

1

# 1 A community effort to improve inter-laboratory 2 standardization of the carbonate clumped isotope 3 thermometer using carbonate anchors

4 S.M. Bernasconi<sup>1</sup>, M. Daëron<sup>2</sup>, K. Bergmann<sup>3</sup>, M. Bonifacie<sup>4</sup>, A.N. Meckler<sup>5</sup>, others in  
5 alphabetical order<sup>3</sup>

6 <sup>1</sup>Geological Institute, ETH Zürich, Sonneggstrasse 5, 8092 Zürich.

7 <sup>2</sup>Laboratoire des Sciences du Climat et de l'Environnement, LSCE/IPSL, CEA-CNRS-  
8 UVSQ, Université Paris-Saclay, Orme des Merisiers, F-91191 Gif-sur- Yvette Cedex, France.

9 <sup>3</sup>Department of Earth, Atmospheric and Planetary Sciences, Massachusetts Institute of  
10 Technology, Cambridge, MA 02139, USA.

11 <sup>4</sup>Université de Paris, Institut de Physique du Globe de Paris, CNRS, F-75005 Paris, France

12 <sup>5</sup>Department of Earth Science, University of Bergen, Bergen, Norway

13

14 Corresponding author: Stefano M. Bernasconi (stefano.bernasconi@erdw.ethz.ch) Mathieu  
15 Daëron (daeron@lsce.ipsl.fr)

16 †Additional author notes should be indicated with symbols (for example, for current  
17 addresses).

## 18 Key Points

- 19 • The exclusive use of carbonate reference materials is a robust method for the  
20 standardization of  $\Delta_{47}$  measurements yielding very good analytical reproducibility  
21 between laboratories, with residual inter-laboratory discrepancies quantitatively  
22 consistent with in-lab analytical uncertainties.
- 23 •  $\Delta_{47}$  measurements using different acid temperatures, different designs of sample  
24 preparation lines and different mass spectrometer models are statistically  
25 indistinguishable.
- 26 • We propose updated guidelines to standardize and report  $\Delta_{47}$  measurements in future  
27 studies using exclusively carbonates.
- 28 • We propose new consensus  $\Delta_{47}$  values for a set of 6 carbonate reference materials  
29 including the four ETH standards and two IAEA standards.

## 30 Abstract

31 Increased adoption and improved methodology in carbonate clumped isotope thermometry  
32 has greatly enhanced our ability to interrogate a suite of Earth-system processes. However,  
33 inter-laboratory discrepancies in quantifying carbonate clumped isotope ( $\Delta_{47}$ ) measurements  
34 persist, and their specific sources remains unclear. To address inter-laboratory differences as  
35 a clumped isotope community, we first provide community consensus values for four  
36 carbonate standards relative to heated and equilibrated gases with 1,447 individual  
37 measurements from nine laboratories. Then we analyzed the four carbonate standards and an  
38 additional three standards, spanning a broad range of  $\delta^{47}$  and  $\Delta_{47}$  compositions, a total of  
39 5,202 times on 26 mass spectrometers representing 23 unique laboratories. We use three  
40 standards to calculated values for the other four standards and find that the use of carbonate  
41 reference materials is a robust method for standardization that yields inter-lab discrepancies  
42 consistent with intra-lab analytical uncertainty. The use of carbonate reference materials,  
43 along with measurement and data processing practices described herein, provides the  
44 carbonate clumped isotope community with a robust way to achieve inter-laboratory  
45 agreement as we continue to use and improve this powerful geochemical tool.

## 51 Plain Language Summary

### 52 1 Introduction

55 Carbonate clumped isotope ( $\Delta_{47}$ ) thermometry is the most developed branch of the rapidly  
56 evolving field of clumped isotope geochemistry. Given the broad range of applications in  
57 Earth Sciences (e.g. Affek & Eiler, 2006; Eagle et al., 2010; Ferry et al., 2011; Ghosh,  
58 Adkins, et al., 2006; Ghosh, Garzione, et al., 2006; Grauel et al., 2013; Guo & Eiler, 2007;  
59 Huntington et al., 2011; Passey & Henkes, 2012) Carbonate clumped isotope thermometry is  
60 the most developed branch of the rapidly evolving field of clumped isotope geochemistry.  
61 Given the broad range of applications in Earth Sciences (Affek & Eiler, 2006; Eagle et al.,  
62 2010; Ferry et al., 2011; Ghosh, Adkins, et al., 2006; Ghosh, Garzione, et al., 2006; Grauel  
63 et al., 2013; Guo & Eiler, 2007; Huntington et al., 2011; Passey & Henkes, 2012) and the  
64 improvement of analytical methods including automation (Bernasconi et al., 2013, 2018;  
65 Defliese & Lohmann, 2015; Dennis et al., 2011; Fiebig et al., 2019; Ghosh, Adkins, et al.,  
66 2006; He et al., 2012; Hu et al., 2014; Huntington et al., 2009; Meckler et al., 2014; Müller,  
67 Fernandez, et al., 2017; Passey et al., 2010; Petersen et al., 2019; Petersen et al., 2016;

Schmid & Bernasconi, 2010), the last 5-10 years have seen an increasing number of laboratories implementing this technique. The great potential of this thermometric method can only be fully exploited if the precision and accuracy of the obtained data is sufficient to solve differences of a few degrees in formation temperatures. In addition, like for any other geochemical method, widely available reference materials that match the sample matrices are necessary so that data can be robustly compared across laboratories (Meier-Augenstein & Schimmelmann, 2019). Currently the situation in the field of clumped isotope geochemistry of carbonates, is far from satisfactory. Published values for the ETH reference materials, the only carbonates that have been recently measured in many different laboratories, show a range of values for the same sample of up to 0.053 ‰ (see Bernasconi et al. 2018, Thaler et al. 2020 for recent comparisons). This clearly calls for better standardization procedures to improve laboratory comparability.

