait » ou attendu, mais de tester diverses options, en faisant varier les règles de l'analyse, et en extraire le niveau d'information souhaité et adapté à l'utilisation que l'on souhaite en faire. Chaque résultat est en soi, une partie du problème et une partie de la solution, qui ne prend tout son sens que s'il est confronté aux

autres résultats. Chaque graphe met ici en valeur une « catégorie » particulière de scores, qu'il s'agisse des extrêmes ou de résultats plus intermédiaires.

Divers niveaux de qualité de l'information sont ici révélés selon la méthode de calcul sélectionnée. Si l'on considère la séquence chronologique de façon globale, il semble évident que la qualité de la donnée augmente à mesure qu'elle concerne des périodes de plus en plus récentes. Ce phénomène est particulièrement visible dans les AMC 3 et 4, où la qualité de l'information progresse par paliers successifs. Ces paliers ne correspondent cependant pas au passage d'une grande période (Néolithique, Âges des Métaux...) à une autre, et n'impliquent pas des résolutions temporelles similaires.

Le niveau le plus bas de la qualité de l'information correspond au début du Néolithique, et perdure jusqu'à la période Halaf tardive. La dernière partie du Néolithique est ainsi incluse dans le second palier, qui se termine au début de l'âge du Fer. Le niveau le plus élevé dans la qualité de l'information couvre ainsi l'âge du Fer, l'Antiquité et la période Islamique. Un zoom plus détaillé montre que les courbes fluctuent au sein de chaque palier, et appellent à un niveau d'observation plus fin : dans le premier palier, la qualité de l'information augmente au PPNB ancien et final, et baisse dans les 400 premières années du Néolithique céramique. Le second palier est, quant à lui, plus stable, et peut être considéré à un niveau d'observation intermédiaire. Le dernier palier demande à être considéré plus précisément, à un niveau d'observation de 200 ans pour en percevoir les pics (périodes Romaine, Byzantine et Ayyoubide) et les chutes (Périodes Romaine tardive, Fatimide-Seljoukide, et Mamelouke).

Ces remarques nous conduisent à remettre en question la manière dont les dynamiques des systèmes de peuplement et les changements dans l'occupation du sol passés sont classiquement appréhendés, notamment dans la longue durée, et dans les approches comparatives. Une représentation claire et globale, une « big picture » des bases de données archéologiques devraient avoir de grandes répercussions dans la manière dont nous considérons la continuité spatiale et temporelle de ces données. Ce qui peut, et doit s'accompagner d'une remise en question sur la manière dont nous choisissons nos périodes et nos échelles temporelles d'étude, afin de ne pas confondre les fluctuations des phénomènes humains et environnementaux, et les fluctuations de la qualité et du nombre de données qui les décrivent. Le choix des périodes et des échelles d'étude peut également être orienté par ces informations.

#### 4.2 Application n°2 de l'AMC : approche comparative entre les données des Marges Arides, et celles de la Syrie du Sud.

Afin de tester l'intérêt de cette méthode dans une approche comparative, nous avons extrait deux jeux de données de la base de données PaléoSyr : le premier concerne les sites archéologiques localisés dans les Marges Arides du Croissant Fertile, une zone de transition entre steppe et désert (zone la plus au Nord sur la Figure 19), et le second, les sites de la partie la plus méridionale de la Syrie. Les Marges Arides et la Syrie du Sud ayant fait l'objet de diverses études comparatives (cf. partie I), il s'agit ici de comparer la qualité des données sur lesquelles ont porté ces études.

