

## IV. The Absolute Chronology of the Excavations: Radiocarbon Dating and Stratigraphic Age Modelling

---

*Bernhard Weninger – Giorgos Toufexis – Christos Batzelas*

### IV.1. Introduction

To establish an absolute timescale for the Neolithic stratigraphy of PMZ we have at our disposal altogether ten  $^{14}\text{C}$  ages on short-lived animal bones (Tab. IV.1). The  $^{14}\text{C}$  measurements were processed in 2017 at the  $^{14}\text{C}$ -AMS laboratory of the Klaus Tschira Archaeometry Centre of the University of Heidelberg under the direction of Ronny Friedrich. The majority of  $^{14}\text{C}$  ages have standard deviations of  $\pm 30$  BP or better. During sampling, great care was taken to avoid any kind of ‘old carbon’ effect due to old wood (inner tree rings), or recycled old wood (secondary use, etc.).

### IV.2. The $^{14}\text{C}$ Database

The data table begins with a serial number (ID1 to ID10) which is to be used below to simplify the identification of the samples (e.g. ID1=MAMS-32119). Then follows the laboratory number (MAMS Labcode) and corresponding conventional  $^{14}\text{C}$  age. In accordance with the advice given in the MAMS  $^{14}\text{C}$  data sheets, in Tab. IV.1 we do not provide the values measured for the  $\delta^{13}\text{C}$  fractionation of the samples, which are machine-specific values (they describe the AMS beam). The next columns show the dated material (bone), the carbon content (%), the measured C/N (carbon/nitrogen) relation, and the collagen content (%). According to the information provided by the Mannheim laboratory, the collagen was extracted using a weak acid dissolution, followed by ultra-filtration and separation of the fraction  $>30\text{kD}$ . The extracted organic carbon was then dry-frozen and burnt in an elemental analyser to produce  $\text{CO}_2$  which was further catalysed to produce graphite.

According to the laboratory, the measured values for carbon content, C/N relation, and collagen content are all within the ranges to be expected for well-preserved collagen (C/N: 2.6–3.2, collagen content  $>1\%$ ). The comment that the conspicuously low collagen content of ID10 (MAMS-32133) might be due to the higher wet acidity of the settlement ditch from which the bone derives, compared to the apparently drier house sediments from whence the other samples derive, is our own observation, albeit hypothetical. Altogether, despite some exceptions, the collagen preservation appears to be better for the stratigraphically higher samples. The following columns named ‘Trench, Unit’, ‘Depth [cm]’, ‘Phase’, and ‘Context’, contain the archaeological information that we have used for the stratigraphic (age-depth) modelling. The final column of Tab. IV.1 shows the calibrated ages for the single  $^{14}\text{C}$  measurements, as referenced to the calBC-scale, with calibrated dating uncertainties given at 68% confidence. These calibrated ages were derived by using CalPal software (Version 2009.5) and the IntCal13 calibration curve.<sup>222</sup>

---

<sup>222</sup> Reimer et al. 2013. We have resisted the temptation of updating the PMZ-chronology to IntCal20. For the interval 5900–5450 calBC the calculated change in the chronology of PMZ would have amounted to  $-0.1 \pm 2.7$  BP (i.e. one  $^{14}\text{C}$ -month). For a preliminary presentation of the radiocarbon dates of the site see Toufexis et al. in press.

Tab. IV.1 Radiocarbon ages (N=10) from PMZ, arranged according to MAMS LabCode (Mannheim <sup>14</sup>C-AMS) (B. Weninger)

ID	Lab-Code	<sup>14</sup> C Age [BP, 68%]	Material	C [%]	C/N	Collagen [%]	Trench, EU	Depth [cm]	BPh/ BSPH	Context	Calendric Age [calBC, 68%]
1	MAMS-32119	6538 ± 27	Bone	27.5	3.2	3.1	A, 255	491	VIIc	Below Surface F20	5502 ± 24
2	MAMS-32120	6651 ± 28	Bone	22.6	3.2	1.1	A, 244	522	VIIb	Established via depth	5583 ± 32
3	MAMS-32123	6870 ± 27	Bone	23.5	2.8	5.4	A, 219	591	VIb	From Surface F24	5751 ± 33
4	MAMS-32124	6855 ± 30	Bone	22.9	2.8	2.0	A, 204	657	Vd	Below Surface F26	5733 ± 34
5	MAMS-32125	6935 ± 36	Bone	18.3	2.7	1.1	A, 195	745	Va	Below Surface F29	5813 ± 49
6	MAMS-32126	6938 ± 30	Bone	18.0	2.7	3.7	A, 192	775	IVb	Below Surface F30	5814 ± 46
7	MAMS-32129	6892 ± 33	Bone	15.0	2.8	1.9	A, 185	876	IIIb	Between Surfaces F33a-d	5777 ± 41
8	MAMS-32130	6948 ± 29	Bone	20.0	3.2	1.5	A, 184	894	IIIb	Between Surfaces F33a-d	5824 ± 46
9	MAMS-32131	6948 ± 29	Bone	24.1	3.2	0.8	A, 179	967	II	Below Surface F34	5824 ± 46
10	MAMS-32133	6965 ± 28	Bone	24.2	3.2	0.4	A, 176	1022	I	Ditch	5843 ± 45

We note in advance that, in our discussion of the PMZ chronology, the only use made of the (unmodelled) single-sample calibrated ages shown numerically in Tab. IV.1, and graphically in Fig. IV.3, is to demonstrate the existence of some strong, artificial chronological distortion (order of magnitude ~100yrs). This age distortion, which is mainly relevant for the youngest PMZ phases, would be immediately introduced into any chronology that is based on the single-sample calibrated ages. This kind of age distortion is not new, but its existence is not widely acknowledged. It can be understood as due to systematic lock-in effects of the multiple <sup>14</sup>C readings on the wiggles of the calibration curve.<sup>223</sup> An elementary description of this effect would be that, since there is a higher number of <sup>14</sup>C-scale readings on the temporally extended wiggles or plateaus of the calibration curve, there is an a priori higher probability that the plateau readings become integrated in our archaeological chronology. This applies, whether or not the plateau readings represent the true sample ages. From a more fundamental mathematical perspective, the age distortion is methodologically preordained (i.e. mathematically inevitable) due to the generally (variable) non-commutative algebraic structure of the calibration curve. In the case of PMZ data, the age distortion is strongest for the two youngest single ages. It exists naturally for all age considerations (i.e. not only for explicitly undertaken age modelling) that are based on calibrated single ages, but can be easily avoided, by use of the chronological results obtained from linear stratigraphic (metric) age-depth modelling, as described below (Fig. IV.5). Application of this linear age model would, strictly speaking, require some independent demonstration that the sediment deposition at PMZ is both continuous and without occupation gaps. Only then could we expect that the PMZ tell growth can be adequately described as resulting from sedimentation processes that add up to