The data normalization scheme used in clumped isotope geochemistry of carbonates in most laboratories is based on the comparison of the composition of the CO<sub>2</sub> liberated from carbonates by reaction with phosphoric acid with that of a set of CO<sub>2</sub> reference or standard gases with different bulk and clumped isotope compositions. These reference gases are prepared either by heating CO<sub>2</sub> at 1000°C (heated gases; HG) or CO<sub>2</sub> equilibration with water at low temperatures (equilibrated gases at *e.g.* 25°C, 50°C; EG). By comparison of the measured compositions with the theoretical predictions of the equilibrium thermodynamic abundance of multiply substituted isotopologues in heated and equilibrated gases (Wang et al. 2004 and updates in Petersen et al. 2019), the measurements are standardized to the scale that was named the “absolute reference frame” (ARF) by (Dennis et al., 2011). In more recent publications the ARF is often referred to as the “Carbon Dioxide Equilibration Scale” (CDES) which is the terminology we adopt here. This approach was designed to allow different laboratories to link their measurements to an internationally recognized scale firmly anchored to theory using relatively easy and established laboratory protocols to produce CO<sub>2</sub> reference gases of known composition. In spite of the introduction of the CDES, however, the method still has numerous open questions (see (Bernasconi et al., 2018; Petersen et al., 2019 for two recent discussions).

Two of the main problems still limiting the reliability of this method to yield accurate temperature reconstructions are the lack of internationally recognized reference materials for

a precise inter-laboratory calibration. Moreover, published  $\Delta_{47}$ -temperature calibrations produced in different laboratories have differed in both temperature dependence (slope), and absolute values (intercept). Possible reasons for the differences in slope and intercepts of the  $\Delta_{47}$  temperature dependence have been widely discussed in the literature (e.g. Bonifacie et al., 2017; Daëron et al., 2016b; Fernandez et al., 2017; Katz et al., 2017; Kelson et al., 2017; Petersen et al., 2019; Schauer et al., 2016). Discrepancies have been attributed to analytical artefacts such as CO<sub>2</sub>-acid re-equilibration at different acid digestion temperatures (see Swart et al., (2019) for a recent discussion). Other main factors proposed to influence the calculated slope of the calibrations are the limitations of the datasets used in the individual studies, in particular in terms of number of samples and replicates and of the temperature range covered by the available samples ((Bonifacie et al., 2017; Fernandez et al., 2017). However, the discrepancies in the intercept of the calibrations, for example between (Kelson et al., 2017) and (Peral et al., 2018) and in laboratory comparability still remain a problem that can only be improved by using a more robust standardization method.

Petersen et al., (2019) in a recent effort to solve differences in calibrations, compiled raw data of a number of published temperature calibrations and recalculated them all in a consistent way using the revised “IUPAC” correction parameters to correct for the <sup>17</sup>O abundance (Daëron et al., 2016a; Schauer et al., 2016). The goal was to test whether data processing differences and/or the use of consistent but incorrect <sup>17</sup>O- correction parameters in the calculations were the root causes of inconsistencies. The result of this study was that differences among calibrations were reduced but not eliminated by the recalculation, implying that other factors must be responsible for the remaining discrepancies. These differences have pushed many laboratories to use laboratory-specific calibrations performed with the same analytical approach, as they take at least partially into consideration possible procedural differences (Petersen et al., 2019). However, if a laboratory changes analytical procedures or has not realized an in-house calibration, this approach is problematic. Good inter-laboratory reproducibility is a natural requirement for sufficiently mature analytical methods and striving to ensure that  $\Delta_{47}$  measurements meet this goal is now key to its broad acceptance and quantitative usefulness.

While the definition of the CDES was a major milestone, a known problem with this approach is that while the CO<sub>2</sub> reference gases equilibrated at known temperature (HT or EG)

can be confidently used for correction of mass spectrometric fractionations/nonlinearities and for effects of the purification procedures, they cannot account for the effects of the phosphoric acid reaction on the composition of the produced CO<sub>2</sub>. Among the factors responsible for discrepant calibrations and laboratory comparability two important ones cannot be tested with a gas-based standardization. These are (1) the absolute value and temperature dependence of the phosphoric acid fractionation factor (see Petersen et al., 2019 for a recent compilation) and (2) possible CO<sub>2</sub> equilibration effects during acid digestion of the sample. Swart et al. (2019) presented evidence that equilibration of CO<sub>2</sub> with water or hot metal surfaces during phosphoric acid reaction and transfer of the CO<sub>2</sub> to the mass spectrometer could be a factor leading to the alteration of the apparent temperature dependence of clumped isotopes in carbonates and on the absolute value of calculated  $\Delta_{47}$ . As many laboratories use custom built extraction lines with very different designs and volumes of tubing and of acid vessels, these factors are impossible to precisely quantify for each laboratory.

We propose that these issues can be circumvented if carbonates, which undergo the same acid digestion as the samples, are used for normalization instead of gases, consistent with the principle of identical treatment of sample and standards (Carter & Fry, 2013; Werner & Brand, 2001a). In addition, normalizing results to some accepted values for the solids, as is commonly done with conventional oxygen isotopes in carbonates, makes it no longer necessary to precisely quantify acid fractionation factors at different temperatures (Bernasconi et al., 2018).

A carbonate standardization approach was introduced by (Schmid & Bernasconi, 2010) and improved by (Meckler et al., 2014), with the following benefits (1) the use of carbonates can more easily be fully automatized, eliminating time-consuming and possibly error-prone manual preparation of CO<sub>2</sub> reference gases (equilibrated at known temperature) by individual users on separate extraction lines and (2) in some automated system designed for small sample sizes, the gases had to be measured through a different capillary than the carbonates with potential biases that would go unrecognized and (3) in those systems the CO<sub>2</sub> reference gases are measured at constant beam intensity whereas the samples are measured with decreasing beams. These features argue in favor of carbonate standardization *a priori*, but it remains critical to assess *a posteriori* whether the results of this approach are as robust and accurate as expected, and whether they significantly improve the inter-laboratory

reproducibility of  $\Delta_{47}$  measurements. Discussions at the Sixth International Clumped Isotope Workshop (Paris, 2017) led to the present inter-laboratory comparison exercise (InterCarb) to evaluate the benefits and drawbacks of a carbonate-based standardization approach as an alternative to the use of gas standards.

