Chapitre 1.1
Introduction : « quand »

Philippe LANOS

Archéosciences Bordeaux UMR 6034, CNRS, Université Bordeaux Montaigne,
Pessac/Géosciences Rennes UMR 6118, CNRS, Université de Rennes, Rennes

La question de la date d'un événement est une question fondamentale en archéologie : il s'agit de placer un événement ou une série d'événements sur une échelle du temps appelée calendrier, ou encore échelle calendaire (Lefort, 1998). L'unité utilisée est le plus souvent l'année, mais ce peut être aussi des milliers d'années (noté ka) ou bien des mois ou encore des jours, ceci en fonction du phénomène étudié et de la précision des datations. Il faut aussi définir une origine, un instant 0 pour la mesure du temps. Dans le calendrier grégorien (où l'année zéro n'existe pas), on exprimera une date en années BC (Before Christ = avant Jésus-Christ = av. J.-C.) ou AD (*Anno Domini* = après Jésus-Christ = ap. J.-C.). L'ensemble des dates constitue alors une chronologie d'événements. La définition du temps est une question difficile dont les différents aspects, qui ne font pas l'objet du présent manuel, sont bien décrits par exemple dans l'ouvrage très pédagogique « la vraie nature du temps » de V. Boqueho (2018).

Pour dater, il faut réunir deux conditions : bien identifier l'événement que l'on veut dater et choisir une ou des méthodes de datation pouvant atteindre, estimer, la date de cet événement qui est appelée date cible (Dean, 1978). Les méthodes de datation, qui constituent un champ de recherche important au sein des disciplines archéométriques, reposent sur des lois physico-chimiques ou biologiques extrêmement variées (radiocarbone, archéomagnétisme, luminescence, dendrochronologie, méthodes basées sur la radioactivité, Résonance de spin électronique, etc.). Ces lois permettent de relier des mesures au temps *via* des équations d'âge (âge radiocarbone, âge luminescence, …) ou bien *via* des référentiels de calibration, d'étalonnage (date calibrée radiocarbone, date archéomagnétique, date dendrochronologique, …). La datation peut aussi ne pas s'appuyer sur une mesure physico-chimique : c'est le cas avec la typo-chronologie où un intervalle de date de fabrication d'une céramique, par exemple, est donné par ré-

férence à un corpus bien daté par ailleurs. Ici, le processus de mesure est remplacé par un processus complexe où l'archéologue va identifier l'objet à dater avec son type d'appartenance lui-même daté grâce à un corpus de référence. Nous pouvons aussi disposer de la date de frappe d'une monnaie, ou encore de la date fournie par un document épigraphique. La transcription mathématique de ce processus n'est pas sans poser des difficultés, en particulier lorsque l'on veut mettre en œuvre la statistique bayésienne (chap. 12).

En chronologie, les dates sont exprimées en années dans une échelle calendaire, le calendrier grégorien. Les dates avant Jésus-Christ sont notées BC (Before Christ, av. J.-C. ou avant notre ère) et sont représentées par des nombres négatifs (150 BC = -150). Les dates après Jésus-Christ sont notées AD (*Anno Domini*, après J.-C. ou de notre ère) et sont représentées par des nombres positifs (1532 AD = +1532). il est important de faire la distinction entre l'âge A et la date t. L'âge est calculé comme la différence entre une date t et une année de référence t_0 choisie sur l'échelle calendaire. Quelle que soit la valeur de t_0, si l'on veut que l'âge A soit exprimé positif pour des dates t qui précèdent t_0, il faut utiliser la relation simple suivante : $A = t_0 - t$. Dans le cas de la datation par le radiocarbone, une mesure de taux résiduel de ^{14}C est transformée en âge apparent (exprimé en années BP = Before Present = 1950) qui est à son tour transformé en date calibrée BC/AD (abbrégé en "date cal." BC ou AD) via un processus dit de calibration (*cf.* 12). On peut ensuite revenir dans une échelle d'âge en prenant $t_0 = 1950$ pour année de référence et en appliquant la formule précédente $A = t_0 - t$. On obtient alors un âge positif qui est un âge dit "âge calibré BP". Dans certaines publications, on voit aussi des âges dits "calibrés B2k", ce qui veut dire que l'année de référence choisie est $t_0 = 2000$.

