



**HAL**  
open science

# Analytical effects on clumped isotope thermometry: Comparison of a common sample set analyzed using multiple instruments, types of standards, and standardization windows

William F. Defliese, Aradhna Tripathi

## ► To cite this version:

William F. Defliese, Aradhna Tripathi. Analytical effects on clumped isotope thermometry: Comparison of a common sample set analyzed using multiple instruments, types of standards, and standardization windows. *Rapid Communications in Mass Spectrometry*, 2020, 34 (8), pp.e8666. 10.1002/rcm.8666 . hal-02933380

**HAL Id: hal-02933380**

**<https://hal.science/hal-02933380v1>**

Submitted on 5 Oct 2021

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Defliese William F. (Orcid ID: 0000-0002-7730-4964)

## **Analytical effects on clumped isotope thermometry: Comparison of a common sample set analyzed using multiple instruments, types of standards, and standardization windows**

**William F. Defliese\***

Department of Earth, Planetary, and Space Sciences, University of California, Los Angeles, CA, USA

Current Address: School of Earth and Environmental Sciences, The University of Queensland, St Lucia, QLD, Australia

**Aradhna Tripathi**

Department of Earth, Planetary, and Space Sciences; Department of Atmospheric and Oceanic Sciences; Institute for the Environment and Sustainability; California Nanosystems Institute; University of California, Los Angeles, CA, USA

European Institute of Marine Sciences (IUEM), Université de Brest, UMR 6538, Domaines Océaniques, Rue Dumont D'Urville, and IFREMER, Plouzané, 29280 France

\* **Author for Correspondence:** [w.defliese@uq.edu.au](mailto:w.defliese@uq.edu.au)

### ABSTRACT

**Rationale:** Carbonate clumped isotope geothermometry is being increasingly applied in multiple disciplines in the geosciences. However, there are potential interlaboratory issues arising from different standardization procedures that may contribute to the multiple  $\Delta_{47}$ -temperature calibrations reported in the literature. We investigate this issue by comparing a common temperature calibration sample set across three different mass spectrometers, using multiple standardization methods.

**Methods** The same temperature calibration sample set was analyzed on three different mass spectrometers. Several standardization methods were utilized, including the use of carbonate versus gas standards, and different types of background correction were applied to the raw data.

This is the author manuscript accepted for publication and has undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the [Version of Record](#). Please cite this article as doi: [10.1002/rcm.8666](https://doi.org/10.1002/rcm.8666)

**Results** All standardization types applied resulted in statistically indistinguishable  $\Delta_{47}$ -temperature slopes, with the exception of standardization calculations that did not correct for background effects. Some instruments and standardizations showed different intercepts relative to each other. The use of carbonate standards improved comparability between different instruments relative to gas standards.

**Conclusions** Our results show that background effects are the largest factor potentially affecting  $\Delta_{47}$  results, and there may be an improvement in interlaboratory precision by using carbonate standards. Critically, all techniques utilized for standardizing  $\Delta_{47}$  results converge on a common slope as long as background effects are properly corrected. The use of carbonate standards is recommended as a component of standardization procedures.

## Introduction

Carbonate clumped isotope geothermometry is a promising isotopic technique that is being applied to fundamental problems in earth and atmospheric sciences<sup>1</sup>. The measurement of the organization of multiple rare isotopes within a molecule opens up avenues of scientific inquiry that are not accessible with single isotope systems, such as the longstanding problem of determining the respective influences of water isotopic values and carbonate growth temperatures on the oxygen isotope composition of carbonate minerals. The carbonate clumped isotope thermometer has been deployed in a growing number of applications, such as paleoclimatology<sup>2-4</sup>, sedimentology and basin analysis<sup>5-7</sup>, and tectonics<sup>8-10</sup> because it is completely independent of the sample's bulk isotopic composition, and allows temperatures to be measured directly without knowledge of the isotopic composition of water. In this application, the overexpression of  $^{13}\text{C}$ - $^{18}\text{O}$  bonds within the carbonate lattice relative to a randomized mixture of isotopes (denoted  $\Delta_{47}$ ) is controlled by the formation temperature of the carbonate mineral, assuming that it formed at equilibrium<sup>11,12</sup>.

However, measurement of multiple rare isotopes within a molecule requires very high precision, making analytical requirements a challenge compared with single rare isotope systems such as carbon ( $^{13}\text{C}/^{12}\text{C}$ ) or oxygen ( $^{18}\text{O}/^{16}\text{O}$ ). The target analytes are very rare in abundance, so high amplification of signals along with very long integration times are required. The standard procedure for carbonate samples involves liberating  $\text{CO}_2$  from solid carbonate using >100 weight percent phosphoric acid, multiple steps of cryogenic separation to remove water, and purification to remove organics and overlapping  $m/z$  contaminants via gas chromatography or use of a molecular resin trap<sup>13</sup>. Measurements of many sample/working gas cycles and/or multiple

sample replicates may be necessary in order to achieve meaningful instrumental errors that approach shot noise limits<sup>14</sup>.

Standardization of  $\Delta_{47}$  data is a complex process requiring multiple steps, and several procedures have been proposed to both provide a common reference frame between labs and to account for instrumental effects in measurement. The first effect to be recognized was a mass spectrometric correction for non-linearity, where measured  $m/z$  47 values show a dependence on  $m/z$  44 ions, as expressed in a plot of  $\Delta_{47}$  versus  $\delta^{47}$  values<sup>13-18</sup>. This effect is related to instrumental design and is caused by a secondary scattering of ions as they travel through the flight tube creating negative currents on the Faraday collectors, and the magnitude of the effect is variable through time, depending on filament age, source tuning, and other instrument specific parameters. Three different types of correction schemes have been used. An initial correction scheme for this non-linearity involves the measuring CO<sub>2</sub> gases of varying bulk isotopic composition which have been heated at 1000 °C for several hours to create a randomized distribution of isotopes (originally proposed by Ghosh et al)<sup>13</sup>. A suite of these gases would be measured, the non-linearity measured, and subtracted out of the final sample  $\Delta_{47}$  values as, by definition, samples that have been equilibrated at the same temperature have the same  $\Delta_{47}$  value, even if they have different bulk isotopic ratios. This procedure must be done continually throughout the measurement period to account for changes in the non-linearity correction.

A variant of this approach replaces the equilibrated gases with two carbonate standards of differing bulk isotopic ratios that have been heated to a high temperature and then rapidly quenched, ensuring similar  $\Delta_{47}$  values despite differing  $\delta^{47}$  values<sup>18</sup>. A more recent approach termed the pressure baseline (PBL) correction involves analyzing the peak shape produced during a slow magnet scan at various gas pressures (corresponding to different signal sizes), identifying the magnitude of negative currents impacting the peak shapes, and subtracting these currents from future measurements<sup>15,17,19</sup>. This procedure has the benefit of being less time consuming to implement; however, it must be performed daily. Many labs use some combination of the above approaches, and often a PBL is applied in addition to equilibrated gases or paired carbonate standards.

The second issue realized was that the non-linearity correction did not fully account for inter-lab differences in  $\Delta_{47}$ . Specifically, the non-linearity correction did not fully anchor the  $\Delta_{47}$  scale, with scale compression or stretching observed. To provide a more complete standardization scheme, Dennis et al<sup>16</sup> proposed that gas or carbonate standard values be defined absolutely to create an absolute reference frame<sup>16</sup>, using one of two methods. The first approach is to equilibrate CO<sub>2</sub> gases at variable temperatures. In addition to heated gas standards, equilibrated CO<sub>2</sub> gas standards could be produced at different temperatures with variable bulk composition by taking advantage of isotopic exchange that occurs with water in the H<sub>2</sub>O + CO<sub>2</sub>

$\leftrightarrow$   $\text{H}_2\text{CO}_3$  reaction. These gases, equilibrated at different temperatures, are used for the non-linearity correction, and then each assigned an absolute  $\Delta_{47}$  value based on theoretical predictions of values given the temperature of equilibrium<sup>20</sup>, so that an absolute reference frame (ARF) can be built to account for scale compression or stretching of  $\Delta_{47}$  values. A carbonate standard-based method of standardization was also proposed<sup>16,18</sup> that involves the use of in-house carbonate standards of varying bulk isotopic composition and  $\Delta_{47}$  values to build the ARF. Carbonate standards can be used alone without the use of supplementary equilibrated gases<sup>21,22</sup>, or can be used as an additional correction (secondary transfer function, or STF) on top of the use of equilibrated gases. The carbonate standards approach has the benefit that it does not require the time-consuming step of manufacturing equilibrated gases, and furthermore treats all standards identically to samples (which may or may not be the case with equilibrated gases, depending on lab group and procedure), but it still requires analyzing multiple standards during the run, reducing sample throughput.

The identification of different potential sources of inter-laboratory offsets has led to new methods being proposed for standardization, as described above. However, the impact of the lack of a consistent practice in standardization is unclear. It is uncertain whether different labs need to use the same standardization methods to achieve consistency, or whether there are multiple valid methods of standardization. Thus, consensus on best practice has not yet been reached. Yet quantifying the impacts of standardization method on resultant  $\Delta_{47}$  values is critical for resolving why there are still discrepancies in  $\Delta_{47}$ -temperature calibrations, and for quantifying the magnitude of equilibrium and kinetic isotope effects. There are three end-member hypotheses for the origin of discrepancies in reported calibrations: 1) there are measurable equilibrium and kinetic isotope effects that give rise to the range of values observed in the literature<sup>11,23-27</sup>, 2) inter-laboratory differences in sample digestion give rise to the variation observed in  $\Delta_{47}$ <sup>28-31</sup>, and 3) inter-laboratory differences in  $\Delta_{47}$  measurement and/or standardization give rise to the variation observed in  $\Delta_{47}$ <sup>32,33</sup>.