<sup>223</sup> Weninger et al. 2011; Weninger et al. 2015.

produce a time-constant accumulation rate. In effect, although intended to overcome the fallacies associated with naïve application of calibrated single ages, the applied stratigraphic age-depth modelling would itself require further studies, since it would otherwise be founded on similarly undemonstrated assumptions as are associated with the use of single calibration ages. Although lacking these studies, the application of an explicit age-depth model at PMZ nevertheless has a major advantage over single-age dating in that the achieved site chronology, once established, immediately becomes testable. Although the explicitly formulated model assumptions may, indeed, be faulty, they are transparent and open to discussion. A further advantage of explicit age modelling is the possibility to identify potential outliers, both in terms of erratic  $^{14}\text{C}$  measurements and of archaeologically misaligned samples. Yet, we state this as a methodological caveat not only as regards many published  $^{14}\text{C}$ -based archaeological studies but also for our own results, mentioning at the same time that we are not alone. The point hereby is that, in our view, so-called ‘archaeological age modelling’ of archaeological  $^{14}\text{C}$  dates – although widely applied – is still very much in its infancy. In published studies to date, the main (and often only) component of archaeological age modelling under study is the mathematical one (most often Bayesian). In other words, the large majority of presently so-called ‘archaeological age models’ are, in actual fact, most often little more than ‘statistical descriptions of the available  $^{14}\text{C}$ -data.

Now, as for the archaeological component of stratigraphic age-depth modelling, what we presently know from experience (without modelling) is that the majority of  $^{14}\text{C}$ -dated tell mounds in Anatolia and southeast Europe have an average sediment accumulation rate of  $\sim 1\text{cm/yr}$ , with uncertainty in the order of  $\pm 20\%$ . This is presently little more than an empirical rule-of-thumb, awaiting further confirmation and, in particular, awaiting further quantitative sedimentological description. Nevertheless, it can be used for the purposes of forecasting. With around 5m of Neolithic sediments at PMZ, what we may therefore expect is an overall time-depth of  $500 \pm 100\text{yrs}$ . Looking ahead, this forecasting is once again confirmed, when based on the tabled values for the single ages (cf. Tab. IV.1 and Fig. IV.3). The more refined modelling result, namely that some 350yrs of sediment accumulation are observable in PMZ, represents a value on the ‘faster-than-average accumulation side’ of the forecasting. In addition, given that PMZ has altogether 19  $^{14}\text{C}$ -dated building phases and building subphases, and given that the average phase length is calculated as 20 years, we also find that the calculated average phase length corresponds neatly with the expected time span of one human generation. In other words, the above-mentioned rule-of-thumb apparently also applies to the promising possibility of using measured tell-heights, beyond first-order dating, for the purpose of demographic forecasting (based on e.g. estimated 3D sediment volumes).

### IV.3. Data Processing and Calibration Software

The present analysis is based on application of the  $^{14}\text{C}$  age calibration software CalPal (Version 2019.5). CalPal software is written in Fortran 95 using Intel® Visual Fortran XE Compiler XE with integrated IMSL® 6.0 Numerical Libraries by RogueWare®. The WinOS interface is provided by Winteracter® version 12.0e. For purposes of high-resolution  $^{14}\text{C}$  age modelling, which occasionally requires Monte Carlo runtimes in the order of 10–20 hrs, CalPal-software is equipped with runtime forecasting time/date Fortran routines that were developed by QTSsoftware®. CalPal is installed on a Celsius W530® workstation with a Xeon E3-1281v3® 3.7 GHz CPU with 61cm display. The majority of CalPal routines are otherwise designed for PC desktop and notebook use (with minimum 12” display). The CalPal package is equipped with a number of import/export interfaces, also written in Fortran 95, and that support data flow with the following external (commercial) programs: Excel® ( $^{14}\text{C}$  database management), Globalmapper® (archaeological site mapping with 30” SRTM® data), and Un-Scan-IT® (automated graphic digitisation of climate records).

Important for the present Gaussian Monte Carlo Wiggle Matching (GMCWM) application is the introduction (in CalPal Version 2017.1) of some more flexible Fortran 95 subroutines for calculation of the best-fitting Chi-squared probabilities. These new routines support calculation of so-called non-central chi-square probability distributions, which enable a more sensitive adjustment of the statistical variables used in GMCWM. From literature we know that the non-central chi-square distribution is well adapted to cases where the distribution function is a mixture of random variables, but for which a small amount of bias cannot be excluded, if only due to limited data availability. This would appear sensible in our context, since archaeological and calibration-curve  $^{14}\text{C}$  data sets typically contain measurements that have generally small but in any case unknown variability in terms of e.g. interlaboratory offsets, local  $^{14}\text{C}$  reservoir effects, sequence-internal bias etc. Conveniently, for the purposes of CalPal Fortran programming, the required subroutines are implemented ready-for-use in the IMSL<sup>®</sup> 6.0 libraries, whereby CalPal makes specific use of the routine called CSNDF, which allows the passing of a (variable) non-centrality parameter  $\lambda$ . In the CalPal GMCWM dialogue, this passing of the variable  $\lambda$  from the screen to the CSNDF routine is implemented by a trackbar. Furthermore, when the non-central chi-squared functionality