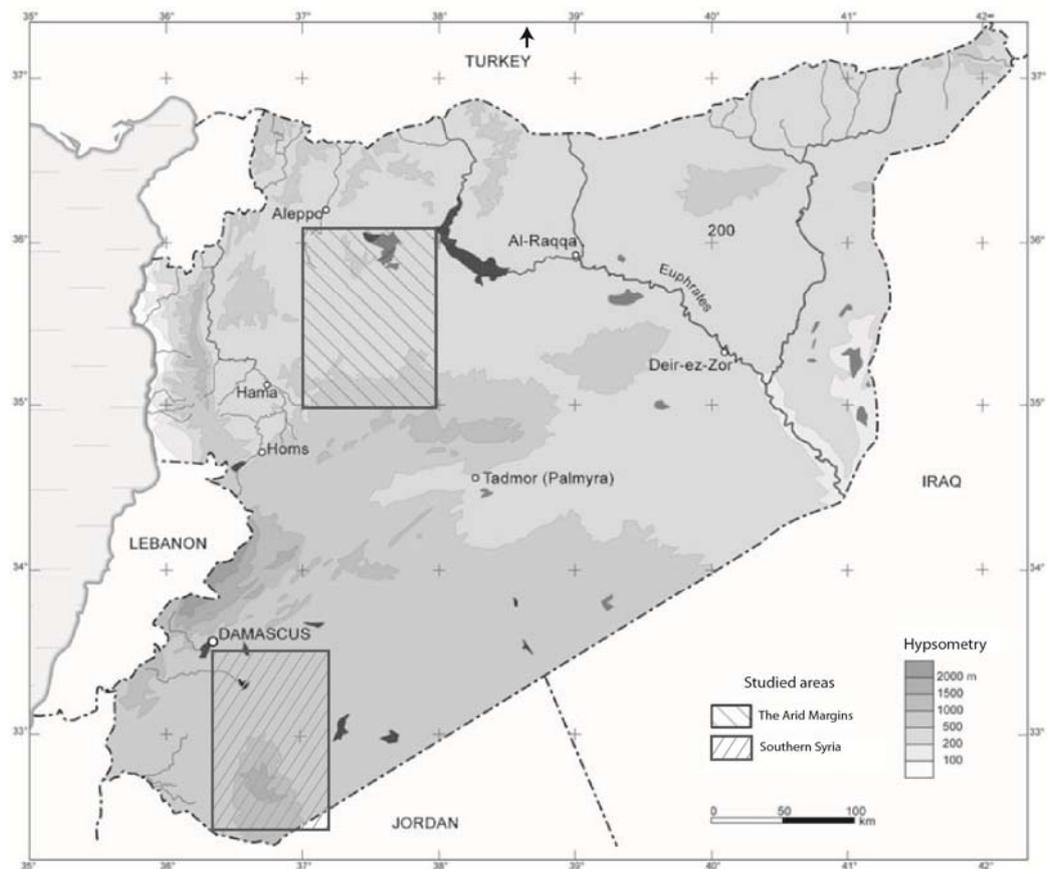


Figure 19 - Carte de situation des zones comparées (d'après Geyer et al., 2005)

Les méthodes de calcul décrites ci-dessus ont ainsi été appliquées successivement sur les deux jeux de données. Les résultats sont présentés sur la Figure 20.