The primary goal of this study was to test whether the exclusive use of carbonate reference materials can solve inter-laboratory discrepancies and provide an alternative to the measurement of heated and equilibrated gases for the entire community. This is particularly important because of the increasing number of laboratories that using the commercial small-sample automated devices which cannot easily be standardized using the HG-EG approach. The InterCarb exercise also provides a natural opportunity to define the best community-derived absolute  $\Delta_{47}$  values for the ETH standards of (Meckler et al., 2014): although these standards are already used in many laboratories, their current nominal  $\Delta_{47}$  values are based on measurements from the ETH laboratory only. The InterCarb exercise can similarly establish community accepted values for other common carbonate reference materials, some of which have been in use for several years, in order to provide the community with a self-consistent set of carbonate reference materials with a broad range of bulk and clumped isotope compositions.

## 1.2. Nomenclature and data processing

Clumped isotope compositions are reported as an excess abundance of the  $\text{CO}_2$  isotopologue of cardinal mass 47 (dominantly the isotopologues  $^{13}\text{C}^{18}\text{O}^{16}\text{O}$ ) compared to a stochastic distribution according to the formula:

$$\Delta_{47}(\text{‰}) = \frac{R^{47}}{R^{47*}} - 1$$

where  $R^{47}$  is the abundance of the minor isotopologues 47 relative to the most abundant isotopologue with mass 44. The expected stochastic ratio  $R^{47*}$  is calculated using the measured abundance of  $^{13}\text{C}$  and  $^{18}\text{O}$  in the sample (Affek & Eiler, 2006). The measured abundance of isotopologues with  $m/z$  47 in the sample with respect to the working gas (WG) in the mass spectrometer is reported in the traditional delta notation as:

$$\delta^{47}(\text{‰}) = \left[ \left( \frac{R^{47}_{\text{sample}}}{R^{47}_{\text{working gas}}} \right) - 1 \right] \times 1000 .$$

201

202 The  $\delta^{47}$  scale is a measure of the difference between the sample of interest and the WG of the  
 203 specific instrument, therefore, it cannot be compared across laboratories. The same notation  
 204 is used for 45, 46, 48 and 49 beams.

205

206 The CO<sub>2</sub> gas-based standardization scheme for clumped isotope thermometry in carbonates  
 207 relies on a set of CO<sub>2</sub> reference gases with different bulk composition ( $\delta^{47}$ ), preferably chosen  
 208 by the user to encompass the  $\delta^{47}$  bulk composition of unknown samples, that have been  
 209 heated at 1000°C to reach a near-stochastic distribution of all isotopologues and one set  
 210 equilibrated with water at low temperatures to reach equilibrium enrichments in the mass-47  
 211 isotopologues (Dennis et al., 2011). The heated gases having a stochastic distribution of all  
 212 isotopologues define the zero point of the scale ( $\Delta_{47} = 0.0266 \text{ ‰}$ ), the equilibrated gases a  
 213 high point, e.g., at 25 °C  $\Delta_{47} = 0.9196 \text{ ‰}$ ), with theoretical values linking measurements to  
 214 theory calculated by Wang et al. (2004), and revised by Petersen et al., (2019). A wide range  
 215 in  $\delta^{47}$  compositions of gases used for normalization is generally chosen so that when plotted  
 216 in a  $\delta^{47}$  vs-  $\Delta_{47}$  plot the broad range of compositions can be used to correct for an apparent  
 217 dependence of  $\Delta_{47}$  on  $\delta^{47}$  which is caused by negative or sometimes positive signal  
 218 backgrounds on the m/z 47 collector (Bernasconi et al., 2013; He et al., 2012). The large  
 219 range in  $\Delta_{47}$  (i.e. 25°C, 1000°C), on the other hand, is necessary to correct for scale  
 220 compression caused by processes of scrambling and molecule recombination in the source of  
 221 the mass spectrometer or elsewhere in the sample preparation/dual-inlet pipeline (Dennis et  
 222 al., 2011). With properly chosen CO<sub>2</sub> reference gases with widely varying  $\delta^{47}$  composition it  
 223 is possible to cover the entire range of natural carbonate compositions, avoiding  
 224 extrapolations in the ( $\delta^{47}$ ,  $\Delta_{47}$ ) compositional space (Fig. 1). Note that measurement errors  
 225 (typically not better than 0.010‰) being relatively large compared to the natural  
 226 compositional range (absolute value is less than 0.7‰), the large (>0.9 ‰) difference in  $\Delta_{47}$   
 227 of the CO<sub>2</sub> reference gases minimizes analytical errors introduced by uncertainties resulting  
 228 from the measurement of HG and EG.