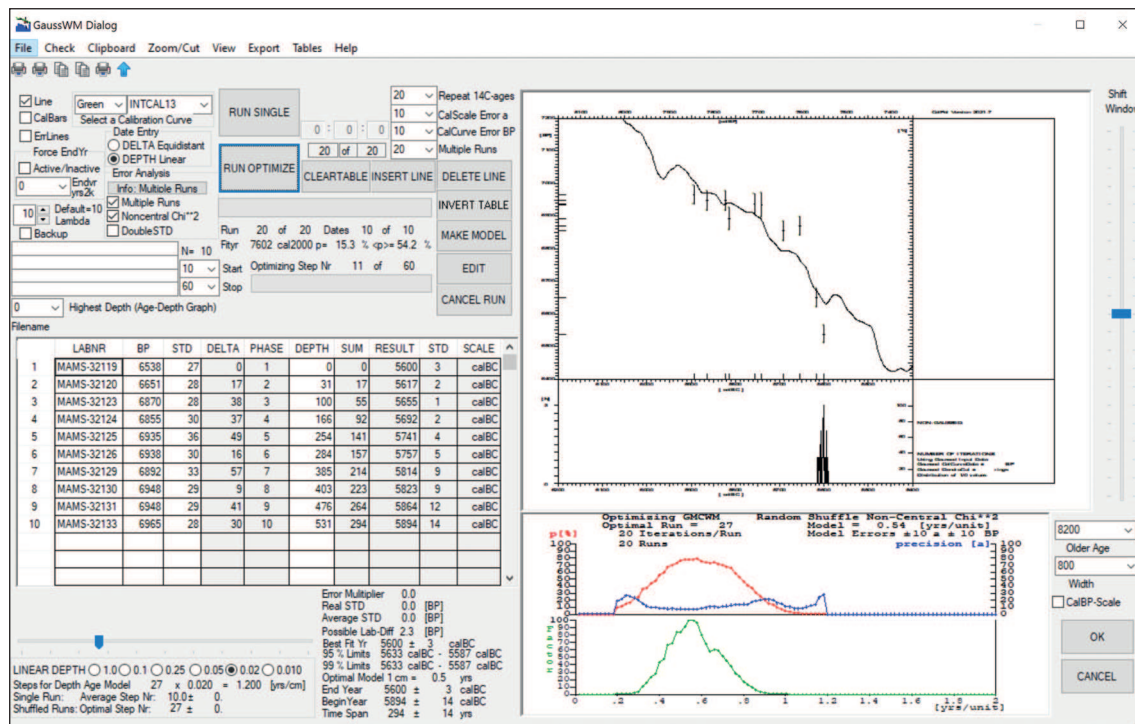


Fig. IV.1 Screenshot of GaussWM dialogue (CalPal Version 2021.6) to illustrate the application of GMCWM to PMZ data. Following import of GaussWM data from an Excel-format file, the study data is shown in stratigraphic order in the spreadsheet, with associated depth values given in the column with heading DEPTH. During the runtime the sequence of  $^{14}\text{C}$ -ages is expanded stepwise, with the best-fitting results at each step shown in the upper graph, and corresponding statistical parameters shown in the lower graph (red=probability, blue=dating precision, green=optimisation factor=probability/precision). Following selection of age-model and runtime parameters, the GaussWM analysis is started with the button RUN OPTIMIZE. During the runtime the developing results are shown (animated) in graphic windows (right) and in spreadsheet rows DELTA, SUM RESULT, STD. Final results are stored to numeric (ASCII, xls-format) and graphic files (selectable format). At the end of the runtime, the overall best-fitting sequence is derived by numeric analysis of the stored sets of statistical variables, the best-fitting graph is established, and corresponding statistical variables are stored to file. The dialogue (above) shows in the upper graph the age-position of the PMZ study data (error bars) in relation to the calibration curve (green line: IntCal13 calibration curve). For purposes of final, visual quality control of the achieved results, a separate dialogue (not shown, cf. Weninger 2017, Appendix) allows joint graphic projection of the GaussWM results with the raw data sets of the laboratories involved in the construction of the high-precision calibration curve (present study: QL-Seattle, UB-Belfast, Hd-Heidelberg) (B. Weninger)

is chosen (by checkbox), the GMCWM dialogue will respond with the proposal to use  $\lambda=10$  as the default value, whereby the user is provided with the information that  $\lambda=0$  corresponds to traditional chi-squared analyses. For explorative purposes in the application of GaussWM, other  $\lambda$  values can be entered into a spin-down integer field named *Lambda* (Fig. IV.1). Finally, we have checked that  $^{14}\text{C}$ -age calibrations and, in particular, the GaussWM results described in the following (based on Intcal13 and CalPal Version 2017.5) can be accurately reproduced by application of recent updates to Intcal20 and CalPal Version 2021.6, with max. deviations in the order of numeric rounding ( $\pm 5$  yrs, 95%).

#### IV.4. Single-Age $^{14}\text{C}$ Age Calibration

##### IV.4.1. Overview: Dispersion Calibration

Fig. IV.2. provides an overview of the radiocarbon ages from PMZ (Tab. IV.2). It shows the  $^{14}\text{C}$ -scale histogram and summed calibrated  $^{14}\text{C}$  age probability distribution for all PMZ data (N=10). For this kind of data representation, it is important to note that the BarCode ages (small vertical lines on the calendric time-scale) – defined as central values of 95%-confidence intervals – are often strongly age-distorted (centennial-scale), due to the non-commutative (age-folding) properties of the  $^{14}\text{C}$  age calibration curve. However, in the case of the present PMZ data, such age distortion apparently applies mainly to the two youngest dates (ID1 and ID2), both of which have calibrated median values that are centred inside the two V-shaped re-entry wiggles at around 5600 and 5500 calBC, respectively. Due to the strongly non-commutative character of the underlying probabilistic calibration algebra, at these positions we can therefore immediately suspect the possible existence of some particularly strong age distortion. Closer inspection of the internal time structure of the two wiggles allows us to estimate the potential magnitude of the age distortion as (max.) approx. 100–120yrs, whereby the actually occurring age distortion in the PMZ age model will be dependent on a number of parameters and, in particular, on the quality of the applied IntCal13 calibration curve. As it turns out (compare Fig. IV.3 and Fig. IV.5), the calibration-induced distortion of the age model based on single dates is  $\sim 100$ yrs, i.e. close to the theoretically possible local maximum. This is a quite extraordinary finding. It can be explained as due to a chance combination of the measured  $^{14}\text{C}$ -age (ID1) with the structure of the (potentially non-ideal) calibration curve in the region of 5500 calBC. At this critical point, as can be seen from Fig. IV.2, it is conspicuous that the wiggle in the IntCal13 curve at around 5600 calBC is constructed such that all (N=6) available high-precision laboratory tree-ring measurements in the near (decadal) vicinity of this wiggle are all situated below the calibration curve, whereas all (N=4) of the next younger high-precision calibration ages are situated above the IntCal13 curve. Finally, the strong downward bend of the IntCal13 curve between 5670 and 5600 calBC (and which is so clearly evident not only in PMZ, but also in the IntCal20-update), provides some quite compelling support for the PMZ age model.