This paper represents a contribution to partially testing hypotheses 2 and 3, with the caveat that all three hypotheses may in fact be true. Here we report the results of a multi-year study where we quantify the differences that can be caused by the use of different standardization techniques. We achieve this by analyzing a common set of synthetic calcite and aragonite samples on two different mass spectrometers and reanalyzing results from a published dataset<sup>28</sup> from a third mass spectrometer. The three mass spectrometers each have a different preparation line. Further, we analyze the effects of different standardization methods, including using gas standards only, carbonate standards only, and a mixture of gas and carbonate standards for corrections, and we explore the impact of using different standardization intervals for corrections. We report the consequent impacts on calculated  $\Delta_{47}$ -temperature calibrations, and

show that carefully chosen measurement windows yield results for the same set of samples that are largely comparable across multiple instruments.

## Methods

### *Samples*

Calcite and aragonite samples synthesized and measured in the Lohmann Stable Isotope Lab at the University of Michigan (Ann Arbor, MI, USA) were reported in Defliese et al<sup>28</sup>, and these data were reanalyzed for this study. In the case of the 5 °C calcite sample, we used a parallel precipitate that was grown at using the same apparatus and methods as the sample reported in Defliese et al<sup>28</sup>; however; this sample was not used for that study. We measured splits of these samples on multiple instruments in the Tripathi Stable Isotope Lab at UCLA (Los Angeles, CA, USA). These carbonates consist of calcite and aragonite that were grown at controlled temperatures using the ‘passive degassing’ technique. Samples were grown in a 4-L beaker at 5, 25, 50, and 70 °C, while being allowed to exchange freely with atmosphere. Some of the samples used in Defliese et al<sup>28</sup> were exhausted, including all replicates at 50 °C, so they were not included in this study. The mineralogy of each sample used is given in Table 1. For full details on the synthesis of these materials, please refer to Defliese et al<sup>28</sup>.

### *Mass spectrometers and operating procedures*

For this study, we used data from three mass spectrometers, two of which are housed in the Tripathi Lab at UCLA, and one of which is located in the Lohmann Lab at the University of Michigan (UMich). Details of sample digestion and mass spectrometric analysis are described below.

#### *University of Michigan, Ann Arbor*

##### Sample Digestion

Carbonate samples and standard materials (equilibrated gases) were prepared through use of a custom-built offline manual prep line. Briefly, carbonates were digested in 105 weight percent H<sub>3</sub>PO<sub>4</sub> in a common acid bath apparatus, held at 75 °C using an enclosed circulating water bath. CO<sub>2</sub> and the co-produced H<sub>2</sub>O were frozen in a U-trap simultaneous with digestion to avoid potential equilibration with H<sub>2</sub>O and scrambling of <sup>13</sup>C-<sup>18</sup>O bonds. CO<sub>2</sub> gases were subject to two stages of cryogenic separation using a liquid nitrogen/1-propanol slush held at -90 °C to remove CO<sub>2</sub>. Dehydrated CO<sub>2</sub> gas was further cleaned to remove organic and other potentially overlapping contaminants by use of a gas chromatograph (Supelco Supel-Q PLOT, 30m x 0.53 mm; Sigma-Aldrich, MO, USA; for samples prior to July 2011) or U-trap filled with PoraPak Q

resin (Waters Corp., Milford, MA, USA; 50-80 mesh, for samples in July 2011 and after). Samples cleaned using the gas chromatograph were carried using an ultra-high-purity He stream at 7 mL/min at a temperature of -20°C, and collected downstream of the gas chromatograph in a multi-loop trap immersed in liquid N<sub>2</sub>. Samples cleaned using the PoraPak were frozen into a cold finger upstream of the PoraPak U-trap, and allowed to warm to room temperature. CO<sub>2</sub> was equilibrated with the PoraPak for 5 minutes at -30 °C, and then liquid nitrogen was placed on a cold finger downstream of the PoraPak U-trap to draw the CO<sub>2</sub> through the trap. Yields were checked by manometer before and after cleaning to ensure complete transfer; any sample that was incomplete was discarded. Cleaned CO<sub>2</sub> was transferred to a glass cold finger for storage, and was analyzed on the mass spectrometer within 24 hours of cleaning.

#### Mass spectrometry

The UMich instrument is a Thermo Fisher Scientific (Bremen, Germany) MAT 253 isotope ratio mass spectrometer. It has been modified to simultaneously collect masses 44-49<sup>14</sup>, with the following resistors on each Faraday cup: 3 x 10<sup>8</sup>, 3 x 10<sup>10</sup>, and 10<sup>11</sup> Ω for masses 44, 45, and 46, and 10<sup>12</sup> Ω resistors for masses 47-49. Analyses were performed in dual inlet mode using standard stainless-steel capillaries, and the pressure was balanced between each acquisition to achieve a signal of 16V on the mass 44 cup for both sample and standard bellows. Each replicate measurement had 8 acquisitions, with each acquisition consisting of 10 sample-standard comparisons, with an 8 second integration time and 16 second changeover delay. Each acquisition was also preceded by peak centering and a 120 second background delay. The total integration time for each replicate was 640 seconds.

#### Types of standards run

The Δ<sub>47</sub> results for the UMich instrument were standardized to the Absolute Reference Frame<sup>16</sup> by comparison with CO<sub>2</sub> gases of varying bulk isotopic composition which were equilibrated at 25 °C and 50 °C by exchange with water, and at 1000 °C by heating in an oven for a minimum of 2 hours. In typical operation, two equilibrated gases were run per day. All equilibrated gases were subject to the same cleaning and handling procedure as carbonate samples, as discussed above, except that they were not reacted with phosphoric acid.

*University of California, Los Angeles*

#### Sample Digestion

Both instruments at UCLA are coupled to custom-built automatic reaction and prep lines, which operate identically, though they differ slightly in dimension and layout. Each line uses a common acid bath reaction chamber held at 90 °C by use of a heating element and thermocouple. The resultant CO<sub>2</sub> and H<sub>2</sub>O pass through a water trap held at ~-50 °C by a mixture of ethanol and

dry ice, and are frozen into a U-trap by liquid nitrogen throughout the duration of the reaction. CO<sub>2</sub> is then entrained in a stream of ultra-high-purity He, and cleaned by use of a gas chromatograph filled with PoraPak Q (50-80 mesh) held at -20 °C. The cleaned CO<sub>2</sub> is collected downstream of the gas chromatograph and passed through a second water trap before being frozen in a multi-loop trap by liquid nitrogen. CO<sub>2</sub> is cryogenically cleaned an additional time with a dry ice/ethanol slush to remove water, and passed to the mass spectrometer for analysis. Equilibrated gases were introduced by use of an external port that bypassed the common acid bath; otherwise they were cleaned and processed in an identical manner to carbonate samples and standards.

### Mass spectrometry

The two instruments at UCLA are a Thermo Fisher Scientific MAT 253 mass spectrometer and a Nu Instruments (Wrexham, UK) Perspective IS mass spectrometer. The MAT 253 has been modified to simultaneously collect masses 44-49, as described in Huntington et al<sup>14</sup>. The signal is passed through resistors of  $3 \times 10^8$ ,  $3 \times 10^{10}$ , and  $10^{11} \Omega$  for masses 44, 45, and 46, while  $10^{12} \Omega$  resistors are used for masses 47-49. Analyses were performed in dual inlet mode using nickel capillaries, and the pressure was balanced between each acquisition to achieve a signal of 16V on the mass 44 cup for both sample and standard bellows. Each replicate measurement had 9 acquisitions, with each acquisition consisting of 10 sample-standard comparisons, with an 8 second integration time and 16 second changeover delay. The total integration time for each replicate was 720 seconds.

The Nu Instruments Perspective IS mass spectrometer at UCLA is a relatively new design, with several differences from the MAT 253. It has a Nier-type ion source and magnetic sector; however, it features a pair of quadratic lenses after the magnetic sector, and cups 47, 48, and 49 are shielded by small electrostatic sector analyzers, which can be tuned to screen out secondary background ions. The resistors are the same as for the MAT 253. Analyses were performed in dual inlet mode using stainless steel capillaries. Measurement of each replicate consisted of 80 cycles of sample-standard comparison, each with 20 seconds integration time and a 6 second changeover delay, for a total integration time of 1600 seconds per replicate. The pressure on the bellows was continuously adjusted after each changeover using the previous acquisition as feedback using a Newtonian zeroing technique, and balanced to achieve a signal of 16V on the mass 44 cup for both sample and standard bellows. This is a key difference from the MAT 253, which only pressure balances after every 10 changeovers as part of a separate routine that is not part of the acquisition cycle.

### Types of standards run

The  $\Delta_{47}$  results for both UCLA instruments were standardized to the Absolute Reference Frame<sup>16</sup> by comparison with CO<sub>2</sub> gases of varying bulk isotopic composition which were equilibrated at 25 °C by exchange with water, and at 1000 °C by heating in an oven for a minimum of 2 hours, and by the use of carbonate standards (ETH 1-4, as well as in-house standards Carrara Marble, Carmel Chalk, Veinstrom, and TV03). In typical operation, one equilibrated gas and two carbonate standards were run per day. All equilibrated gases were subject to the same cleaning and handling procedure as carbonate samples/standards, as discussed above, except that they were not reacted with phosphoric acid.