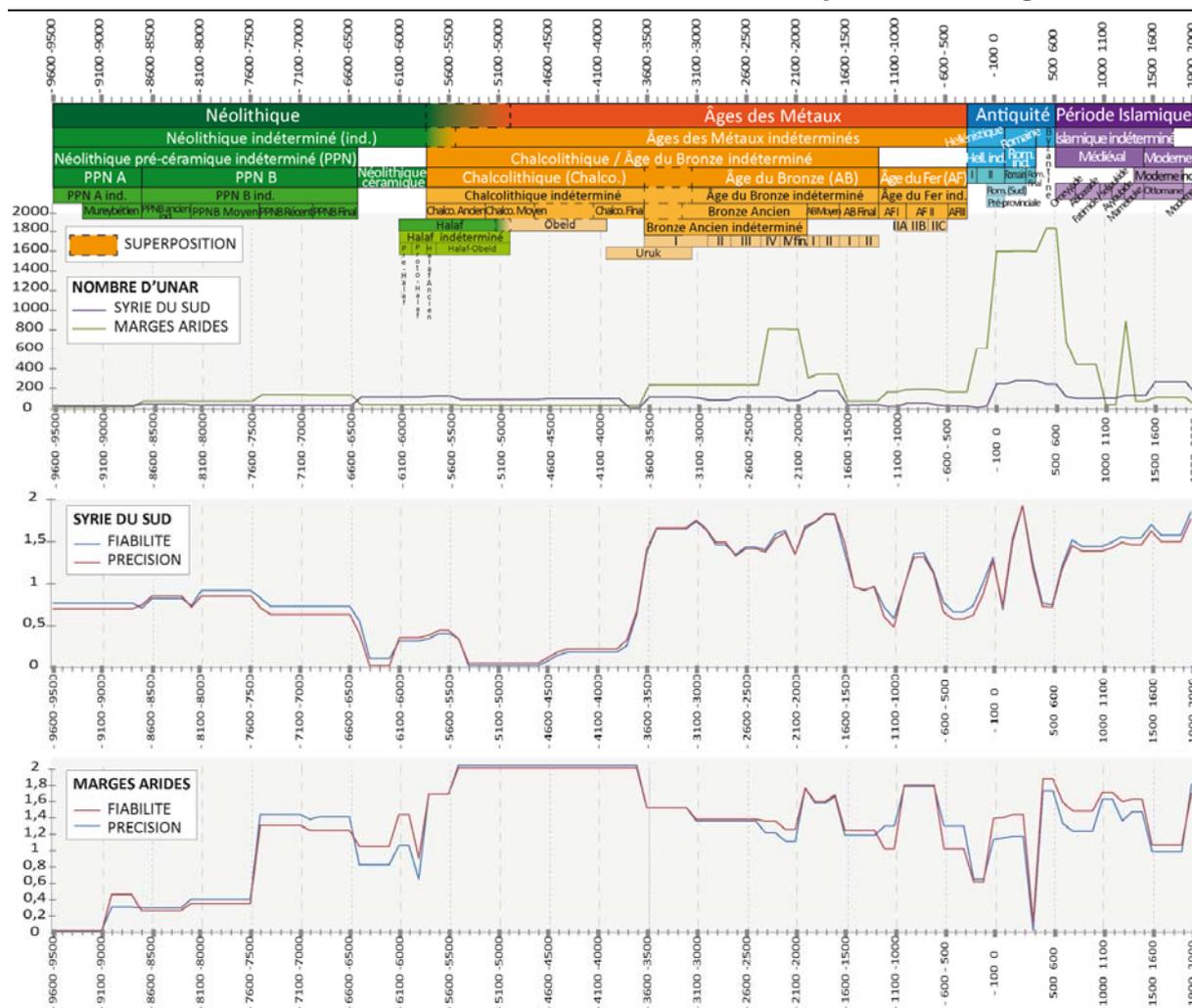


Figure 20 - Résultats de l'application n°2 de l'AMC: approche comparative entre les Marges Arides et la Syrie du Sud

Le premier graphique nous indique que le nombre de données varie fortement au cours du temps, et diffère entre les deux jeux de données considérés. Bien que les données concernant les Marges Arides (7888 UNAR) soient bien plus nombreuses que celles décrivant la Syrie du Sud (1851 UNAR), diverses observations doivent être considérées avant toute interprétation des résultats. En effet, les variations des courbes et leurs différences peuvent indiquer une réalité de terrain, c'est-à-dire une véritable baisse dans l'occupation du sol à une période donnée par exemple, mais également un artefact de la base de données provoqué par une saisie incomplète de données.

Nous savons ici que la différence entre les deux courbes à l'Antiquité est la conséquence d'un état d'enregistrement des données plus avancé pour l'équipe des Marges Arides – contrairement à l'âge du Bronze Final ou la période Mamelouke, où les oppositions dans les courbes témoignent d'une véritable réalité de terrain. Le pic observé à l'âge du Bronze IV pourrait alors être bel et

bien lié à une expansion démographique en Syrie Centrale, corrélée avec une augmentation de la densité et de la taille des sites.

Les graphiques témoignent d'un phénomène intéressant : les choix de saisie des deux équipes archéologiques y sont lisibles. En effet, les courbes représentant la précision et la fiabilité en Syrie du Sud se superposent presque parfaitement, tandis que celles des Marges Arides sont bien moins proches. Ce phénomène peut s'expliquer par le fait que l'équipe de Syrie du Sud fut plus « tranchée » dans ses attributions chronologiques, contrairement à l'équipe des Marges Arides qui a conservé plus de niveaux d'incertitude dans ses données.

En outre, il semble clair que l'échelle chronologique mobilisée par l'équipe de Syrie du Sud est plus fine que celle déployée par l'équipe des Marges Arides pour la période Chalcolithique, les âges du Bronze et du Fer, et que l'inverse se produit aux périodes Romaine, Byzantine et Islamique. La chronologie moins fine au Chalcolithique et aux âges du Bronze I à III dans les Marges Arides peut s'expliquer par le peu de sites datés de ces périodes, et par la nature de ces sites : de très petits campements, dépourvus d'éléments de datation. Distinguer le Chalcolithique de l'Âge du Bronze Ancien aurait donc véhiculé une fausse idée de la précision des données en Syrie du Sud, contrairement aux Marges Arides où cette distinction peut être réalisée plus aisément.

## CONCLUSION

L'objectif de cette étape des traitements était de mettre en place une approche permettant de considérer un ensemble de données non pas directement pour leur dimension interprétative mais pour le poids qu'il est possible de leur accorder dans l'interprétation, en d'autres termes, leurs niveaux de fiabilité et de précision. Il s'agit ainsi dans un premier temps d'analyser les données non pour le contenu de l'information qu'elles nous transmettent, mais pour leur aptitude à matérialiser la réalité que l'on cherche à retracer, en quantifiant cette aptitude. Nous soulignons ainsi l'intérêt de mettre les données en perspective, notamment lorsqu'elles ont été collectées et saisies par des équipes différentes –dont chacune présente ses propres choix, méthodes et objectifs- afin de s'assurer que les phénomènes que l'on observe et interprète ne sont pas des artefacts mais bien les marqueurs interprétables d'une réalité passée. Une représentation claire et globale des données nous offre cette opportunité, et nous a permis :