Interestingly, and now excluding ID1, for all other PMZ dates, on average, the calibrated single-age values are already in approximately correct stratigraphic order, even for the otherwise so often problematic method of dispersion calibration (Fig. IV.2). This property of PMZ data is altogether promising, in terms of the possibility of constructing a high-resolution archaeological chronology, when based on quantitative (metric) age-depth modelling.



## IV.4.2. Single-Date Age-Depth Model

As already indicated, the main aim of the present stratigraphic  $^{14}\text{C}$  analysis of PMZ data by application of GMCWM is to construct an age-depth model of the Middle and Late Neolithic deposits of PMZ, such that i) each of the nine building phases (BPh I–IX) is assigned a calendar age and its associated dating uncertainty; ii) similarly, each of the alternatively defined 19 subphases (I, II, IIIa–c, IVa–b, Va–e, VIa–b, VIIa–c, VIII, IX) requires an assigned calendar age and dating uncertainty, and finally iii), what is also required is a graph (and corresponding numeric equation) from which the calendric age of any tell deposit can be calculated, together with an estimate of the achieved dating precision, provided that its stratigraphic depth (cm) is given.

To achieve this goal, let us now test whether this chronology can be achieved simply by constructing a graph that shows the calibrated ages versus the stratigraphic depth from which the dated (bone) samples were taken. To be sure, such an approach represents the currently most widely applied method for construction of age-depth models in many disciplines e.g. in geomorphology, marine geology, as well in archaeological research. A graphic representation of the results achieved by this uncomplicated, unpretentious and, indeed, seemingly natural manner of age-depth representation of the PMZ data is shown in Fig. IV.3. This figure also provides the table (Tab. IV.2) of the numeric values and archaeological variables used in the chronology construction ( $^{14}\text{C}$  age, calibrated age, sample depth (absolute scale [cm]), sample depth (relative scale [cm]), rescaled with the depth of the youngest dated sample set to  $\Delta=0$  [cm]), sample phase and subphase assignment).

Put differently, for PMZ date MAMS-32119 (ID1), the (vertically) oversmoothed construction of IntCal13 leads to a (horizontal) age distortion of around 100 yrs, such that the calculated calibrated  $^{14}\text{C}$  age of  $5502 \pm 24$  calBC is around 100 yrs too young. Given that this age distortion is statistically significant, it is no wonder that the single-age-based age model shows an unexpectedly strong bend away from the initial linearity 5850–5700 calBC, at around 5700 calBC and for younger ages (Fig. IV.3).

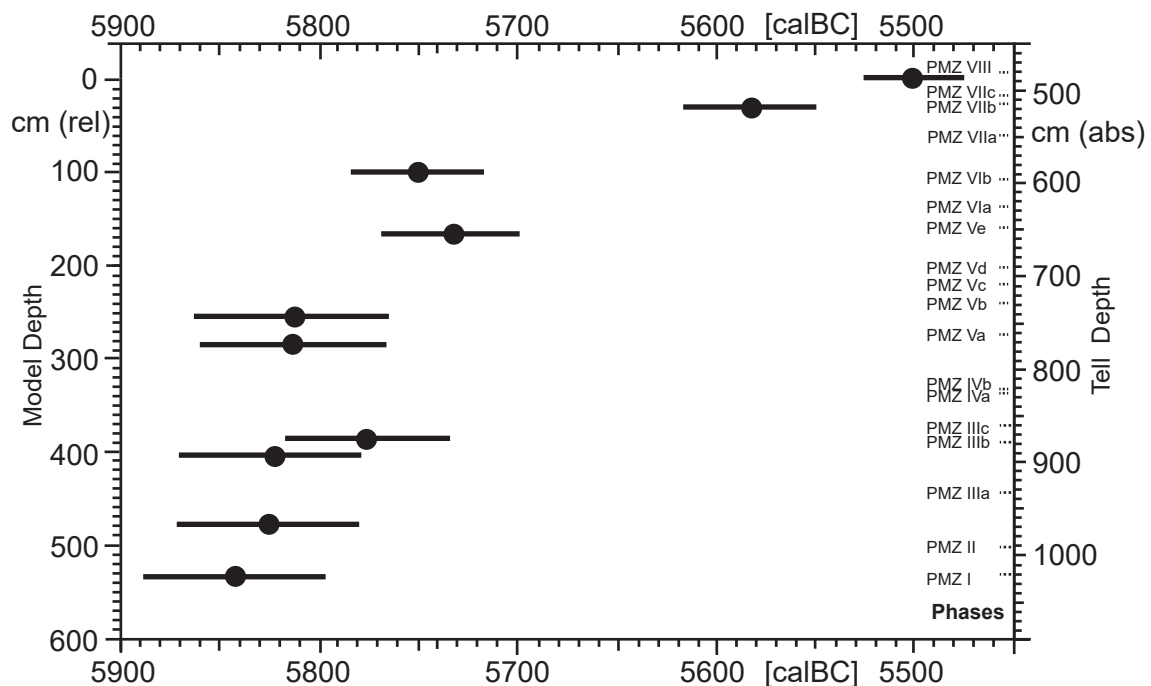


Fig. IV.3 PMZ age-depth model based on calibrated single dates (Tab. IV.2.3). Note the strongly age-distorted shape of the age-depth relationship, but which is simply due to the non-linear character (in this case: ‘bent shape’) of the underlying  $^{14}\text{C}$  age calibration curve, in combination with one by chance badly designed wiggle at around 5600 calBC (cf. text) (B. Weninger)