### ***Data Processing and Standardization***

*University of Michigan*

We reprocessed the UMich data that was previously published in Defliese et al<sup>28</sup>. With the exception of one variation, exact the same data processing techniques have been used. Following recent suggestions<sup>34,35</sup>, we recalculated isotope ratios using the isotope parameter set of Brand et al<sup>36</sup>, in contrast to using Gonfiantini et al<sup>37</sup> which was originally used in Defliese et al<sup>28</sup>. Data was placed on the Absolute Reference Frame by use of equilibrated gases prepared at 25, 50, and 1000 °C, as described previously. No pressure baseline corrections were carried out on this instrument. In typical operation, one 1000 °C gas and one 25 or 50 °C gas were run daily, for a total of two equilibrated gases analyzed each day. The Absolute Reference Frame was then calculated daily, using a moving window of the previous 14 analyzed equilibrated gases, which approximated a one-week window. We did not calculate the effects of a longer standardization window for the UMich instrument, as this instrument is not a dedicated clumped isotope device, and clumped isotope measurements were conducted for two- to three-week periods every 4-6 months. We found that the reference frame calibrations were unstable and required constant monitoring to provide a stable standardization to the ARF. All samples were treated with an acid fractionation factor of 0.062 ‰, following Defliese et al<sup>28</sup>, and acid digestion at 75 °C. Stable isotope ratios for these samples are listed in Table 1.

*University of California, Los Angeles*

At UCLA, data was acquired over a period of six months on each instrument. We then constructed reference frames using 10 different standardization techniques, using aliquots of the same set of samples and standards, as described below. All isotope ratios were calculated using the isotope parameters of Brand et al<sup>36</sup>, and an acid fractionation factor of 0.082 ‰ was applied to all samples, following Defliese et al<sup>28</sup>, and acid digestion at 90 °C. We used a mixture of the ETH standards provided by Stefano Bernasconi (ETH, Zurich, Switzerland) and in-house UCLA

standards Carrara Marble, Carmel Chalk, Veinstrom, and TV03 as carbonate standards for this exercise. For the purpose of this exercise, we used the values for the ETH standards published in Müller et al<sup>38</sup>. The accepted values of the in-house standards are listed in the supporting information.

Historically what has been done at UCLA is to monitor standard drift and choose standardization intervals that reflect instrument stability, by constantly monitoring the reference frame. Typically, standardization windows range from several weeks to months in time. For this study, we chose two end-member standardization intervals to construct the Absolute Reference Frame, with each type of standardization window applied to both instruments. The first type of standardization interval consisted of approximately one-month intervals, where all data collected within each month were binned together and used to construct the ARF. The second type used a 6-month window, where all data collected within the entire 6 months were binned together to construct the ARF. This allowed us to assess the stability of the reference frame through time.

Five different methods were used to construct the Absolute Reference Frame (ARF) for each instrument and standardization window at UCLA. It should be noted that we did not perform pressure baseline corrections for any of these standardization methods. The first method used only the equilibrated gas data to create non-linearity correction and the ARF, as described in Dennis et al<sup>16</sup>. All carbonate standards are ignored in the analysis. The second method used equilibrated gases to construct the non-linearity correction, and used equilibrated gases and carbonate standards to create the ARF. The third method used the equilibrated gases to do the non-linearity correction, and then used carbonate standards for the ARF. The fourth method ignored the non-linearity correction, and used the carbonate standards alone to construct the ARF. The fifth method used the assigned values of ETH 1 and 2 to do the non-linearity correction, and used all the carbonate standards to create the ARF. We note for the MAT 253 at UCLA, there were not enough carbonate ETH standards run in the latter half of the 6-month acquisition window to robustly determine the fifth type of standardization for the shorter time intervals. Therefore, the last method has not been applied in the 1-month standardization series, resulting in 9 standardization techniques. A summary of each technique used is listed in Table 2, with a numerical code given for each unique combination of ARF construction and correction window length. We will refer to this code in the text, where appropriate.

## Results

A Table with all samples and standards used in this study is available in the supporting information. Four of the original samples from Defliese et al<sup>28</sup> were analyzed on the 253

instrument at UCLA (25 °C calcite, and 5, 25, and 70 °C aragonite), versus 5 on the Perspective IS (5 and 25 °C calcite, and 5, 25, and 70 °C aragonite).

The various temperature-versus- $\Delta_{47}$  regressions produced are listed in Table 3 (for the 253 instrument at UCLA) and Table 4 (for the Perspective IS instrument at UCLA). These regressions are compared with the data originally analyzed at the University of Michigan, where new regressions for the Defliese et al<sup>28</sup> data are generated using only the subset of samples concurrently analyzed at UCLA. Defliese et al<sup>28</sup> originally reported the following calibration, using the Gonfiantini et al<sup>37</sup> parameter set and standardization using equilibrated gases only (method 6 in our Table 2):

$$\Delta_{47} = \frac{34841 \pm 2293}{T^2} + 0.3031 \pm 0.0244$$

where T is the temperature in Kelvin. Using the parameter set of Brand et al<sup>36</sup> and equilibrated gases (method 6), the calculated regression for the University of Michigan MAT 253 is slightly different, though not at a statistically significant level ( $p = 0.4579$ ):

$$\Delta_{47} = \frac{34283 \pm 1774}{T^2} + 0.2870 \pm 0.0189$$

Subsetting only the four data points analyzed on the UCLA MAT 253 with the Brand et al<sup>36</sup> parameter set and equilibrated gases (method 6), the calculated regression is also statistically identical ( $p = 0.931$ ) to the original UM calibration:

$$\Delta_{47} = \frac{36693 \pm 2899}{T^2} + 0.2612 \pm 0.0322$$

Subsetting only the five data points analyzed on the UCLA Perspective IS with the Brand et al<sup>36</sup> parameter set and equilibrated gases (method 6), the calculated regression is similarly statistically identical ( $p = 0.7753$ ):

$$\Delta_{47} = \frac{35254 \pm 2476}{T^2} + 0.2751 \pm 0.0284$$

Notably, the addition or subtraction of various data points slightly affects the slope, intercept, and error of each regression, although they all overlap. This is not surprising, as they were all produced and analyzed in the same manner using the same equipment, and it shows that the original work of Defliese et al<sup>28</sup> is internally consistent. These serve as useful comparisons with the regressions calculated using the various standardization techniques, and allow a comparison with the data generated at UCLA.

## Discussion

### Effects of Different Standardization Techniques

The different standardization techniques employed in this study reveal that  $\Delta_{47}$  values are sensitive to standardization practices. However, our analysis reveals that while there are changes with different techniques, the majority of them produce temperature regressions that are statistically identical to each other. To analyze this, for each instrument we performed ANCOVA analysis of each regression compared with every other regression on that instrument. The resulting p-values are presented in Tables 5 and 6 for the Perspective IS and MAT 253, respectively. Critically, the majority of standardization techniques produced regressions that were statistically identical (p-value for differences in slope/intercept of greater than 0.10, reject separate regression hypothesis), with some exceptions that warrant further discussion.

The regressions produced using the Perspective IS instrument are remarkably similar to each other, regardless of the standardization technique used. In fact, all 10 techniques are statistically identical (all p-values greater than 0.1, reject separate regression hypothesis), as shown in Table 5. Most standardization techniques also yield results that are statistically identical to the reprocessed Defliese et al<sup>28</sup> data, with all slopes statistically indistinguishable, and only the intercepts significantly different at the 95% level (Figure 1). These differences in intercept are of the order of the differences reported for the error on acid digestion<sup>23,33</sup> with an average offset of 0.022‰ in this dataset, and may reflect slight differences in acid temperature and/or other measurement discrepancies. It should be noted that these discrepancies are only detectable at the 95% confidence level because the reproducibilities of the Perspective IS and UMich 253 instruments are quite high, as evidenced by the low measurement error for each. The Perspective IS instrument is very stable throughout the whole study interval, as shown by the fact that there is no significant difference between any of the internally generated regressions, regardless of the standardization window (i.e., stability on a timescale of 6 months) or technique employed. It should be noted that this was a very stable time in the instrument's performance, with no downtime due to repairs or maintenance.

The regressions produced using the MAT 253 instrument are not as similar to each other as the Perspective IS, as shown in Table 6. It should be noted that the data on the MAT 253 shows much more scatter in general (as evidenced by much larger errors on temperature regressions) than either the Perspective IS or the University of Michigan MAT 253, which may be due to having analyzed only 4 samples on this instrument compared with 5 on the Perspective IS. As the missing sample is one of the cold end-members in our comparison (5 °C calcite) it may have a large influence on the final results.

There are two regressions that are statistically different from the other 7 produced, and both of them are ARF constructions that ignore the non-linearity correction (methods 4 and 9 in Table 2). There is also a difference between the intercept of the DeFliese et al<sup>28</sup> data and the regressions produced on the UCLA MAT 253 (Figure 2), although the slopes of the regressions are statistically identical, as long as the non-linearity correction is performed in some manner. Similar to the Perspective IS, the MAT 253 produces internally consistent regressions when non-linearity corrections are preformed, regardless of the time window (i.e., 1-month versus 6-month window) for standardization.