- D'évaluer les différentes formes que prend la qualité de l'information (précision et fiabilité de la connaissance sur les UNAR) et les divers éléments qui peuvent nous permettre de la quantifier :
- De décrire les variations de la qualité de l'information selon l'intervalle temporel observé, et d'obtenir une première représentation synthétique de la base de données. Adopter une vision globale est à ce stade de l'étude indispensable étant donné le nombre de données et leur hétérogénéité ;
- De choisir une période d'étude pertinente et « accessible » : afin de poursuivre les analyses, nous avons choisi de réduire notre champ temporel d'étude à l'intervalle le mieux pourvu en qualité et en quantité de données, quelle que soit la méthode de calcul considérée. Ses bornes temporelles sont ainsi - 3600 av. J.-C., et 1400 ap. J.-C. Ce choix nous a en outre semblé judicieux dans la mesure où il implique trois grands âges, la seconde moitié des Âges des Métaux (âges du Bronze et du Fer), l'Antiquité et la première moitié de la période Islamique, eux-mêmes découpés en diverses sous-périodes. Plusieurs hypothèses de fluctuation du peuplement et de changements dans l'occupation du sol entre ces diverses périodes ont été fréquemment émises, notamment sur la base de l'observation des variations de la quantité de données, c'est-à-dire, d'UNAR et de sites archéologiques découverts sur l'ensemble du territoire, ou des zones plus restreintes. Proposer une démarche permettant de questionner ces hypothèses à la lumière de la qualité des données sur lesquelles elles sont fondées prend ainsi tout son sens dans cet intervalle temporel.

La méthode et l'optique adoptées ont cependant abouti à une représentation privilégiant l'angle temporel: les variations de la qualité de l'information ont été observées sur l'ensemble de la chronologie de la base de données. Il s'agit désormais de renouer avec une vision spatio-temporelle et d'observer les variations de la qualité de l'information à la fois dans l'espace et dans le temps, en réduisant le champ à l'intervalle sélectionné au cours de cette étape.

La Figure 21 synthétise les apports de la première étape des traitements.

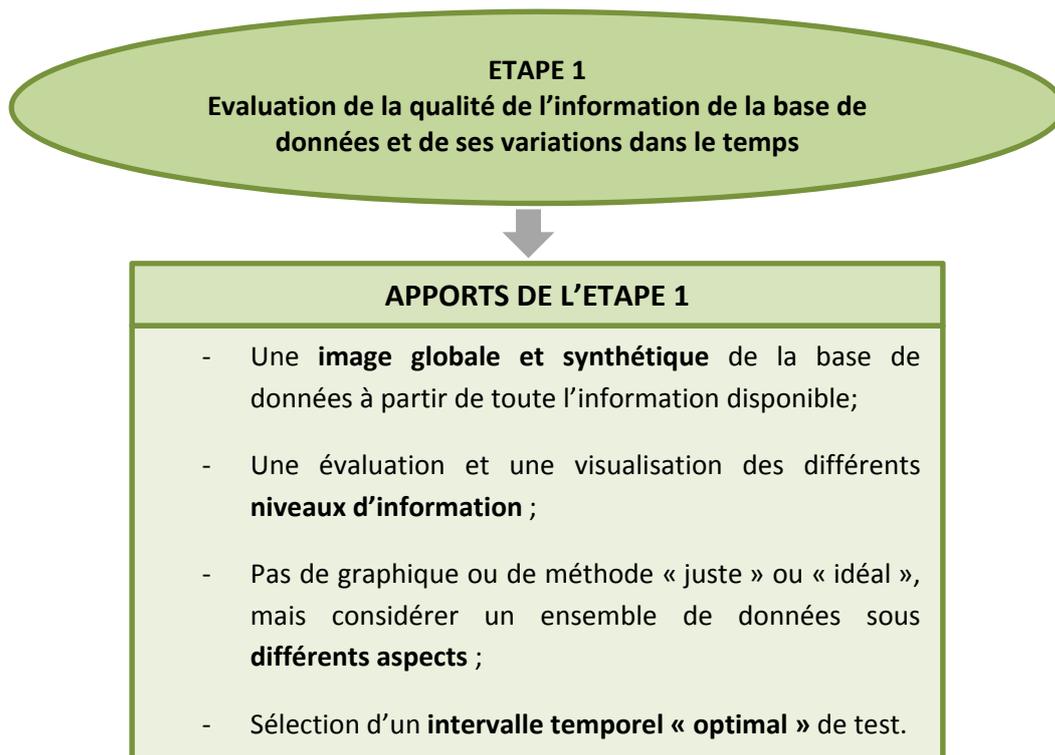


Figure 21 – Synthèse des apports de la première étape des traitements