## IV.5. Gaussian Monte Carlo Wiggle Matching

### IV.5.1. Methodology

Gaussian Monte Carlo Wiggle Matching is an extension of the wiggle-matching method that was first developed for dendrochronological applications by G.W. Pearson.<sup>224</sup> The method was generalised by Bernhard Weninger for archaeological studies.<sup>225</sup> In both cases, the basic approach is to fit a given sequence of <sup>14</sup>C ages to the calibration curve by minimising the summed and squared differences between the <sup>14</sup>C data (of unknown calendric age) and the calibration data (known calendric age). The calendric age that belongs to the calculated chi-squared minimum can then be taken as representing the statistically best-fitting calendric age. Whereas for tree-ring data with annual growth the sample distances on the calendric time-scale can be derived by direct counting of the growth rings, archaeological application is technically more complicated since it requires modelling of the age distances between the sequential variables (for example: stratified sediment layers or architectural phases of a tell site). What is invariably more demanding than the actual age modelling is, in particular, the derivation of realistic estimates for the age model certainties and to this end, the software realisation of the wiggle-matching procedure – depending on application – requires a quite large number of additional input/output procedures. The user dialogue for the presently applied method of optimising GMCWM is illustrated in Fig. IV.1. To begin, in both cases described above (dendrochronological and archaeological wiggle matching), the approach is similar in that a chi-squared distribution is fitted to the data whereby the weighted sum of squares for the two <sup>14</sup>C data sets (known and unknown calendric ages) is used to calculate the probability of the best-fitting calendric age for the study data. What complicates matters is that, due to the non-linearity of the calibration curve, there is not always a unique solution. On this specific point, in recent years it has become increasingly apparent that GMCWM results (and presumably also the results of Bayesian sequencing), even of larger data sets, may exhibit chronological quantisation similar to that of single <sup>14</sup>C ages. The existence of multiple solutions for large data sets would represent the mathematical generalisation of the multiple readings observed for single <sup>14</sup>C ages. Further, we may theoretically expect a rapid increase in the number of solutions in parallel to an increase in dating precision. This forecasting is the actual reason for our recent introduction, as mentioned above, of the more flexible non-central chi-square probability function. This problem is addressed in more detail by Weninger,<sup>226</sup> but does not require further mediation for the PMZ data, which – interestingly – do not show such multiple best-fit solutions. The reason for this is presumably the relatively straightforward structure of the calibration curve in the time window under study (6000–5400 calBC), with an extended more or less linear section 5950–5650 calBC and an s-shaped wiggle at ~5600 calBC, followed by a set of rather wobbly data 5550–5450 calBC. What we consequently observe (as shown below: Fig. IV.4), is that the relative amplitudes of the higher-order solutions (although presumably existing) are so small that they cannot be distinguished from the statistical noise. Further methodological refinements, prior to the introduction of noncentral chi-square analysis, were already implemented in earlier versions of CalPal.<sup>227</sup> These program developments, in combination, today allow utilisation of a variety of technical, graphical, and statistical structures (e.g. Excel-format data import/export, semi-automated construction of archaeological models, calculation of conditional and marginal probabilities, phase boundaries, analysis with/without sample order randomisation, and others), depending on the specific application.

<sup>224</sup> Pearson 1986.

<sup>225</sup> Weninger 1986.

<sup>226</sup> Weninger et al. 2018.

<sup>227</sup> E.g. Weninger 1986; Bronk Ramsey et al. 2001; Benz et al. 2012; Horejs et al. 2015; Jung – Weninger 2015; Krauß et al. 2017.

#### IV.5.2. Results of Gaussian Monte Carlo Wiggle Matching Application to Platia Magoula Zarkou-Data

In the present study we have used the GMCWM method as implemented in CalPal (Version 2019.5) to derive a best-fitting age-depth model for the PMZ data. As it turned out, the GMCWM-application for the PMZ data was unusually straightforward (i.e. there was no need for complex sample order randomisation, or for calibration curve correction studies, or for otherwise typically necessary outlier analysis). The necessary model construction was achieved, in an uncomplicated manner, by data entry into a standardised (CalPal-format) Excel file, wherein the required variables are listed as LabCode (character field),  $^{14}\text{C}$  age (numeric field, BP scale), standard deviation (numeric field, BP scale), and stratigraphic depth (numeric field, cm scale), along with an associated list of incrementally increasing integer values (ID1–10) to encode the lack of need for sample position randomisation during the Monte Carlo processing. This file was then read into CalPal via the ODBC interface, to produce an input/output table, with structure as illustrated in Tab. IV.2. Following some preliminary runs, needed to establish the appropriate case-specific runtime parameters and settings (e.g. time windows, uncertainty parameters, choice of linear age-depth modelling, choice of non-centrality value  $\lambda$ ), the final results were achieved for an extended (typical) runtime of 10 hrs. The statistical parameters for this run are provided in the caption to Fig. IV.4, with numeric results given in Tab. IV.2.

The input/output table with GMCWM results (Tab. IV.2) contains a stratigraphically ordered list (from young to old) of model-input  $^{14}\text{C}$  ages and corresponding metric depth. The youngest sample depth value is referenced to  $\Delta=0\text{cm}$  (compare the two columns with headings ‘ $\Delta[\text{cm}]$ ’ and ‘Depth Tell [cm]’). The column with the heading ‘BestFit’ contains the final numeric results, achieved by statistical processing in combination with a number of mathematical measures and visual-based decisions, all aimed at data reduction of the three basic statistical curves shown in Fig. IV.4. It is interesting, indeed methodologically entrancing, to note the unusually small size of the statistical uncertainties (noted at 95% confidence in the BestFit column). These uncertainties have values ranging from (min.)  $\pm 1$  yr (for ID3) to (max.)  $\pm 14$  yrs (for ID10), with the smallest values in the sequenced centre (c. ID3–7), and the largest values for the outer (oldest/youngest) regions in the sequence (ID1 and ID10). This is the natural result of the applied modelling procedure, which is based on systematic (stepwise) expansion of the data set along the calendric time-scale, such that, as a result of statistical optimisation, the possible ‘wobbling’ of the data set is at a maximum for the sequence beginning/end and is at its lowest for the central sample positions. Put differently, due to the sandwich structure of the data set, the inner dates are ‘held fast’ by the outer dates. This structure is quite acceptable, and is, indeed, an immediate consequence of the GMCWM approach. En passant, we note that such a ‘sandwich-structure’ of modelling uncertainties, with smallest values achieved in the model-centre and largest values at its begin and end, is also observable for Bayesian sequencing. Although seldom commented on by authors, the existence of such quasi inbuilt (i.e. fundamental) error-structures is easily recognised in many published OxCal-graphs.