We also compared the two UCLA instruments with each other, focusing on the 1-month standardization windows (methods 6-9) as this is more reflective of real-world practices than using a 6-month window. For this exercise we only used the 4 samples analyzed in common between the MAT 253 and Perspective IS, excluding the 5<sup>th</sup> sample (5 °C calcite) analyzed on the Perspective IS. The resulting regressions produce slopes that are statistically identical (Table 7), except when non-linearity is ignored on the MAT 253 (method 9), although the intercepts are often different between the two instruments. The cause of this discrepancy is difficult to identify, as each instrument was standardized in an identical manner, using exactly the same standards and procedures for each. There are physical differences between the two extraction lines and mass spectrometers, but both operate on the same principles and should produce similar results when properly standardized. A possible cause of this discrepancy could be related to the phosphoric acid digestions on each instrument – any difference in acid temperature/density/volume could cause an offset in intercept due to raising or lowering the  $\Delta_{47}$  value of each sample and carbonate standards, but would not affect the slopes of the resulting calibrations, assuming that the acid is at a constant temperature/density/volume through time.

If it is in fact somehow related to the acid digestion process, the standardization techniques that rely on carbonate standards will produce regressions that are more in agreement with each other across the two instruments, which is exactly what our data reveals. The MAT 253 standardization technique, which relies on equilibrated gases to do the non-linearity correction but carbonate standards for the ARF (method #8), produces a regression that is most similar to the Perspective IS regressions, producing statistically identical regressions, compared with all other standardization techniques that produce regressions differing at 95% (or higher) confidence. This shows that using carbonate standards can improve the reproducibility between different instrumentation as carbonate standards are treated identically to sample unknowns<sup>18,21,22</sup>, whereas equilibrated gases are not digested by phosphoric acid, leaving an additional variable that is not accounted for in the standardization process.

Our results also show the importance of the approach used to correct for non-linearity, as our standardization schemes that did not address non-linearity produced the most divergent

results, similar to Fiebig et al<sup>19</sup>. Correcting for non-linearity can be done in one of three ways: 1) using equilibrated gases, following Dennis et al<sup>16</sup>; 2) using two (or more) carbonate standards than have differing bulk isotopic ratios, but identical  $\Delta_{47}$  values such as ETH-1 and ETH-2<sup>22,38</sup>; and 3) performing a pressure baseline correction<sup>15,17</sup>. In this study, we used both the 1<sup>st</sup> and 2<sup>nd</sup> methods, which produced results that were statistically identical to each other for both instruments. It should be noted that we observed the greatest agreement between regressions when we used equilibrated gases alone to perform the non-linearity correction, and then carbonate standards alone to build the remaining portion of the ARF (method 8 in Table 2). We suspect that this may reflect the more limited isotopic range encompassed by ETH-1 and ETH-2 versus equilibrated gases, which can potentially cover a much larger range of  $\delta^{13}\text{C}$  and  $\delta^{18}\text{O}$  values when calculating the non-linearity correction slope, although we emphasize that either method is adequate per our analysis.

Our data shows the importance of performing this correction, as our standardization procedures that intentionally disregarded the correction (methods 4 and 9 in Table 2) produced regressions that were statistically different from all others, being the only ones that produced a steeper ‘Ghosh’ style slope similar to that reported in Ghosh et al<sup>13</sup>. The version of the MAT 253 used at UCLA (and University of Michigan) has been shown to be particularly susceptible to these issues<sup>14</sup>, although it should be noted that they are easily resolved with proper procedures and care during operation and standardization. Newer versions of the MAT 253 (such as the 253+, not evaluated here) have features designed to minimize scattering and background effects. The Perspective IS does not seem to suffer anywhere near as much from scattering effects as shown by the fact that all standardization techniques produced similar regressions, which is due to the different geometry and use of electrostatic analyzers to ‘shield’ the Faraday cups from background ions. Although the non-linearity correction is already a standard part of clumped isotope procedures, we emphasize that the non-linearity correction is of critical importance for clumped isotope measurements, and is easy to implement using any of the three methods listed above.

We did not find a significant influence of calibration window (1-month versus 6-month) on regression slope for either UCLA instrument (MAT 253 or Perspective IS), although we did observe offsets in intercept. This surprised us to some degree, as typically instruments do not display such high stability through time, as exemplified by the common slope. Typically, some instrumental conditions will change over the period of several months (downtime for maintenance, power outages, aging of filament, retune focus parameters, etc.) which necessitate the use of a new calibration window. The choice of calibration window can vary significantly from instrument to instrument depending on operating conditions<sup>13,14</sup>, and as such all laboratories should regularly monitor the calibration by continuing to analyze standards, regardless of

apparent stability. Instrumental stability can only be assessed in hindsight, as conditions could change at any moment during a run period. We emphasize this, as while our 6-month calibration window produced results that were fairly similar to the UM regression, the shorter 1-month windows produced better agreement.

### Implications for Prior Work and Multiple Temperature Calibrations

One of the most puzzling issues in the clumped isotope literature has been the problem of multiple  $\Delta_{47}$ -temperature calibrations, which display a wide range in temperature sensitivity. The first two published calibrations<sup>13,39</sup> had almost a 50% difference from each other, and this has been referred to as the ‘two slope’ or “steep and shallow slope” problem. Many studies have attempted to address the underlying cause of these discrepancies, investigating factors such as mineralogy<sup>28-30,33,40-44</sup>, precipitation rate<sup>24,29</sup>, acid digestion temperature<sup>23,28-30,33,38</sup>, the use of varying isotopic constants<sup>32-35</sup>, and the use of natural versus synthetic materials<sup>28</sup>, in addition to other procedural and methodological issues. No single factor has yet been found as the cause of this historical discrepancy in the literature. It is important to note that the ‘two slope’ problem represents the outliers, and recently published calibrations show much better alignment with each other (i.e., Petersen et al<sup>33</sup>, and references within), perhaps due to better standardization and methodological improvements. The emerging consensus on a ‘true’  $\Delta_{47}$ -temperature is a major achievement for clumped isotope thermometry, although there remains work to be done.

This study does not present a solution for a ‘true’  $\Delta_{47}$ -temperature relationship; however, it does allow us to offer some insight into the nature of the historical ‘two slope’ problem. Our results show that even when vastly different standardization procedures and instrumentation are used, the same sample materials generate statistically identical  $\Delta_{47}$ -temperature slopes, with the important exception of procedures that ignore the non-linearity correction. Given that the non-linearity correction has been a standard part of clumped isotope practice for some time<sup>14,16</sup>, this should not result in any changes to current practices for the vast majority of labs worldwide. This work suggests that as long as labs are following a rigorous standardization procedure to construct a proper reference frame<sup>14,16,21,45</sup> and other lab protocols, differences in measured  $\Delta_{47}$ -temperature slopes probably reflect real differences in the samples themselves<sup>25,46</sup>, and are largely not affected by laboratory practices.

Our data also shows it may be possible to ‘steepen’ or ‘shallow’ the slope of a  $\Delta_{47}$ -temperature regression by undercorrecting or overcorrecting for non-linearity. This is particularly the case if there is a linear relationship between the sample’s  $\delta^{47}$  value (related to  $\delta^{13}\text{C}$  and  $\delta^{18}\text{O}$  values) and  $\Delta_{47}$  value, i.e. if higher temperature samples have more negative  $\delta^{47}$  values. This can result in a situation where all the calibration samples, even when uncorrected for non-linearity, form a linear relationship with respect to formation temperature and  $\Delta_{47}$  value, as

exemplified by the carbonates produced by Defliese et al<sup>28</sup> and used in this study. As such, the slope of the non-linearity correction plays a very important role in determining the slope of the final  $\Delta_{47}$ -temperature relationship, and under- or overcorrecting will affect the final slope. In the case of this study, as the non-linearity slope is positive, when uncorrected it results in a ‘steepening’ of the calculated temperature relationship. This relationship arises because the same sources of bicarbonate and deionized water were used to precipitate the Defliese et al<sup>28</sup> carbonates, so the  $\delta^{47}$  values (and particularly the  $\delta^{18}\text{O}$  values) become more negative at higher temperatures. Any laboratory-grown samples that use the same carbon and water sources for all temperatures could be sensitive to the same effects, and this phenomenon may be a possible explanation for the origin of the historical ‘two slope’ problem.

### Update to Defliese et al (2015) Temperature Calibration

As part of this work, we recalculated the original data presented in Defliese et al<sup>28</sup> using the Brand et al<sup>36</sup> isotopic parameters as recommended<sup>34,35</sup>. We can combine those results with measurements of those same samples from UCLA used in this study to present updated  $\Delta_{47}$ -temperature calibrations (Figure 3). We used the  $\Delta_{47}$  values produced by standardization with equilibrated gases and short time intervals for data from the University of Michigan (method 6 in Table 2), as carbonate standards were not analyzed at the University of Michigan. For the data from UCLA, we used the data standardized with both equilibrated gases and carbonate standards, and short time intervals (method 7 in Table 2). This analysis produced the following calibrations:

$$\Delta_{47} = \frac{36620 \pm 4449}{T^2} + 0.2999 \pm 0.0489 \text{ (Aragonite)}$$

$$\Delta_{47} = \frac{34720 \pm 5306}{T^2} + 0.3081 \pm 0.0590 \text{ (Calcite)}$$

$$\Delta_{47} = \frac{35810 \pm 3404}{T^2} + 0.3038 \pm 0.0376 \text{ (All Data)}$$

All three of these calibrations are within error of each other, and are generally slightly steeper in slope (but still within error) than the calibration produced solely using data from the University of Michigan. Notably, recently published calibration studies<sup>23,29,33,40</sup> have steeper slopes and smaller intercepts that do not overlap with these calibrations, suggesting that there are probably real differences in the absolute  $\Delta_{47}$ -temperature relationships among these carbonates, given that the Perspective IS (which should be immune to issues arising from the non-linearity correction) shows the same result as the UMich MAT 253 (Figure 3). If there is a universal calibration, it is likely that calibration studies encompassing a larger number of samples, measurements, and range of temperatures, such as Petersen et al<sup>33</sup>, better capture the true  $\Delta_{47}$ -temperature relationship<sup>45</sup>.