Les résultats de datation communiqués par les laboratoires sont le plus souvent exprimés par une valeur moyenne accompagnée d'une erreur d'estimation de cette moyenne, ou encore par un intervalle d'incertitude déterminé à un certain niveau de confiance (le plus souvent à 95%). Lorsque le résultat de datation est entaché d'incertitude, on exprime le résultat sous la forme d'une distribution de probabilité (ou densité de probabilité) qui résume toute l'information que l'on a sur la date d'un événement. Une datation de plus en plus précise verra sa distribution se resserrer jusqu'à devenir un « bâton », plus précisément un rectangle (distribution uniforme) dont la base sur le temps tend vers zéro (l'instant) : cette fonction devient une distribution nommée « Dirac ». Toute distribution de probabilité est de surface égale à 1, autrement dit, il y a 100% de chances que la date vraie se situe dans le support de la date, c'est-à-dire dans l'intervalle où la date peut exister. Une date historique à l'année près, au jour près, à la minute près peut se représenter par un « Dirac ». C'est une datation précise, mais qui dit précis ne veut pas forcément dire fiable. En effet, il peut y avoir une erreur systématique, irréductible au sens où elle ne peut être réduite par un processus d'échantillonnage, due à des causes plus ou moins bien connues : erreur de prélèvement, d'attribution, de conservation, etc. Ainsi, la fiabilité d'une date ne peut pas être évaluée à partir des seules informations statistiques internes au processus de datation. Cette fiabilité ne peut être évaluée que par comparaison à d'autres datations sur des objets provenant des mêmes contextes. Des datations qui s'écartent peuvent alors apparaître comme des valeurs aberrantes (*Outlier* en anglais) qu'il convient de

détecter en développant des modèles statistiques de combinaison des dates (*cf.* 1.2, 1.4, 1.5 et 12).

Ainsi, les outils statistiques mobilisés en datation sont essentiellement focalisés sur la gestion des erreurs dont les causes sont nombreuses et différentes d'une méthode à une autre. Outre les outils classiques de la statistique descriptive (moyenne, écart-type par exemple, *cf.* 4), on doit aussi faire appel aux distributions de probabilité (*cf.* 7) pour décrire les erreurs portant sur les mesures effectuées ainsi que sur les paramètres dont dépendent ces mesures, dont en particulier le temps. L'estimation du temps, autrement dit d'une date cible (date d'intérêt pour l'archéologue, l'historien, le paléo-environnementaliste…), à partir d'une ou plusieurs mesures provenant d'une ou plusieurs méthodes de datation, se fera, d'un point de vue statistique, soit par une approche probabiliste dite fréquentiste où on cherchera la date qui maximise la vraisemblance des observations (*cf.* 8), soit par une approche dite bayésienne où on cherchera la distribution *a posteriori* de la date en combinant la fonction de vraisemblance (loi d'échantillonnage ou loi des observations) avec une distribution de probabilité *a priori* sur la date cible, *a priori* qui est défini au préalable sur la base de connaissances antérieures apportées par l'archéologue (*cf.* 12).

Pour illustrer la façon de combiner les dates et de tenir compte des incertitudes, nous présentons dans les sections 1.2 à 1.6 cinq exemples de chronologies établies dans des contextes archéologiques très différents. Ces exemples traitent selon les cas des erreurs qui interviennent dans le processus de datation ou bien de la façon de combiner les datations obtenues avec des informations archéologiques comme la datation relative qui est fondée sur la stratigraphie, sur la sériation ou sur tout autre critère pouvant définir une relation d'ordre sur le temps. Nous insistons aussi sur la nécessité de bien décrire les données de datation (contextes prélevés, adéquation entre événement ciblé et datations réalisées, protocoles de mesure suivis en laboratoire…) et de les rendre accessibles *via* des bases de données (tableaux de données) documentées qui permettent de reproduire et de critiquer les raisonnements suivis. Par exemple, concernant le radiocarbone, il est important de pouvoir disposer des mesures d'âge avec leur erreur associée exprimée à 1 sigma afin de pouvoir transformer cette mesure en date calendaire avec la toute dernière courbe de calibration mise à jour.