An open question is then, of course, whether the small size of the statistical errors is indeed acceptable, and what do such small values mean? To begin, and given that similarly small errors are also observed for studies that apply the Bayesian sequencing method,<sup>228</sup> the problem is apparently not specific to GMCWM, but has a more general character. In effect, so it seems, what both methods are measuring cannot be uncritically interpreted as immediately real ‘dating uncertainties’ or ‘errors’, with traditional (probabilistic) meaning. It appears that these uncertainties (derived as they are from millions of model comparisons) have a rather more *conditional* meaning, less in terms of achieved (numeric) dating precision, and more in terms of what might better be called (semi-quantitative) ‘model robustness’. In this respect, the title of the paper by

<sup>228</sup> E.g. Sevink et al. 2011.

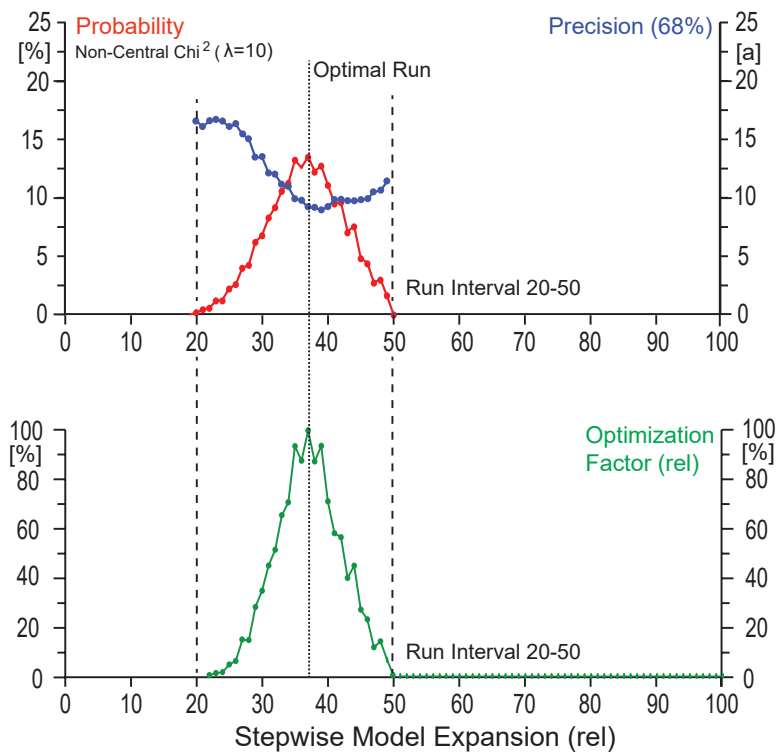


Fig. IV.4 Statistical results for GMCWM analysis of PMZ data. Upper graph: Dating probability (red) and dating precision (blue). Lower graph: Probability/precision (B. Weninger)

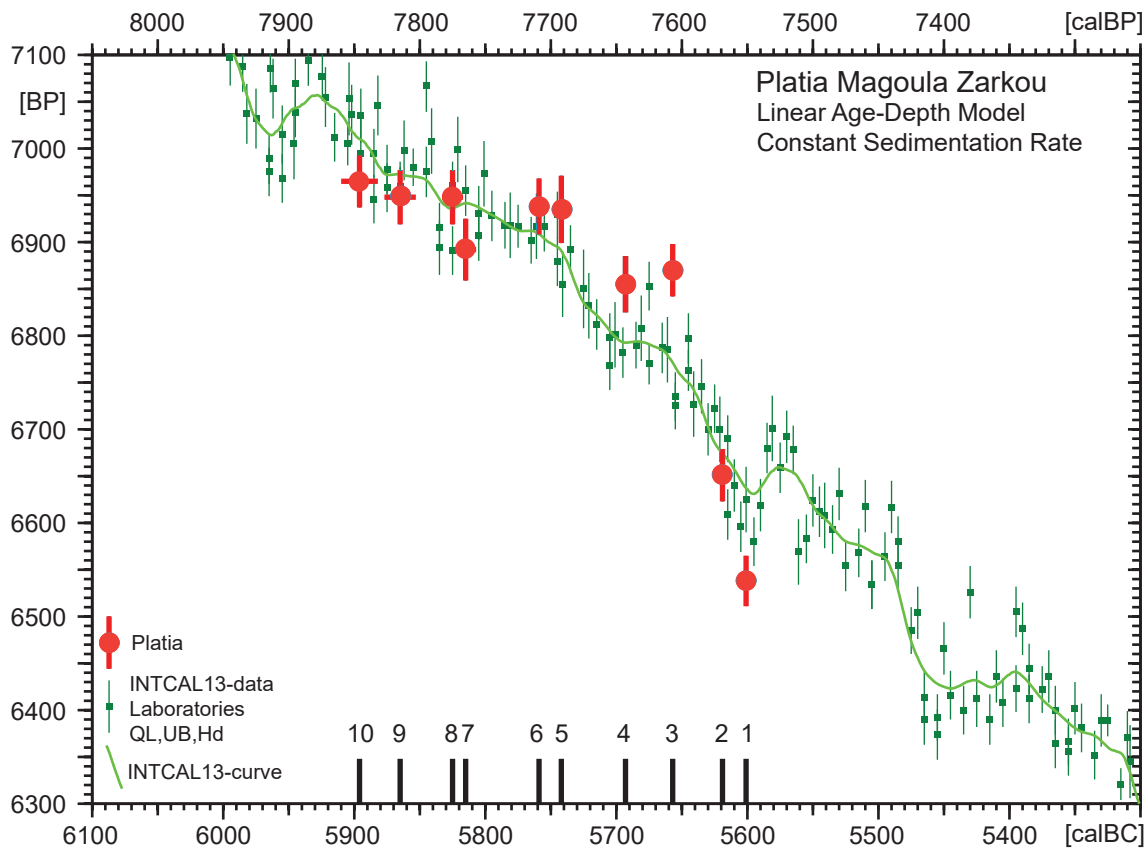


Fig. IV.5 Final GMCWM linear age-depth model with a focus on showing the relationship between the calibrated data (dots with 68%-error bars) and the tree-ring calibration curve (IntCal13, green line. ID numbers 1–10) represent dated bone samples according to Tab. IV.2.2 (B. Weninger)