## Recommendations for Future Work

Carbonate clumped isotope measurements are in the process of becoming a mature technique, with new laboratories opening worldwide, each with slightly varying instrumental setup and procedures. Our results show that convergent results can be achieved using a variety of instrumentation and standardization procedures, which bears good news for the future application of the technique. However, our results do highlight that rigorous standardization is necessary to produce accurate data. Carefully chosen measurement windows using carbonate standards or a mix of carbonate and gas standards are ideal, and can yield results for the same set of samples that are comparable across multiple instruments.

In particular, the non-linearity correction must be monitored and corrected for, via one of the three techniques discussed above, and the ARF carefully constructed with reference frame variability characterized. In this analysis, conditions were typically more stable over a one month time frame (UCLA) than a 6 month time frame. The new generation of instruments improve long-term stability of the reference frame.

Our data also shows that the use of carbonate standards may offer an improvement in interlaboratory standardization (in our case, intralaboratory) versus the use of equilibrated gases alone, as demonstrated by the improvement in regression significance when using carbonate standards to build the ARF. This will probably vary depending on the skill of the worker making the equilibrated gases, as well as the physical setup of laboratory apparatus and instrumentation. As such, carbonate standards offer the advantage of being processed in a manner identical to sample materials, eliminating the unknown effect of phosphoric acid digestion, and limiting the ‘human factor’ associated with making equilibrated gases on manual vacuum lines. They are also much easier to transport and provide direct comparison across laboratories compared to equilibrated gases, which require a separate specialized vacuum line to prepare. As such, in step with other recent proposals<sup>21</sup>, we recommend the frequent use of carbonate standards within each laboratory’s standardization scheme, whether in conjunction with equilibrated gases or replacing them entirely.

## Conclusions

As clumped isotope geochemistry has evolved over time, and the number of users has expanded, there have been many different standardization techniques applied. We evaluated the effects of several of these different techniques on a common set of synthetic carbonates, and find that there can be significant effects caused by the use of different standardization techniques.

Specifically, the non-linearity correction must be performed accurately when using a MAT 253 instrument, whether via equilibrated gases, carbonate standards, or the use of a pressure baseline correction. Our results show that undercorrecting for non-linearity can cause an apparent steepening in the slope of  $\Delta_{47}$ -temperature calibrations, and hypothetically a shallowing if overcorrected. In addition, our analysis suggests that the use of carbonate standards to place  $\Delta_{47}$  data in the absolute reference frame can improve interlaboratory data comparisons, as carbonate standards are not subject to the same types of human influences as equilibrated gases, while also capturing variability in acid digestion that is not recognized by equilibrated gases. Fortunately, the majority of standardization procedures that we investigated produced results that were comparable with one another, showing that prior data utilizing these methods with careful monitoring of reference frame variability is robust.

### Acknowledgements

The work was funded by the Department of Energy through BES grant DE-FG02-13ER16402. Travel to present this work at a conference was partially supported by the “Laboratoire d’Excellence” LabexMER (ANR-10-LABX-19) and co-funded by a grant from the French government under the program “Investissements d’Avenir”.

### References

1. Eiler JM. Paleoclimate reconstruction using carbonate clumped isotope thermometry. *Quat Sci Rev.* 2011;30:3575-3588. doi:10.1016/j.quascirev.2011.09.001
2. Finnigan S, Bergmann K, Eiler JM, et al. The Magnitude and Duration of Late Ordovician-Early Silurian Glaciation. *Science.* 2001;331:903-906. doi:10.1126/science.1200803
3. Henkes GA, Passey BH, Grossman EL, Shenton BJ, Yancey TE, Pérez-Huerta A. Temperature evolution and the oxygen isotope composition of Phanerozoic oceans from carbonate clumped isotope thermometry. *Earth Planet Sci Lett.* 2018;490:40-50. doi:10.1016/j.epsl.2018.02.001
4. Horton TW, Defliese WF, Tripathi AK, Oze C. Evaporation induced  $^{18}\text{O}$  and  $^{13}\text{C}$  enrichment in lake systems: A global perspective on hydrologic balance effects. *Quat Sci Rev.* 2016;131:365-379. doi:10.1016/j.quascirev.2015.06.030
5. Bergman SC, Huntington KW, Crider JG. Tracing paleofluid sources using clumped isotope thermometry of diagenetic cements along the Moab fault, Utah. *Amer J Sci.* 2013;313:490-515. doi:10.2475/05.2013.03

6. Budd DA, Frost III EL, Huntington KW, Allwardt PF. Syndepositional deformation features in high-relief carbonate platforms: Long-lived conduits for diagenetic fluids. *J Sed Res.* 2013;82:12-36. doi:10.2110/jsr.2013.3
7. Dale A, John CM, Mozley PS, Smalley PC, Muggeridge AH. Time-capsule concretions: Unlocking burial diagenetic processes in the Mancos Shale using carbonate clumped isotopes. *Earth Planet Sci Lett.* 2014;394:30-37. doi:10.1016/j.epsl.2014.03.004
8. Carrapa B, Huntington KW, Clementz M, et al. Uplift of the Central Andes of NW Argentina associated with upper crustal shortening, revealed by multiproxy isotopic analyses. *Tectonics.* 2014;33:1039-1054. doi:10.1002/2013TC003461
9. Huntington KW, Lechler AR. Carbonate clumped isotope thermometry in continental tectonics. *Tectonophysics.* 2015;647:1-20. doi:10.1016/j.tecto.2015.02.019
10. Huntington KW, Saylor J, Quade J, Hudson AM. High late Miocene-Pliocene elevation of the Zhada Basin, southwestern Tibetan Plateau, from carbonate clumped isotope thermometry. *GSA Bull.* 2015;127:181-199. doi:10.1130/B31000.1
11. Hill P, Tripathi A, Schauble E. Theoretical constraints on the effects of pH, salinity, and temperature on clumped isotope signatures of dissolved inorganic carbon species and precipitating carbonate minerals. *Geochim Cosmochim Acta.* 2014;125:610-652. doi:10.1016/j.gca.2013.06.018
12. Schauble EA, Ghosh P, Eiler JM. Preferential formation of  $^{13}\text{C}$ - $^{18}\text{O}$  bonds in carbonate minerals, estimated using first-principles lattice dynamics. *Geochim Cosmochim Acta.* 2006;70:2510-2529. doi:10.1016/j.gca.2006.02.011
13. Ghosh P, Adkins J, Affek H, et al.  $^{13}\text{C}$ - $^{18}\text{O}$  bonds in carbonate minerals: A new kind of paleothermometer. *Geochim Cosmochim Acta.* 2006;70:1439-1456. doi:10.1016/j.gca.2005.11.014
14. Huntington KW, Eiler JM, Affek HP, et al. Methods and limitations of 'clumped'  $\text{CO}_2$  isotope ( $\Delta_{47}$ ) analysis by gas-source isotope ratio mass spectrometry. *J Mass Spectrom.* 2009. doi:10.1002/jms.1614
15. Bernasconi SM, Hu B, Wacker U, Fiebig J, Breitenbach SFM, Rutz T. Background effects on Faraday collectors in gas-source mass spectrometry and implications for clumped isotope measurements. *Rapid Commun Mass Spectrom.* 2013;27:603-612. doi:10.1002/rcm.6490
16. Dennis KJ, Affek HP, Passey BH, Schrag DP, Eiler JM. Defining an absolute reference frame for 'clumped' isotope studies of  $\text{CO}_2$ . *Geochim Cosmochim Acta.* 2011;75:7117-7131. doi:10.1016/j.gca.2011.09.025
17. He B, Olack GA, Colman AS. Pressure baseline correction and high-precision  $\text{CO}_2$  clumped-isotope ( $\Delta_{47}$ ) measurements in bellows and micro-volume modes. *Rapid Commun Mass Spectrom.* 2012;26:2837-2853. doi:10.1002/rcm.6436