229

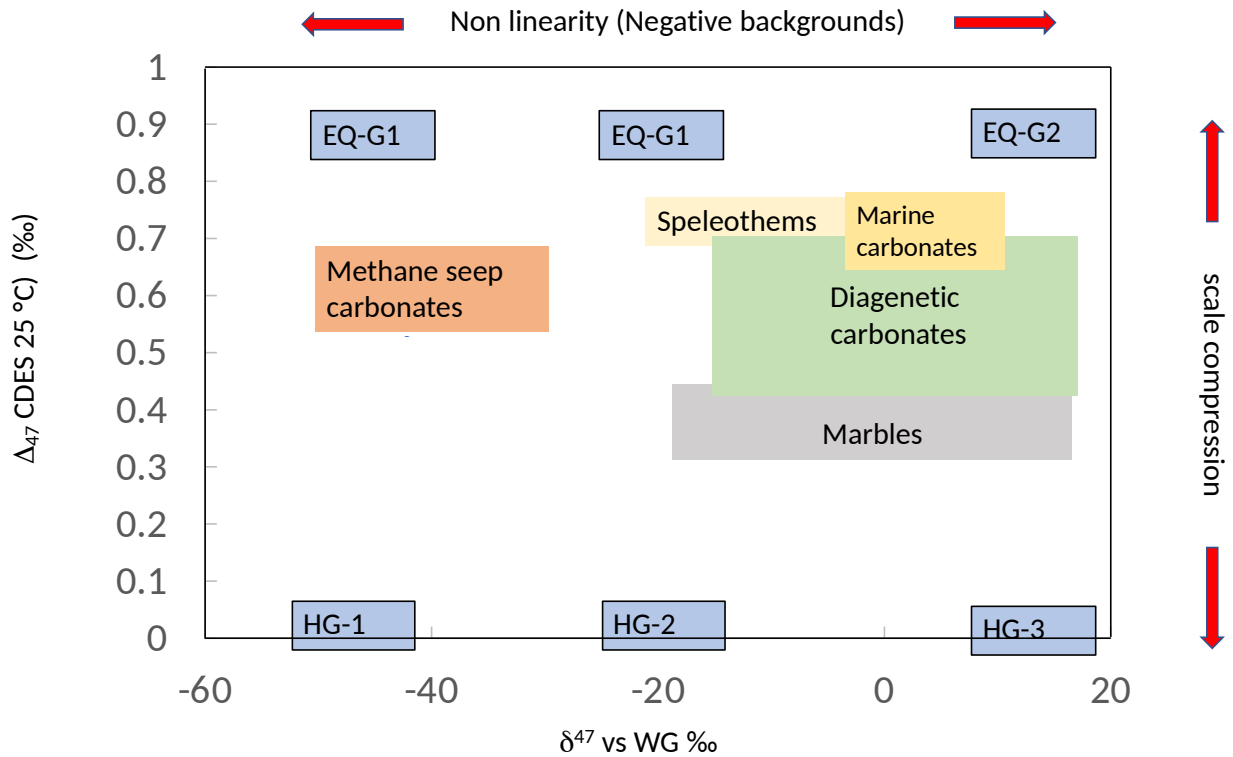


Fig 1  $\delta^{47}$  vs.  $\Delta_{47}$  composition of heated and equilibrated gases in a range commonly used in many laboratories showing that the observed range in measured clumped isotope compositions in natural carbonates can be completely bracketed by heated and equilibrated  $\text{CO}_2$  reference gases from which  $\delta^{47}$  composition have been chosen by the user.

Meckler et al., (2014) attempted to achieve a similar framework as the  $\text{CO}_2$  gas-based standardization but with carbonate standards. They described four carbonates that were developed at ETH Zürich to serve as replacements for HG-EGs and demonstrated that very good long- and short-term reproducibility can be achieved using only carbonates for data correction. Bernasconi et. al (2018) discussed in detail these standards and postulated, based on a limited inter-laboratory dataset, that carbonate standardization should improve inter-laboratory data comparability in most cases. This claim seems arguably strengthened by the results of Meinicke et al., (2020) Peral et al. (2018), Piasecki et al. (2019) and (Jautzy et al., 2021). The first three studies produced independent foraminifera-based and the fourth a synthetic carbonate-based  $\Delta_{47}$ -temperature calibration anchored to the same set of carbonate standards, yielding statistically indistinguishable slopes and intercepts despite the use of independent sample sets and different analytical systems. The same conclusion was reached



by (Spooner et al., 2016)(Spooner et al., 2016) found that carbonate standardization removed data biases between the Caltech and the WHOI laboratories, which were present when not removed by normalization to the CDES by heated and equilibrated gases was used..

A possible limitation of carbonate standardization is that available carbonates have a smaller range in  $\delta^{47}$  and, perhaps more importantly, a smaller range in  $\Delta_{47}$  compositions than what is achievable with heated and equilibrated gases. In some cases, standardization procedures may require extrapolation to compositions that are not within the  $\delta^{47}$ - $\Delta_{47}$  space created by carbonate standards (Fig. 2). In addition, the range of  $\Delta_{47}$  compositions for carbonates is only on the order of 0.45 ‰ between 0 and 1000°C. The smaller range in  $\Delta_{47}$  compared to HG-EG requires higher precision and thus high standard replication and a standard:sample ratio  $\geq 1$  to keep normalization errors small (see Daëron (this volume) for details).

## 1.2. InterCarb goals and design:

InterCarb was designed after the Paris' Clumped Isotope Workshop in 2017 to carefully evaluate the potential of carbonates to serve as a standardization scheme that improves inter-laboratory agreement for 'unknown' carbonates both inside and outside of the  $\delta^{47}$ - $\Delta_{47}$  space defined by the anchor samples (Fig. 2). The main questions posed are:

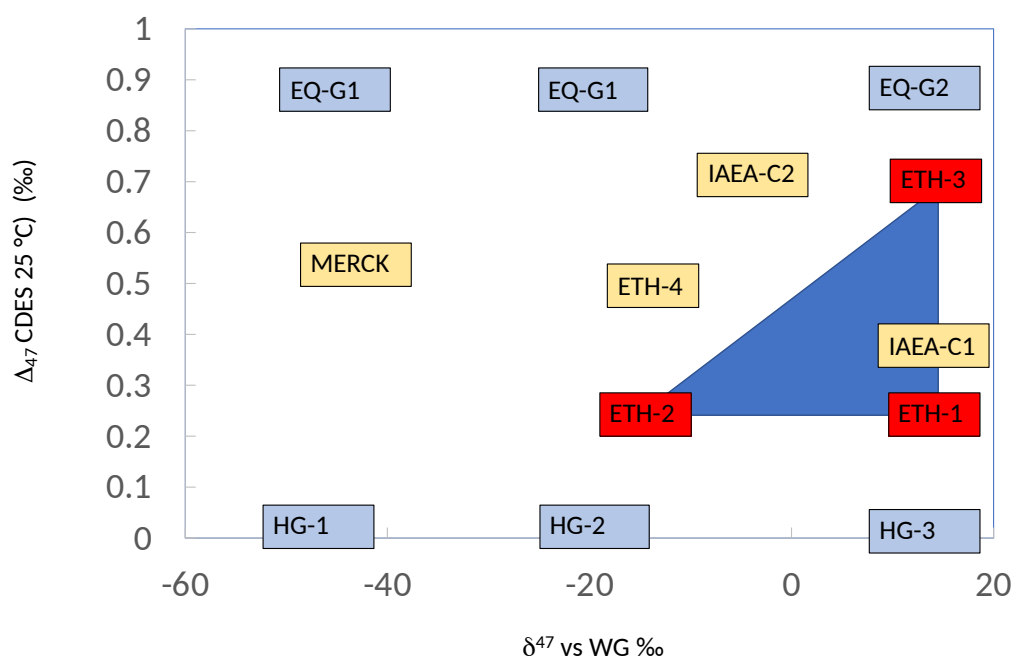
1. Is it possible to produce consistent carbonate clumped isotope measurements across laboratories using exclusively carbonate reference materials? In other words, does the observed inter-laboratory scatter in  $\Delta_{47}$  values match that expected from intra-laboratory analytical precision limits?
2. How well does the carbonate standardization approach perform when extrapolating beyond the  $\delta^{47}$ - $\Delta_{47}$  compositional space sampled by a set of carbonate reference materials?
3. Do carbonate reference materials fully correct effects arising from different reaction temperatures, sample preparation protocols, and analytical equipment?
4. Can we define a self-consistent set of widely available reference materials with community-agreed on compositions accurately anchored to the CDES scale?