Jan Sevink et al. 2011, ‘Robust date for the Bronze Age Avellino eruption (Somma-Vesuvius): 3945 ±10 calBP (1995 ±10 calBC)’ is well chosen, even though the actual meaning of the acclaimed dating precision of 10yrs itself remains to be clarified. From what is presently clear, given that such exaggeration of dating precision accords with widespread archaeological praxis in error notation, we cannot rely on the expectation that the (otherwise: nominally) standardised ‘±’ character is an abbreviation for reproducible dating uncertainty at 68% confidence. A solution to the vexing problem of how to realistically quantify the dating uncertainties that are associated with calibrated <sup>14</sup>C ages – even when derived from advanced statistical studies, e.g. via GMCWM or by Bayesian modelling – in our view represents one of the presently most important and still unresolved enigmas of archaeological <sup>14</sup>C analysis. We will return to this question in Chapter IV.6 (Discussion). Let us first take a look at the statistical properties of the GMCWM results.

The set of three graphs (Fig. IV.4: which we regularly call the ‘GMCWM statistics box’) shows the statistical functions that contribute to the final (numerically abbreviated) dating results (data input: Tab. IV.2). The technical procedures applied in deriving these functions are as follows. Following selection of an appropriate scaling factor (with the dimension cm/yr), the GMCWM algorithm expands the sample sequence (internally scaled to cm/yr) stepwise along the calendric scale, whereby at each step the corresponding probability of the age model is calculated and stored to file. As can be seen from Fig. IV.4, the stepwise model expansion is based on a fixed number of max. 100 steps (both graphs have x-axes scaled to the same 0–100 steps). In the present study, only steps 20–50 were required to achieve a useful time window, and this limitation of the time window was happily chosen to achieve a major (70%) reduction in runtime. In parallel to the width of the best-fitting calendric-scale histogram (calculated at each step by repeating the fitting process for an (optional) number of times, here: N=100) under different conditions (here: Gaussian CalScale distance error of 10 [a]; Gaussian CalCurve remeasurement per point of ±10 [BP], 68% confidence). The entire procedure was repeated for an (optionally) large number of times (here: 1000 runs), with a total runtime of 10 hrs.

The finally achieved age-depth model for the PMZ data is shown in Fig. IV.5, with numeric ‘Best Fit’ results provided in Tab. IV.2 (above).

#### IV.6. Discussion

As it turns out, the calculated best-fitting calendric ages for each of the sample positions within the sequence are so highly reproducible (‘robustness’= ±10yrs, 95% confidence) that, already above, we have expressed some reservations about the meaning of these values. In addition, we have also indicated above that the underlying puzzle has little to do with any specific properties of the applied GMCWM method. Namely, similarly small values (±10yrs) are obtained, as exemplified in the study by Sevink et al. 2011 that is based on Bayesian sequencing of some few (N=9) dates, for the Bronze Age eruption of the Avellino volcano. Note that, of the nine dated samples, only one is short-lived (leaves) and the remaining eight <sup>14</sup>C ages were processed on peat and wood.<sup>229</sup> This is not a unique result: the application of Bayesian sequencing is often claimed to be capable of deriving dating uncertainties with magnitudes in the order of some very few decades.<sup>230</sup> The reality of this acclaimed precision (or rather: its interpretation) is not immediately plausible given the many accompanying uncertainties in archaeological age models, the present need to calibrate annual-width samples on low-density and mostly decadal-scale calibration curves, the often observed over-smoothing of these calibration curves, the unexplored questions of geographic variability in carbon reservoirs, and other remaining unresolved technical aspects of <sup>14</sup>C dating, e.g.

<sup>229</sup> Sevink et al. 2011.

<sup>230</sup> E.g. Bayliss 2009; Blockley – Housley 2009.

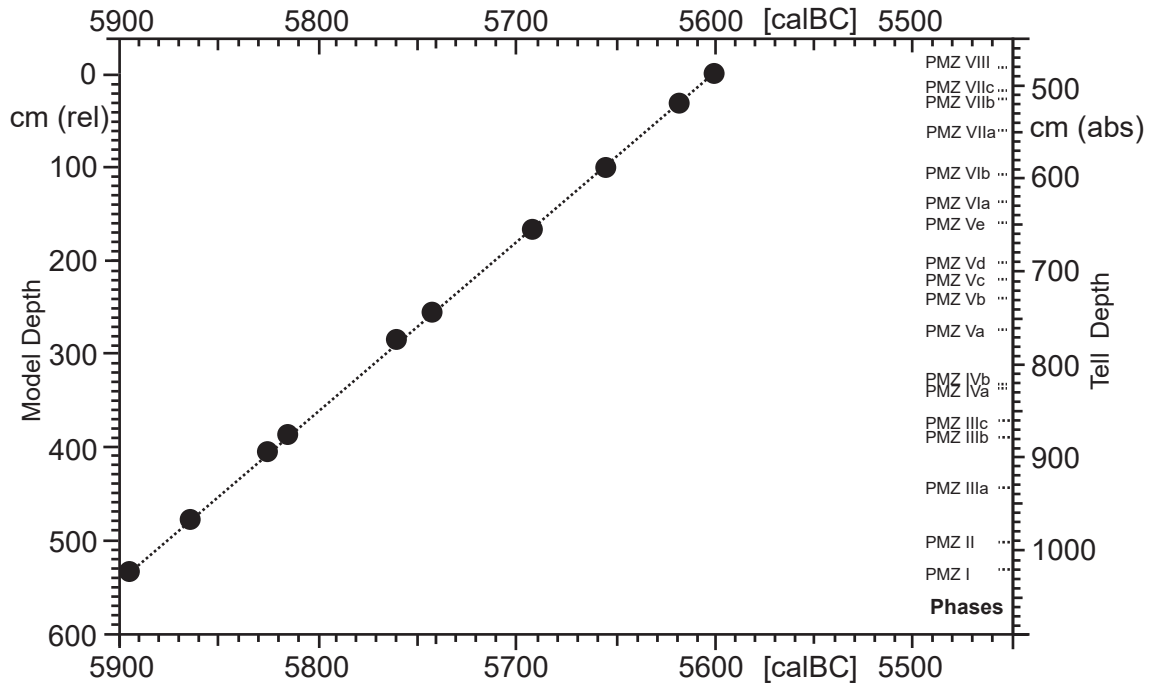


Fig. IV.6 Final GMCWM linear age-depth model with a focus on showing the age-depth-phase relationship of PMZ-stratigraphy. Modelled ages are shown as dots, representative for the respective calibrated central values of 95% intervals (cf. Tab. IV.2.2). ID numbers (1–10) represent dated bone samples (cf. Tab. IV.2.2). The statistically optimised best-fitting linear curve (dotted line) that connects the calibrated sample ages and their respective stratigraphic depth is so precise that, for simplicity, it is drawn by hand (B. Weninger)