18. Schmid TW, Bernasconi SM. An automated method for ‘clumped-isotope’ measurements on small carbonate samples. *Rapid Commun Mass Spectrom.* 2010;24:1955-1963. doi:10.1002/rcm.4598
19. Fiebig J, Hofmann S, Löffler N, Lüdecke T, Methner K, Wacker U. Slight pressure imbalances can affect accuracy and precision of dual inlet-based clumped isotope analysis. *Isot Environ Health Stud.* 2016;24:12-28 doi:10.1080/10256016.2015.1010531
20. Wang Z, Schauble EA, Eiler JM. Equilibrium thermodynamics of multiply substituted isotopologues of molecular gases. *Geochim Cosmochim Acta.* 2004;68:4779-4797. doi:10.1016/j.gca.2004.05.039
21. Bernasconi SM, Müller IA, Bergmann KD, et al. Reducing Uncertainties in Carbonate Clumped Isotope Analysis Through Consistent Carbonate-Based Standardization. *Geochem, Geophys, Geosys.* 2018;19:2895-2914. doi:10.1029/2017GC007385
22. Meckler AN, Ziegler M, Millán MI, Breitenbach SFM, Bernasconi SM. Long-term performance of the Kiel carbonate device with a new correction scheme for clumped isotope measurements. *Rapid Commun Mass Spectrom.* 2014;28:1705-1715. doi:10.1002/rcm.6949
23. Müller IA, Rodriguez-Blanco JD, Storck J-C, et al. Calibration of the oxygen and clumped isotope thermometers for (proto-) dolomite based on synthetic and natural carbonates. *Chem Geol.* 2019;525:1-17. doi:10.1016/j.chemgeo.2019.07.014
24. Tang J, Dietzel M, Fernandez A, Tripathi AK, Rosenheim B. Evaluation of kinetic effects on clumped isotope fractionation ( $\Delta_{47}$ ) during inorganic calcite precipitation. *Geochim Cosmochim Acta.* 2014;134:120-136. doi:/10.1016/j.gca.2014.03.005
25. Tripathi A, Hill P, Eagle R, et al. Beyond temperature: Clumped isotope signatures in dissolved inorganic carbon species and the influence of solution chemistry on carbonate mineral composition. *Geochim Cosmochim Acta.* 2015;166:344-371. doi:10.1016/j.gca.2015.06.021
26. van Dijk J, Fernandez A, Storck J-C, et al. Experimental calibration of clumped isotopes in siderite between 8.5 and 62 °C and its application as paleo-thermometer in paleosols. *Geochem Cosmochim Acta.* 2019;254:1-20. doi:10.1016/j.gca.2019.03.018
27. Watkins JM, Hunt JD. A process-based model for non-equilibrium clumped isotope effects in carbonates. *Earth Planet Sci Lett.* 2015;432:152-165. doi:10.1016/j.epsl.2015.09.042
28. Defliese WF, Hren MT, Lohmann KC. Compositional and temperature effects of phosphoric acid fractionation on  $\Delta_{47}$  analysis and implications for discrepant calibrations. *Chem Geol.* 2015;396:51-60. doi:10.1016/j.chemgeo.2014.12.018
29. Kelson JR, Huntington KW, Schauer AW, Saenger C, Lechler AR. Toward a universal carbonate clumped isotope calibration: Diverse synthesis and preparatory methods

- suggest a single temperature relationship. *Geochim Cosmochim Acta*. 2017;197:104-131. doi:10.1016/j.gca.2016.10.010
30. Murray ST, Arienzo MM, Swart PK. Determining the  $\Delta_{47}$  acid fractionation in dolomites. *Geochim Cosmochim Acta*. 2016;174:42-53. doi:10.1016/j.gca.2015.10.029
  31. Swart PK, Murray ST, Staudigel PT, Hodell DA. Oxygen Isotopic Exchange Between CO<sub>2</sub> and Phosphoric Acid: Implications for the Measurement of Clumped Isotopes in Carbonates. *Geochem, Geophys, Geosys*. 2019;20:3730-3750. doi:10.1029/2019GC008209
  32. Olack G, Colman AS. Modeling the Measurement:  $\Delta_{47}$ , Corrections, and Absolute Ratios for Reference Materials. *Geochem, Geophys, Geosys*. 2019;24:3569-3587. doi:10.1029/2018GC008166
  33. Petersen SV, Defliese WF, Saenger C, et al. Effects of Improved <sup>17</sup>O Correction on Interlaboratory Agreement in Clumped Isotope Calibrations, Estimates of Mineral-Specific Offsets, and Temperature Dependence of Acid Digestion Fractionation. *Geochem, Geophys, Geosys*. 2019;20:3495-3519. doi:10.1029/2018GC008127
  34. Daëron M, Blamert D, Peral M, Affek HP. Absolute isotopic abundance ratios and the accuracy of  $\Delta_{47}$  measurements. *Chem Geol*. 2016;442:83-96. doi:10.1016/j.chemgeo.2016.08.014
  35. Schauer AJ, Kelson J, Saenger C, Huntington KW. Choice of <sup>17</sup>O correction affects clumped isotope ( $\Delta_{47}$ ) values of CO<sub>2</sub> measured with mass spectrometry. *Rapid Commun Mass Spectrom*. 2016;30:2607-2616. doi:10.1002/rcm.7743
  36. Brand WA, Assonov SS, Coplen TB. Correction for the <sup>17</sup>O interference in  $\delta^{13}\text{C}$  measurements when analyzing CO<sub>2</sub> with stable isotope mass spectrometry (IUPAC Technical Report). *Pure Appl Chem*. 2010;82:1719-1733. doi:10.1351/PAC-REP-09-01-05
  37. Gonfiantini R, Stichler W, Rozanski K. Standards and intercomparison materials distributed by the International Atomic Energy Agency for stable isotope measurements. IAEA TECDOC 825. International Atomic Energy Agency, Vienna, Austria; 1995.
  38. Müller IA, Fernandez A, Radke J, et al. Carbonate clumped isotope analyses with the long-integration dual-inlet (LIDI) workflow: scratching at the lower sample weight boundaries. *Rapid Commun Mass Spectrom*. 2017;31:1057-1066. doi:10.1002/rcm.7878
  39. Dennis KJ, Schrag DP. Clumped isotope thermometry of carbonatites as an indicator of diagenetic alteration. *Geochim Cosmochim Acta*. 2010;74:4110-4122. doi:10.1016/j.gca.2010.04.005
  40. Bonifacie M, Calmels D, Eiler JM, et al. Calibration of the dolomite clumped isotope thermometer from 25 to 350 °C, and implications for a universal calibration for all (Ca, Mg, Fe)CO<sub>3</sub> carbonates. *Geochim Cosmochim Acta*. 2017;200:255-279. doi:10.1016/j.gca.2016.11.028

41. Eagle RA, Eiler JM, Tripathi AK, et al. The influence of temperature and seawater carbonate saturation state on  $^{13}\text{C}$ - $^{18}\text{O}$  bond ordering in bivalve mollusks. *Biogeosciences*. 2013;10:4591-4606. doi:10.5194/bg-10-4591-2013
42. Fernandez A, Tang J, Rosenheim BE. Siderite ‘clumped’ isotope thermometry: a new paleoclimate proxy for humid continental environments. *Geochim Cosmochim Acta*. 2014;126:411-421. doi:10.1016/j.gca.2013.11.006
43. Henkes GA, Passey BH, Wanamaker Jr AD, Grossman EL, Ambrose Jr WG, Carroll ML. Carbonate clumped isotope compositions of modern marine mollusk and brachiopod shells. *Geochim Cosmochim Acta*. 2013;106:307-325. doi:10.1016/j.gca.2012.12.020
44. Kluge T, John CM. Effects of brine chemistry and polymorphism on clumped isotopes revealed by laboratory precipitation of mono and multiphase calcium carbonates. *Geochim Cosmochim Acta*. 2015;160:155-168. doi:10.1016/j.gca.2015.03.031
45. Fernandez A, Müller IA, Rodriguez-Sanz L, van Dijk J, Looser N, Bernasconi SM. A Reassessment of the Precision of Carbonate Clumped Isotope Measurements: Implications for Calibrations and Paleoclimate Reconstructions. *Geochem, Geophys, Geosys*. 2017;18:4375-4386. doi:10.1002/2017GC007106
46. Kimball J, Eagle R, Dunbar R. Carbonate “clumped” isotope signatures in aragonitic scleractinian and calcitic gorgonian deep-sea corals. *Biogeosciences*. 2016;13:6487-6505. doi:10.5194/bg-13-6487-2016

Sample	# Analyzed Michigan MAT 253	$\delta^{18}\text{O}$	$\delta^{13}\text{C}$	$\Delta_{47}$	# Analyzed UCLA MAT 253	$\delta^{18}\text{O}$	$\delta^{13}\text{C}$	$\Delta_{47}$	# Analyzed UCLA Perspective	$\delta^{18}\text{O}$	$\delta^{13}\text{C}$	$\Delta_{47}$
5 °C Calcite <sub>a,b</sub>	4 <sup>a</sup>	- 6.82±0.20	- 5.58±0.06	0.740±0.003	-				3 <sup>b</sup>	- 5.23±0.08	- 1.92±0.05	0.767±0.010
25 °C Calcite	4	- 10.88±0.14	- 4.83±0.11	0.671±0.013	2	- 10.81±0.02	- 6.37±0.06	0.780±0.061	6	- 10.80±0.05	- 6.35±0.01	0.682±0.010
5 °C Aragonite	4	- 4.27±0.08	1.17±0.08	0.723±0.008	5	- 4.20±0.11	1.38±0.03	0.800±0.030	4	- 4.02±0.07	1.40±0.03	0.753±0.007
25 °C Aragonite	4	- 7.83±0.09	- 0.47±0.06	0.670±0.009	4	- 7.74±0.03	- 0.42±0.01	0.715±0.041	5	- 7.70±0.05	- 0.42±0.05	0.692±0.019
70 °C Aragonite	4	- 16.61±0.14	- 2.69±0.08	0.575±0.011	3	- 16.71±0.01	- 2.60±0.01	0.651±0.037	4	- 16.61±0.12	- 2.58±0.02	0.596±0.014

**Table 1:**  $\delta^{13}\text{C}$ ,  $\delta^{18}\text{O}$  and  $\Delta_{47}$  values of the synthetic carbonates used in this study based on measurements from 3 mass spectrometers. These data are based on the measurements originally reported in Defliese et al<sup>28</sup>, recalculated using the Brand et al<sup>36</sup> isotopic parameters. The number of replicates measured on each instrument is reported. All errors are  $\pm 1$  standard error. a: The 5° C calcite originally measured in Defliese et al<sup>28</sup>. b: A parallel precipitate of the 5° C calcite reported in Defliese et al<sup>28</sup>; this material was grown using the same method and apparatus but was not used for the Defliese et al<sup>28</sup> study.