5. Does inter-laboratory reproducibility of carbonate clumped isotope analyses corrected using carbonate reference materials improve on the relatively large reported inter-laboratory differences using HG-EG standardization?

### Approach:

It was decided to distribute and analyze a common set of 7 carbonate standards with a large range of  $\delta^{47}$  and  $\Delta_{47}$  compositions (Figure 2), treating three of them as “anchors” to the CDES and processing the remaining four as “unknowns”. Due to the relatively widespread use in different laboratories the three reference materials ETH-1, ETH-2 and ETH-3 (Meckler et al. 2014; Bernasconi et al. 2018) were chosen as anchors. They are still available today in relatively large quantities (>600 g), have been in use at ETH since 2013 and in many other laboratories for several years. Importantly, they have been thoroughly tested for homogeneity based on thousands of measurements in 80 to 150  $\mu$ g aliquot sizes in different laboratories.

The “unknown” InterCarb reference materials were chosen to cover a wide natural range in  $\delta^{47}$  and  $\Delta_{47}$  compositions. These samples had to be available in large quantities, inexpensive, and if possible distributed by an organization with a long-term perspective in order to ensure future data quality and availability for the increasing number of laboratories.



*Fig 2. Approximate compositions of the anchor samples (red) and the unknowns (yellow) in  $\delta^{47}$  -  $\Delta_{47}$  space for a mass spectrometer with a working gas (WG) with a bulk isotope composition similar to that used in many laboratories (e.g. CO<sub>2</sub> provided by Oztech Trading corporation with  $d18O$  close to X‰ and  $d13C$  close to Y‰). Note the smaller achievable range compared to heated and equilibrated gases and the large extrapolation necessary for the determination of the composition for MERCK. Heated and equilibrated CO<sub>2</sub> reference gases have a larger  $\Delta_{47}$  range allowing for more robust stretching calculations with identical numbers of standard:sample analyses.*

## **2 Materials and Methods**

### **2.1. Sample description**

The anchor samples ETH-1 (Carrara marble heated at 600°C) ETH-2 (synthetic carbonate heated at 600°C) and ETH-3 (Upper cretaceous chalk) are described in detail in Bernasconi et al. (2018).

IAEA-C1 (marble from Carrara, Italy) is distributed by the International Atomic Energy Agency (IAEA) as a mechanically crushed and milled down to a dust-free fraction with grains ranging from 1.6 to 5 mm. All the provided 50g were milled and thoroughly homogenized in a ball mill at ETH Zürich to a grain size of less than 100 µm and filled in plastic vials of 0.5 g aliquots for distribution.

IAEA-C2 is a freshwater travertine from Bavaria distributed by IAEA as a powder which was homogenized further in a ball mill to a grain size of less than 100 µm at ETH Zürich, and filled in 0.5 g aliquots in plastic vials for distribution.

ETH-4 is synthetic carbonate with intermediate formation temperature and the same bulk isotope composition as ETH-2 (see Bernasconi 2018 for details).

MERCK (lot no. B1164559 515) is a ultra-pure commercially available synthetic calcite and was chosen for its very depleted  $\delta^{13}C$  and  $\delta^{18}O$ , of approximately -41.7 ‰ and -15.5 ‰

(VPDB), respectively. This sample represents an extreme case of extrapolation from the  $\delta^{47}$ - $\Delta_{47}$  space defined by the anchor materials (Fig. 2).

## 2.2 Instrumentation.

The reported data were produced with a variety of custom built (14 laboratories) and commercial (10 laboratories) preparation systems (ThermoFisher Scientific Kiel IV device and Nu instruments Nucarb). Reaction temperatures were generally 90 °C for “large-sample” custom preparation systems and 70°C for the Kiel and the NuCarb. Four mass spectrometer types were used: Thermo Fisher scientific MAT253 and 253Plus, Nu instruments Perspective and Elementar Isoprime 100. All participants contributed results they considered to be of a “publication-grade” quality.

## 2.3 Data processing, correction and error assessment.

It should be stressed that the InterCarb experiment, by design, is not intended to grade the analytical “performance” of individual laboratories. Each participating laboratory (or mass spectrometer, in the case of laboratories with several instruments) was thus randomly assigned an anonymous identifying number. Within each laboratory, analyses were grouped in different analytical sessions defined by the participants themselves. An analytical session is generally defined by a time where the behavior of the analytical system (preparation system, source tuning, backgrounds, isotope scrambling in the source) is considered to be similar. The database record of each analysis consists of a laboratory identifier; a session identifier; an analysis identifier; the name of the analyzed sample; the mass spectrometer model; the acid reaction temperature; the mass of reagent carbonate; and background-corrected  $\delta^{45}$ ,  $\delta^{46}$  and  $\delta^{47}$  values.

The only instrumental corrections to the raw data applied independently by each participating laboratory were background corrections (or “Pressure Baseline Correction”: PBL) to the ion currents/voltages (Bernasconi et al., 2013; He et al., 2012). The PBL is strongly dependent on instrument design and configuration, and varies temporally depending on many factors. This correction, therefore, can only be carried out by each participating laboratory according to its own established procedures and monitoring.

367

368 To avoid artefacts arising from different calculation/standardization procedures, rounding  
 369 errors, and  $^{17}\text{O}$  correction parameters, raw data from all laboratories were processed by a  
 370 single Python script (SI file *intercarb-src.zip*) based on data reduction, standardization and  
 371 error propagation methods described in detail in the companion paper (Daëron, this volume).  
 372 Here we briefly summarize these calculations.