the precision and accuracy of interlaboratory calibration. A solution is, perhaps, to simultaneously keep in mind both the technical sources of the measuring uncertainties, and their deployment in the specific scientific framework under study. Since these two contexts can be different, they can also be related to different meanings, e.g. for our understanding, use, and, in particular, our inferential application of the corresponding chronology.

In conclusion, a possible resolution to the puzzle under study is to differentiate (more clearly than is often advocated) between dating uncertainties that arise as a result of comparing the archaeological  $^{14}\text{C}$  data under study with a theoretical model, and the dating uncertainties that exist in the archaeological reality. In the present case, what we have achieved is to demonstrate (no more than) the existence (*sensu strictu*) of a linear age-depth model for the PMZ stratigraphy, and further that – with minimal (10-hour) statistical manipulation – the PMZ data are shown to be in near-perfect agreement with this model. All these nice properties (of the model) are advocated in Figs. IV.5 and IV.6. Unfortunately, what we have not yet shown is that the model really does agree (in detail) with the actual stratigraphy of PMZ, nor have we derived estimates of the amounts, sources, location, and presumable relocation of the sediments that are used in the PMZ buildings. In direct analogy, and although we have no reason to doubt the usefulness and typically enhanced quality of  $^{14}\text{C}$ -based age modelling, seldom is archaeological justification provided to conclude a posteriori why the achieved numeric precision of modelling studies should be immediately valid (beyond its theoretical justification) for the archaeological realities under study. This justification, quite generally, would require a variety of further, independent studies, but would take us into research projects presently largely beyond the reach of contemporary statistical modelling.<sup>231</sup> In

<sup>231</sup> E.g. Pettitt – Zilhão 2015.

effect, understanding our often highly simplified theoretical hopes and guesses is something very different to understanding the complexity of the real world, out there. As for the present PMZ analysis, the reconstructed extremely small variability of the sediment accumulation apparently does not correlate, at least not to any large extent, with the archaeologically defined sequence of periods and phases. Instead, the tell apparently just keeps on growing, and growing..., with an average but entirely constant accumulation rate in the order of 1 cm/yr throughout all its many periods and phases, for some overall 300 yrs. This applies, at least, as observed in the <sup>14</sup>C-sampled Trench A. More generally, for the present generation of dates and archaeological models, in particular concerning tell stratigraphies, although the calculated errors of GMCWM and Bayesian sequencing are typically in the range of a few years/decades, the underlying real dating uncertainties are presumably much larger, in the order of decades/centuries. Whether these uncertainties can be reduced through the development of more realistic and archaeo-bio-physically motivated theories of tell development, is a seemingly natural expectation, the validity of which nevertheless awaits detailed verification.

#### IV.7. Conclusions

PMZ contains ~5m of Neolithic sediments, to which 9 periods and 19 subphases have been assigned based on architectural and pottery-style analysis. In the present chapter, based on a set of N=10 <sup>14</sup>C measurements provided by the Mannheim AMS laboratory for stratified bone samples (Tab. IV.1), we have constructed, with high numeric dating precision, a linear age-depth model for the Neolithic layers. The method applied in age-depth modelling is Gaussian Monte Carlo Wiggle Matching. The results are provided in two graphs (Figs. IV.5–6), whereby Fig. IV.5 focuses on comparing the mathematically constructed optimal age-depth model with the tree-ring calibration curve IntCal13, and Fig. IV.6 demonstrates the achieved high quality of the best-fitting linear age-depth model. In Fig. IV.6 the calculated (model-based) optimal calendar age of each of the <sup>14</sup>C-dated bone samples is represented by a dot with assigned calendric-scale age uncertainty, whereby the error bar length is mostly too small to be recognisable. This age model is the result of applying a (largely standardised) Monte Carlo statistical optimisation to the PMZ data, which required some 10 hours of runtime on a standard PC, and which results in such small dating errors (often < 10 yrs, 68% confidence), that we have reason to doubt their reality. It is pointed out, finally, that an important methodological difference exists between such calculated model uncertainties and the true (but largely unknown) dating errors of the archaeological object, whereby the differences between the two dating concepts are often neglected. For the PMZ data, the calculated dating uncertainties based on extensive (empirical) Monte Carlo modelling are in the range of 1–14yrs (95% confidence; Tab. IV.2). The true (real) dating errors are likely to be in the order of 20–50yrs (95% confidence).

In a nutshell, the Neolithic tell at PMZ (depth: 500 cm) is continuously occupied from 5900 to 5600 calBC.<sup>232</sup> During its Neolithic occupation (length: 300 yrs) the average growth rate of the tell is 1.67 cm/yr (500 cm / 300 yrs), with overall quite small growth variability (e.g. from one phase to the next) estimated to be smaller than ±5 %. When calculated from cultural/demographic perspective, each of the stratigraphically discernible 19 subphases covers 17 ±1 yrs.

Hence, as illustrated in Fig. IV.6, assuming for convenience that the depth of the highest Neolithic Phase VIII is set to Zero [cm], the following age/depth formula can be used to achieve a first order estimate of the calendric age for all single finds (left scale) and/or PMZ phases (right scale):

$$\text{Calendric Age (PMZ depth)} = 0.55 * \text{depth [cm]} + 5600 \pm 20 [\text{calBC, 68\%}].$$

<sup>232</sup> For a start of the PMZ sequence see also Reingruber et al. 2017, 44; Alram-Stern – Toufexis, this volume, 615–617.