Method	6-month windows	1-month windows
EG Only	1	6
EG Non-linearity EG + CS ARF	2	7
EG Non-linearity CS ARF	3	8
No Non-linearity CS ARF	4	9
ETH 1-2 Non-linearity CS ARF	5	10

**Table 2:** Numbering scheme for the different standardization techniques used in this study. In future Tables, the institution and instrument used are identified followed by the numeric code for standardization technique, i.e. 'UCLA T 4' for the 6-month, no non-linearity correction, carbonate standards only ARF method using samples/standards analyzed on the UCLA Thermo MAT 253. Method 10 was not applied to the UCLA MAT 253 as we did not analyze enough ETH standards to perform the non-linearity correction in 2 out of 6 months.

Set	Slope	1 SE	Intercept	1 SE
University of Michigan MAT 253 (4 point)	36693	2899	0.2612	0.0322
UCLA MAT 253 1	31280	13390	0.4222	0.1523
UCLA MAT 253 2	30150	12900	0.4159	0.1467
UCLA MAT 253 3	29480	12620	0.4126	0.1435
UCLA MAT 253 4	51600	9152	0.1114	0.1041
UCLA MAT 253 5	24240	12730	0.4715	0.1448
UCLA MAT 253 6	33590	12540	0.3853	0.1426
UCLA MAT 253 7	33440	11620	0.3646	0.1321
UCLA MAT 253 8	32250	9871	0.3610	0.1123
UCLA MAT 253 9	73241	14720	-0.1396	0.1674

**Table 3:** Slopes and intercepts for the nine different regressions calculated on the MAT 253 mass spectrometer at UCLA, plus the equivalent regression at the University of Michigan. The equation solved is  $\Delta_{47} = \text{Slope}/T^2 + \text{Intercept}$ , where T is in Kelvin.

<b>Set</b>	<b>Slope</b>	<b>1 SE</b>	<b>Intercept</b>	<b>1 SE</b>
<b>University of Michigan MAT 253 (5 point)</b>	35254	2476	0.2751	0.0284
<b>UCLA Perspective 1</b>	36720	4049	0.2757	0.0461
<b>UCLA Perspective 2</b>	37310	4114	0.2624	0.0470
<b>UCLA Perspective 3</b>	38730	4270	0.2486	0.0486
<b>UCLA Perspective 4</b>	37300	4238	0.2652	0.0482
<b>UCLA Perspective 5</b>	39250	4284	0.2426	0.0490
<b>UCLA Perspective 6</b>	36360	3757	0.2865	0.0428
<b>UCLA Perspective 7</b>	36690	3829	0.2794	0.0436
<b>UCLA Perspective 8</b>	37980	3990	0.2680	0.0454
<b>UCLA Perspective 9</b>	37210	3992	0.2768	0.0454

<b>UCLA Perspective 10</b>		38750	3742	0.2594	0.0426					
UM 253		UCLA P1	UCLA P2	UCLA P3	UCLA P4	UCLA P5	UCLA P6	UCLA P7	UCLA P8	UCLA P9

**Table 4:** Slopes and intercepts for the ten different regressions calculated on the Perspective IS mass spectrometer at UCLA, plus the equivalent regression at the University of Michigan for the MAT 253. The equation solved is  $\Delta_{47} = \text{Slope}/T^2 + \text{Intercept}$ , where T is in Kelvin.

UM 253	-	0.0798*	0.3564	0.2146	0.2275	0.2084	0.0067*	0.0238*	0.0090*	0.0094*
UCLA P1	0.0798*	-	0.7402	0.8310	0.8957	0.7948	0.7248	0.9268	0.7336	0.7407
UCLA P2	0.3564	0.7402	-	0.9421	0.9566	0.9245	0.2871	0.5015	0.3186	0.3141
UCLA P3	0.2146	0.8310	0.9421	-	0.9711	0.9961	0.4007	0.6305	0.4627	0.4439
UCLA P4	0.2275	0.8957	0.9566	0.9711	-	0.9469	0.4538	0.6937	0.4876	0.4829
UCLA P5	0.2084	0.7948	0.9245	0.9961	0.9469	-	0.3750	0.5972	0.4446	0.4219
UCLA P6	0.0067*	0.7248	0.2871	0.4007	0.4538	0.3750	-	0.9136	0.9569	0.9879
UCLA P7	0.0238*	0.9268	0.5015	0.6305	0.6937	0.5972	0.9136	-	0.9033	0.9187
UCLA P8	0.0090*	0.7336	0.3186	0.4627	0.4876	0.4446	0.9569	0.9033	-	0.9906
UCLA P9	0.0094*	0.7407	0.3141	0.4439	0.4829	0.4219	0.9879	0.9187	0.9906	-
UCLA P10	0.0055*	0.6918	0.2910	0.4469	0.4570	0.4336	0.9035	0.8588	0.9903	0.9613

5A

Table 5B

	UCLA P1	UCLA P6	UCLA P7	UCLA P8	UCLA P9	UCLA P10
UM 253	0.7601	0.8082	0.7558	0.5668	0.6807	0.4433

**Table 5:** ANCOVA analysis of the regressions produced via the different standardization techniques on the UCLA Perspective IS mass spectrometer (denoted P in this table), plus the equivalent regression at the University of Michigan 253. A: P-values for each regression pair are listed in the matrix, with – for boxes that are self-comparison. Color coding indicates significance, green boxes produced statistically identical values ( $p\text{-value} > 0.1$ , reject different slope/intercept), yellow boxes are borderline different ( $0.1 > p\text{-value} > 0.05$ , statistically different at 90% confidence), and red boxes indicate significantly different values ( $p\text{-value} < 0.05$ , statistically different at 95% or greater confidence). \*Indicates that the regression pair produces a statistically identical slope, but not a statistically identical intercept, with p-values for common slope listed in part B. B: For regression pairs that are statistically different ( $p\text{-value} < 0.05$ ), p-values of a test to determine common slope (i.e., temperature dependence), with same color coding as A. All regression pairs show a common slope, indicating statistical differences only in the intercepts.

6A:

	UM 253	UCLA T1	UCLA T2	UCLA T3	UCLA T4	UCLA T5	UCLA T6	UCLA T7	UCLA T8	UCLA T9
UM 253	-	0.0023*	0.0016*	0.0046*	0.1659	0.0037*	0.0005*	0.0043*	0.0125*	0.0503
UCLA T1	0.0023*	-	0.8181	0.6083	0.0097*	0.5700	0.9270	0.5142	0.1966	0.0074
UCLA T2	0.0016*	0.8181	-	0.9335	0.0333*	0.8852	0.9447	0.8636	0.5013	0.0167
UCLA T3	0.0046*	0.6083	0.9335	-	0.0616	0.9581	0.7868	0.9660	0.7266	0.0246
UCLA T4	0.1659	0.0097*	0.0333*	0.0616	-	0.0423*	0.0194*	0.0799	0.1454	0.4552
UCLA T5	0.0037*	0.5700	0.8852	0.9581	0.0423*	-	0.7057	0.8616	0.6585	0.0160
UCLA T6	0.0005*	0.9270	0.9447	0.7868	0.0194*	0.7057	-	0.7232	0.3305	0.0130*
UCLA T7	0.0043*	0.5142	0.8636	0.9660	0.0799	0.8616	0.7232	-	0.7914	0.0329
UCLA T8	0.0125*	0.1966	0.5013	0.7266	0.1454	0.6585	0.3305	0.7914	-	0.0430
UCLA T9	0.0503	0.0074	0.0167	0.0246	0.4552	0.0160	0.0130*	0.0329	0.0430	-

6B:

UM vs T1	UM vs T2	UM vs T3	UM vs T5	UM vs T6	UM vs T7	UM vs T8	T1 vs T4	T1 vs T9	T2 vs T4	T2 vs T9	T3 vs T9	T4 vs T5	T4 vs T6	T5 vs T9	T6 vs T9	T7 vs T9	T8 vs T9
0.6794	0.6052	0.5609	0.3222	0.8008	0.7760	0.6532	0.2224	0.0456	0.1877	0.0376	0.0334	0.0937	0.2573	0.0189	0.0514	0.0443	0.0296

**Table 6:** ANCOVA analysis of the regressions produced via the different standardization techniques on the UCLA MAT 253 mass spectrometer (denoted T in this table), plus the equivalent regression at the University of Michigan 253. A: P-values for each regression pair are listed in the matrix, with – for boxes that are self-comparison. Color coding indicates significance, green boxes produced statistically

identical values (p-values > 0.1, reject different slope/intercept), yellow boxes are borderline different (0.1 > p-value > 0.05, statistically different at 90% confidence), and red boxes indicate significantly different values (p-value < 0.05, statistically different at 95% or greater confidence). \*Indicates that the regression pair produces a statistically identical slope, but not a statistically identical intercept, with p-values for common slope listed in part B. B: For regression pairs that are statistically different (p-value < 0.05), p-values of a test to determine common slope (i.e., temperature dependence), with same color coding as A.