373

374 Session-averaged background-corrected  $\delta^{45}$  and  $\delta^{46}$  values for each of the three anchor  
 375 samples were first used to calculate the bulk isotope composition of the working gas used in  
 376 each session, based on (a) previously reported  $\delta^{13}\text{C}_{\text{VPDB}}$  and  $\delta^{18}\text{O}_{\text{VPDB}}$  values of ETH-1, ETH-  
 377 2, and ETH-3 (Bernasconi et al., 2018), (b) the IUPAC oxygen-17 correction parameters of  
 378 Brand et al. (2010) , and (c) a temperature-dependent oxygen-18 acid fractionation factor  
 379 between  $\text{CO}_2$  and calcite of (Kim et al., 2015). This recalculation of working gas bulk  
 380 compositions avoids (small) discrepancies potentially introduced by inaccuracies in the  
 381 nominal compositions of the working gases.

382 Raw  $\Delta_{47}$  values were computed according to:

383

$$384 \quad \Delta_{47}^{\text{Raw}} = \left( \frac{R^{47}}{R^{47*}} - 1 \right)$$

385 Where  $R^{47}$  is the measured ratio and  $R^{47*}$  the calculated stochastic ratio of mass 47 over mass  
 386 44 of  $\text{CO}_2$ .

387  $\Delta_{47}^{\text{Raw}}$  values are then normalized to “absolute”  $\Delta_{47}$  values using session-specific relationships  
 388 of the form:

389

$$390 \quad \Delta_{47}^{\text{Raw}} = a \Delta_{47} + b \delta_{47} + c$$

391

392 For each session, the best-fit standardization parameters (a, b, c) are computed from an  
 393 unweighted least-squares regression treating  $\Delta_{47}^{\text{Raw}}$  as the response variable, only considering  
 394 the three anchor samples ETH-1, ETH-2, and ETH-3. Absolute  $\Delta_{47}$  values are then  
 395 computed for all replicates within that session. Standardization parameters for all sessions are  
 396 listed in (Table S1, supplementary information).

397

In the table and figures the uncertainty of the measurements is reported as 1SE of the mean either only considering uncertainties in the analyses of a given sample, or considering fully-propagated errors associated with the establishment of the reference frame. In both cases, the analytical error assigned to each individual raw  $\Delta_{47}$  analysis is equal to the pooled “external” repeatability of raw  $\Delta_{47}$  measurements for all samples (anchors and unknowns) within each session. This treatment of error is a new approach that more fully accounts for error in both the sample measurement and reference frame. A detailed description of these error estimates is presented in a companion paper (Daëron, this volume).

## 2.4 Clumped isotope composition of the ETH anchor materials.

The clumped isotope compositions of the four ETH reference materials relative to the CO<sub>2</sub> reference frame CDES were first reassessed based on new data provided by 9 laboratories that also provided HG and EG data measured during the same sessions as the ETH reference materials. Six of these laboratories reacted carbonates at 90 °C, two at 70 °C and only one at 25 °C. In keeping with tradition, we “project” the carbonate  $\Delta_{47}$  values to an acid reaction temperature of 25 °C using the acid corrections of (Petersen et al., 2019), i.e. +0.66 ‰ and +0.088 ‰, respectively, for reactions at 70 °C and 90 °C.

## 3 Results and Discussion

### 3.1 Redetermination of nominal $\Delta_{47}$ values for the ETH standards relative to heated and equilibrated CO<sub>2</sub> gases

The weighted averages of the 4 standards projected to 25°C comprising 619 analyses of the carbonate standards and 828 heated and equilibrated gases from 9 different laboratories are reported in Table 1 and Fig 4. The large number of analyses and the appropriate consideration of the errors on the anchors (CO<sub>2</sub> gases references) is a first, and allows a robust redetermination of the accepted values of the ETH reference materials with 1SE uncertainties of 2 ppm or less.

*Table 1: Newly determined nominal  $\Delta_{47}$  values of the ETH standards projected to 25°C reaction temperature using a correction factors of 0.088 ‰ and 0.066 ‰ for 90 °C and 70 °C reactions, respectively (Petersen et al. 2019). Reported error values are 1SE after*

430 *propagating analytical uncertainties associated with carbonate unknowns and equilibrated*  
431 *gas standards (Daëron, this volume)*  
432

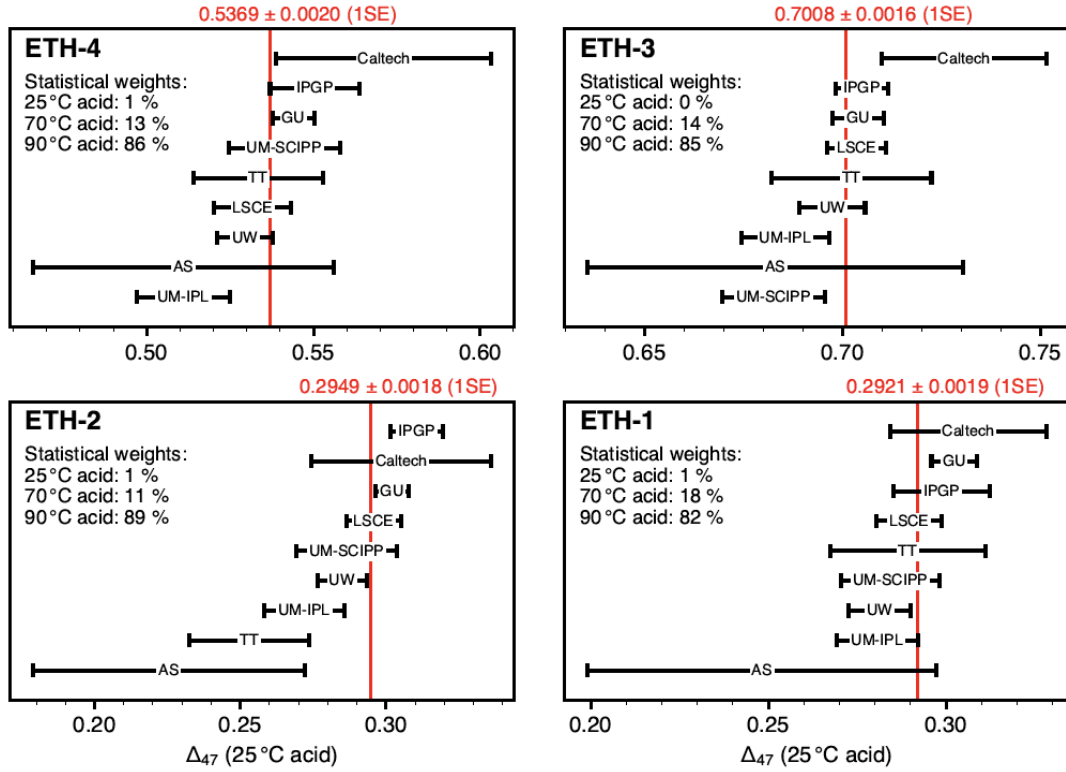
| Sample                 | ETH-1               | ETH-2               | ETH-3               | ETH-4               |
|------------------------|---------------------|---------------------|---------------------|---------------------|
| $\Delta_{47}$ (‰ CDES) | 0.2920 $\pm$ 0.0019 | 0.2949 $\pm$ 0.0018 | 0.7008 $\pm$ 0.0016 | 0.5369 $\pm$ 0.0020 |