Nous sommes conscients que ces exemples sont loin de couvrir tous les aspects des statistiques mobilisées par les méthodes de datation. On pense par exemple à la dendrochronologie qui fait appel aux tests statistiques pour juger de la qualité des synchronisations des séquences de largeurs de cernes d'arbre. L'important dans cet ouvrage est de montrer le raisonnement suivi dans le traitement des erreurs ainsi que dans la présentation des résultats de datation, qui associe étroitement mathématiques et statistiques avec archéométrie et archéologie.

W. Banks et Ph. Lanos (*cf.* 1.2) montrent comment la chronologie joue un rôle clé dans l'examen des relations culture-environnement, en particulier pendant les périodes caractérisées par une variabilité climatique et une réorganisation environnementale rapides et marquées comme c'est le cas du Paléolithique supérieur en France actuelle de 32 000 à 21 000 ans calibrés BP. L'exemple développé prend en compte des mesures de radiocarbone examinées de manière critique et provenant de contextes archéologiques

étudiés en détail dans le but de reconstruire la chronologie des cultures archéologiques du Gravettien moyen au Badegoulien. Le traitement statistique des données chronologiques repose ici sur une modélisation statistique bayésienne (*cf.* 12) reposant sur l'utilisation du logiciel ChronoModel. Cet exemple montre l'étroite interaction entre l'analyse des données archéologiques et archéométriques et la construction du modèle statistique permettant d'établir la chronologie.

L. Bellanger *et al.* (*cf.* 1.3) présentent l'outil de datation CeraRdate pour des données céramiques issues de sites archéologiques du bassin de la Loire. Cet outil s'insère dans un projet plus vaste nommé SPARTAAS (Statistical Pattern Recognition and daTing using Archeological Artefacts assemblageS) qui a pour objectif de développer des outils statistiques dédiés à l'analyse des données mobilières volumineuses dans une démarche de modélisation archéo-statistique. Il s'agit d'estimer la date d'une action inscrite dans le sol – pour l'archéologue, d'un contexte archéologique – en fonction de l'assemblage céramique qui le compose. La méthode CeraRdate consiste à construire un modèle de régression linéaire liant une date connue dans le temps calendaire – ici celle de l'émission d'une monnaie –, à son faciès céramique.

P. Guibert et Ph. Lanos (*cf.* 1.4) abordent la question difficile du traitement des erreurs dans la méthode de datation par luminescence, appliquée ici à la datation de la construction des édifices en archéologie du bâti et en histoire de la construction. Les difficultés méthodologiques sont traitées sous l'angle de la modélisation des erreurs (dites statistiques et systématiques) avec une application à la datation de la tour-clocher de la collégiale Saint-Martin à Angers (France)

A. Quilès (*cf.* 1.5) modélise une chronologie dynamique en prenant exemple sur l'Égypte ancienne, où textes, objets et datations radiocarbone sont confrontés. Deux exemples sont développés : l'étude chronologique du site de Baharya et la chronologie de la 18^{e} dynastie, en développant des modèles chronologiques bayésiens reposant sur l'utilisation du logiciel OxCal. Tout l'enjeu est de poser un dialogue productif entre les approches historiques et archéométriques en les inscrivant dans une même chaîne de raisonnement, au sein d'un modèle statistique confrontant des données hétérogènes issues de sources textuelles, historiques, archéologiques comme des données analytiques (radiocarbone).

Le dernier chapitre présenté par B. Desachy (*cf.* 1.6) montre qu'il existe en archéologie une importante tradition de traitement graphique basé sur les permutations de lignes et de colonnes de tableaux de données transformés en matrices graphiques et croisant des structures, objets ou ensembles avec des traits culturels. L'objectif principal de ces traitements est la sériation chronologique, basée sur l'hypothèse que la diagonalisation de telles matrices graphiques exprime une évolution chronologique dans laquelle s'ordonnent individus et variables. Cette tradition propre aux archéologues et purement empirique à l'origine a rencontré, en France, les travaux de traitement graphique des données de J. Bertin et, plus généralement, les outils mathématiques et informatiques (par exemple Ph. Cibois), pour aboutir aujourd'hui à une approche intégrant des notions statistiques et utilement complémentaires de l'analyse statistique multidimensionnelle (*cf.* 6), à condition de prendre en compte les limites du modèle du temps sériel.