7A:

	UM 253	UCLA P6	UCLA P7	UCLA P8	UCLA P9	UCLA P10	UCLA T6	UCLA T7	UCLA T8	UCLA T9
UM 253	-	0.0628	0.1534	0.0973	0.0892	0.0777	0.0005*	0.0043*	0.0125*	0.0503
UCLA P6	0.0628	-	0.9279	0.9637	0.9911	0.9054	0.0033*	0.0378*	0.1491	0.0280
UCLA P7	0.1534	0.9279	-	0.9343	0.9436	0.8832	0.0022*	0.0253*	0.1014	0.0322
UCLA P8	0.0973	0.9637	0.9343	-	0.9910	0.9866	0.0035*	0.0381*	0.1450	0.0383
UCLA P9	0.0892	0.9911	0.9436	0.9910	-	0.9559	0.0036*	0.0391*	0.1505	0.0335
UCLA P10	0.0777	0.9054	0.8832	0.9866	0.9559	-	0.0032*	0.0358*	0.1351	0.0426
UCLA T6	0.0005*	0.0033*	0.0022*	0.0035*	0.0036*	0.0032*	-	0.7232	0.3305	0.0130*
UCLA T7	0.0043*	0.0378*	0.0253*	0.0381*	0.0391*	0.0358*	0.7232	-	0.7914	0.0329
UCLA T8	0.0125*	0.1491	0.1014	0.1450	0.1505	0.1351	0.3305	0.7914	-	0.0430
UCLA T9	0.0503	0.0280	0.0322	0.0383	0.0335	0.0426	0.0130*	0.0329	0.0430	-

7B:

UM vs T6	UM vs T7	UM vs T8	P6 vs T6	P6 vs T7	P6 vs T9	P7 vs T6	P7 vs T7	P7 vs T9	P8 vs T6	P8 vs T7	P8 vs T9	P9 vs T6	P9 vs T7	P9 vs T9	P10 vs T6	P10 vs T7	P10 vs T9	T6 vs T9	T7 vs T9	T8 vs T9
0.8008	0.7760	0.6532	0.9429	0.9283	0.0093	0.9181	0.9020	0.0010	0.8372	0.8164	0.0013	0.8915	0.8739	0.0011	0.7705	0.7455	0.0148	0.0051	0.0044	0.0029

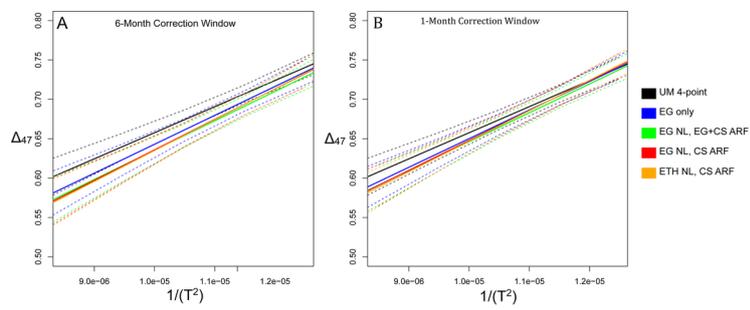
**Table 7:** Comparison of the ANCOVA analysis of the Perspective IS (abbreviated as P) and MAT 253 (abbreviated as T) instruments at UCLA. Only the standardization techniques utilizing 1-month windows are considered, as this is more representative of real-world practices. For this analysis, only the four samples measured on both instruments were considered. A: P-values for each regression pair are listed in the matrix, with – for boxes that are self-comparison. Color coding indicates significance, green boxes produced statistically identical values ( $p\text{-value} > 0.1$ , reject different slope/intercept), yellow boxes are borderline different ( $0.1 > p\text{-value} > 0.05$ , statistically different at 90% confidence), and red boxes indicate significantly different values ( $p\text{-value} < 0.05$ , statistically different at 95% or greater confidence). \*Indicates that the regression pair produces a statistically identical slope, but not a statistically identical intercept, with p-values for common slope listed in part B. B: For regression pairs that are statistically different ( $p\text{-value} < 0.05$ ), p-values of a test to determine common slope (i.e., temperature dependence), with the same color coding as A.

**Figure 1:** Plot of the various  $\Delta_{47}$ -Temperature regressions from the UCLA Perspective IS instrument as well as the recalculated 5-point line from the University of Michigan. Temperature is in degrees Kelvin. Dashed lines are 95% confidence intervals. A: Regressions produced using a 6-month standardization window (methods 1-5). B: Regressions produced using a shorter, 1-month standardization window (methods 6-9).

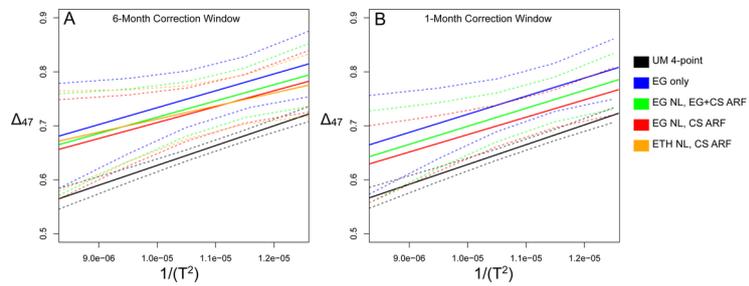
**Figure 2:** Plot of the various  $\Delta_{47}$ -Temperature regressions from the UCLA MAT 253 instrument as well as the recalculated 4-point line from the University of Michigan. Temperature is in degrees Kelvin. Dashed lines are 95% confidence intervals. A: Regressions produced using a 6-month standardization window (methods 1-5). B: Regressions produced using a shorter, 1-month standardization window (methods 6-9).

**Figure 3:** Plots of the  $\Delta_{47}$ -Temperature calibrations using short correction windows, comparing the Brand and Gonfiantini  $^{17}\text{O}$  corrections, and the use of equilibrated gases alone (method 6 in Table 2) versus equilibrated gases and carbonate standards (method 7 in Table 2). Note that all University of Michigan data uses method 6 (equilibrated gases only) in all figures. Dashed lines are 95% confidence intervals. Red line: Data from the University of Michigan (all 8 samples). Grey line: Data from the UCLA Perspective IS (5 samples). Blue line: Data from the UCLA MAT 253 (4 samples). Black line: All combined data. A: Equilibrated gases only ARF (method 6), calculated using the Brand  $^{17}\text{O}$  correction.

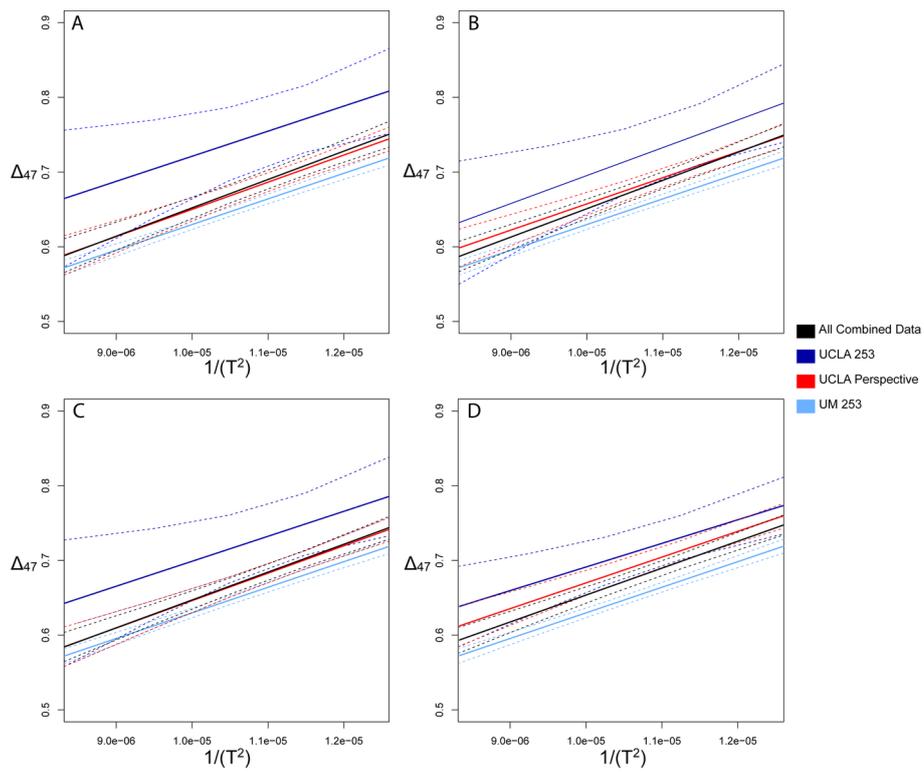
B: Equilibrated gases only ARF (method 6), calculated using the Gonfiantini  $^{17}\text{O}$  correction. C: Equilibrated gases + carbonate standards ARF (method 7), calculated using the Brand  $^{17}\text{O}$  correction. D: Equilibrated gases + carbonate standards ARF (method 7), calculated using the Gonfiantini  $^{17}\text{O}$  correction.



RCM\_8666\_Figure 1.tif



RCM\_8666\_Figure 2.tif



RCM\_8666\_Figure 3.tif