433

434 When compared with the published values in Bernasconi et al. (2018) the average  $\Delta_{47}$  values  
435 ETH-1 and ETH-2 are respectively 0.034 and 0.039 ‰ greater than the original values,  
436 whereas ETH-3 increases by 0.010 and ETH-4 by 0.030 ‰. Such positive offset of D47  
437 values reported in Bernasconi et al. (2018) versus other laboratories has also been observed in  
438 Thaler et al. (2020).

439

440



441

442 Figure 3. New determination of  $\Delta_{47}$  values for the four ETH standards relative to the CDES  
 443 using updated  $\text{CO}_2$  equilibrium values and phosphoric acid fractionation factors from  
 444 Petersen et al. (2019). Error bars correspond to the 95 % confidence limits taking into  
 445 account fully propagated errors (ie. taking into account errors in both unknown and anchor  
 446 analyses). The red numbers are the error-weighted average values. Note different horizontal  
 447 scales for the different samples.

448

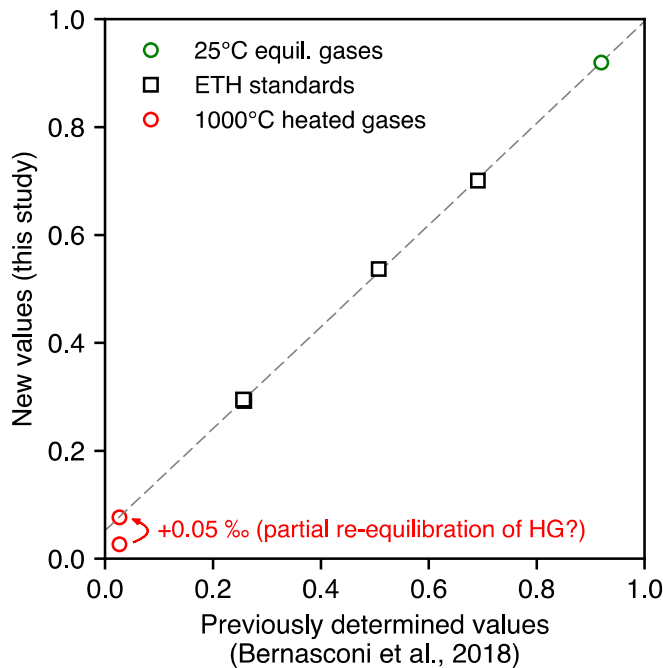
449

450 The observation that these changes in nominal values decrease as  $\Delta_{47}$  increases suggests a  
 451 simple hypothesis to explain this discrepancy: in the original study of Meckler et al. (2014),  
 452 the carbonate samples and the heated/equilibrated  $\text{CO}_2$  gases experienced different analytical  
 453 procedures (i.e. measurements made for HGs passed through a unique set of capillaries and  
 454 used the bellows-mode whereas carbonates were measured using the microvolume and gas-  
 455 depletion). The potential effects of partial re-equilibration of the heated and 25 °C  
 456 equilibrated gases at room temperature could be significant for the former yet remain  
 457 minuscule for the latter, leading to an overestimation of  $\Delta_{47}$  scale compression and thus  
 458 applied stretching of the  $\Delta_{47}$  scale towards theoretical values. The observed changes in



apparent carbonate standard  $\Delta_{47}$  values may therefore simply reflect partial re-equilibration of heated gases at least at the time of measurements at ETH (and reported in Meckler et al., 2014), increasing their values in the original study by about 0.05 ‰ (Figure 4)

**Figure 4:** New nominal  $\Delta_{47}$  values for the ETH standards compared to previously reported ones. Apparent changes in these values scale linearly with the  $\Delta_{47}$  difference between carbonate samples and 25 °C equilibrated CO<sub>2</sub>, suggesting that  $\Delta_{47}$  values of heated gases in the original study may have been biased by  $\sim +0.05$  ‰ through partial re-equilibration at room temperature at ETH at the time the measurements were performed.



It has been suggested previously that ETH 1 and ETH 2 should be indistinguishable in  $\Delta_{47}$  and close to stochastic distribution (Müller, Violay, et al., 2017), based on comparison with stochastic carbonates which showed that the  $\Delta_{47}$  of ETH 1 and ETH 2 is very close (approx. 0.006 ‰ higher) to the values of aliquots of the same carbonates heated at 1000°C. However, additional test measurements in multiple laboratories of samples heated at  $>1000^{\circ}\text{C}$  are necessary to confirm this observation.

One laboratory (TT) observed a large difference in the value for ETH-1 and ETH-2, which is not what we should expect because both carbonates were heated under identical conditions. Their values of ETH-3 and 4 are very similar to other laboratories. The reason for these

inconsistencies is probably due to the fact that ETH-1 was only measured four times with only limited number of HG/EG and ETH-2 and ETH-4 were not measured in the same session. In general, the labs with the smallest number of replicate measurements have uncertainties that are systematically larger. These results highlight the importance of correction procedures in clumped isotope analysis. Sufficient replication of both standards and samples is critical and offsets can arise when comparing results from different sessions. Due to these difficulties it is good practice to spread replicates of the same sample in different sessions over longer periods of time to obtain accurate results and follow a standard:sample ratio  $\geq 1$ .

Based on the results above, the difference between the average of ETH1/2 and ETH-3 is reduced by 0.0265 ‰ thus leading to a compression of the scale by about 6.1 % compared to the values reported by Bernasconi et al. (2018). As a consequence, the slopes of published temperature calibrations produced with carbonate standardization (Bernasconi et al., 2018; Jautzy et al., 2021; Meinicke et al., 2020; Peral et al., 2018; Piasecki et al., 2019) will become slightly shallower, with more positive y-intercepts. If results from previous publications are recalculated with the new standard values (see section 3.4), however, changes in calculated formation temperatures will be negligible. For this reason, when comparing data from publications using old accepted values of the ETH Standards for standardization (either those published by Meckler et al. (2013) or those recalculated with the IUPAC parameters by Bernasconi et al. (2018) to new ones, it is recommended to directly compare the reconstructed temperatures rather than recalculating the  $\Delta_{47}$ . Full recalculation of old measurements require the availability of the entire dataset including standards and to use the same correction procedures such as the averaging method used in the original publications and is described in section 3.5. Based on the results above, the difference between the average of ETH1/2 and ETH-3 is reduced by 0.0265 ‰ thus leading to a compression of the scale by about 6.1 % compared to the values reported by Bernasconi et al. (2018). The slopes of published temperature calibrations produced with carbonate standardization (Bernasconi et al., 2018; Jautzy et al., 2021; Meinicke et al., 2020; Peral et al., 2018; Piasecki et al., 2019) will therefore become slightly shallower, with more positive y-intercepts. If results from previous publications and calibrations are recalculated with the new standard values (see below), changes in calculated formation temperatures will be negligible. For this reason, when comparing data from publications using old accepted values of the ETH Standards for

standardization (either those published by Meckler et al. (2013) or those recalculated with the IUPAC parameters by Bernasconi et al. (2018) to new ones, it is recommended to directly compare the reconstructed temperatures rather than recalculating the  $\Delta_{47}$ . Full recalculation of old measurements requires the availability of the entire dataset including standards used in the original publications.

Table 2. Results of CO<sub>2</sub> reference gases derived values of ETH reference carbonate materials, after acid temperature corrections based on Petersen et al. (2019). N is the number of replicate analyses of carbonate samples. Uncertainties reported here represent both reference frame errors and reproducibility errors.

| Laboratory | ETH-1 |        |        | ETH-2 |        |        | ETH 3 |        |        | ETH 4 |        |        |
|------------|-------|--------|--------|-------|--------|--------|-------|--------|--------|-------|--------|--------|
|            | N     | D47    | ± 95 % | N     | D47    | ± 95 % | N     | D47    | ± 95 % | N     | D47    | ± 95 % |
| AS         | 4     | 0.2482 | 0.049  | 4     | 0.2255 | 0.0466 | 4     | 0.6831 | 0.0474 | 4     | 0.5111 | 0.0452 |
| Caltech    | 8     | 0.3062 | 0.022  | 8     | 0.3052 | 0.0309 | 8     | 0.7307 | 0.0207 | 7     | 0.5711 | 0.0323 |
| GU         | 19    | 0.3023 | 0.0064 | 18    | 0.3021 | 0.0058 | 15    | 0.7039 | 0.0066 | 12    | 0.544  | 0.0063 |
| IPGP       | 5     | 0.2988 | 0.0136 | 11    | 0.3105 | 0.0091 | 20    | 0.7049 | 0.0066 | 5     | 0.5504 | 0.0135 |
| LSCE       | 34    | 0.2896 | 0.0091 | 23    | 0.2957 | 0.0093 | 55    | 0.7036 | 0.0074 | 10    | 0.5318 | 0.0115 |
| TT         | 4     | 0.2893 | 0.0219 | 4     | 0.253  | 0.0206 | 5     | 0.7023 | 0.0202 | 4     | 0.5334 | 0.0194 |
| UM-IPL     | 14    | 0.2806 | 0.0115 | 13    | 0.272  | 0.0139 | 15    | 0.6856 | 0.0111 | 12    | 0.511  | 0.0139 |
| UM-SCIPP   | 16    | 0.2844 | 0.0138 | 13    | 0.2865 | 0.0171 | 15    | 0.6826 | 0.0129 | 15    | 0.5413 | 0.0167 |
| UW         | 58    | 0.2813 | 0.0088 | 51    | 0.2849 | 0.0085 | 59    | 0.6974 | 0.0083 | 47    | 0.5294 | 0.0083 |

, it to reach a community consensus on how to report clumped isotope measurements, to reduce the confusion in the literature related to different scales used in this rapidly evolving field. Currently  $\Delta_{47}$  are reported for different temperatures of phosphoric acid digestion, mostly projected to 25°C (CDES25) but also to 70 °C (CDES70) or 90 °C (CDES90) reaction. In the literature, phosphoric acid correction factors used by different research groups to convert results from 90 to 25°C reaction have varied between 0.069 ‰ (Wacker et al., 2014) to 0.092 ‰ e.g. (Bonifacie et al., 2017) thus representing a significant source of confusion and uncertainty. In contrast, carbonate-based standardization eliminates different phosphoric acid correction factors as a source of uncertainty and provides a consistent framework to report  $\Delta_{47}$  without uncertainties related to the reaction temperature.

One possible option is to keep a projection to 25 °C acid reaction temperature, consistent with most clumped isotope publications to date. At first sight it could appear that maintaining the projection to a reaction temperature of 25 °C would make a comparison of new data to older publications easier. However, we emphasize that previously published  $\Delta_{47}$  data, either normalized with heated and equilibrated gases or with carbonates, cannot be

directly compared to the new ones. All previous carbonate-normalized data have to be recalculated in order to fully take into account the new standard values reported here. For HG/EG normalized data it is not possible to convert to the new scale, unless an adequate number of ETH standards had been measured during the same period of time. In this case, it may be more reasonable to directly compare reconstructed temperatures with a laboratory specific calibration, rather than  $\Delta_{47}$  values as suggested by Petersen et al. (2019).

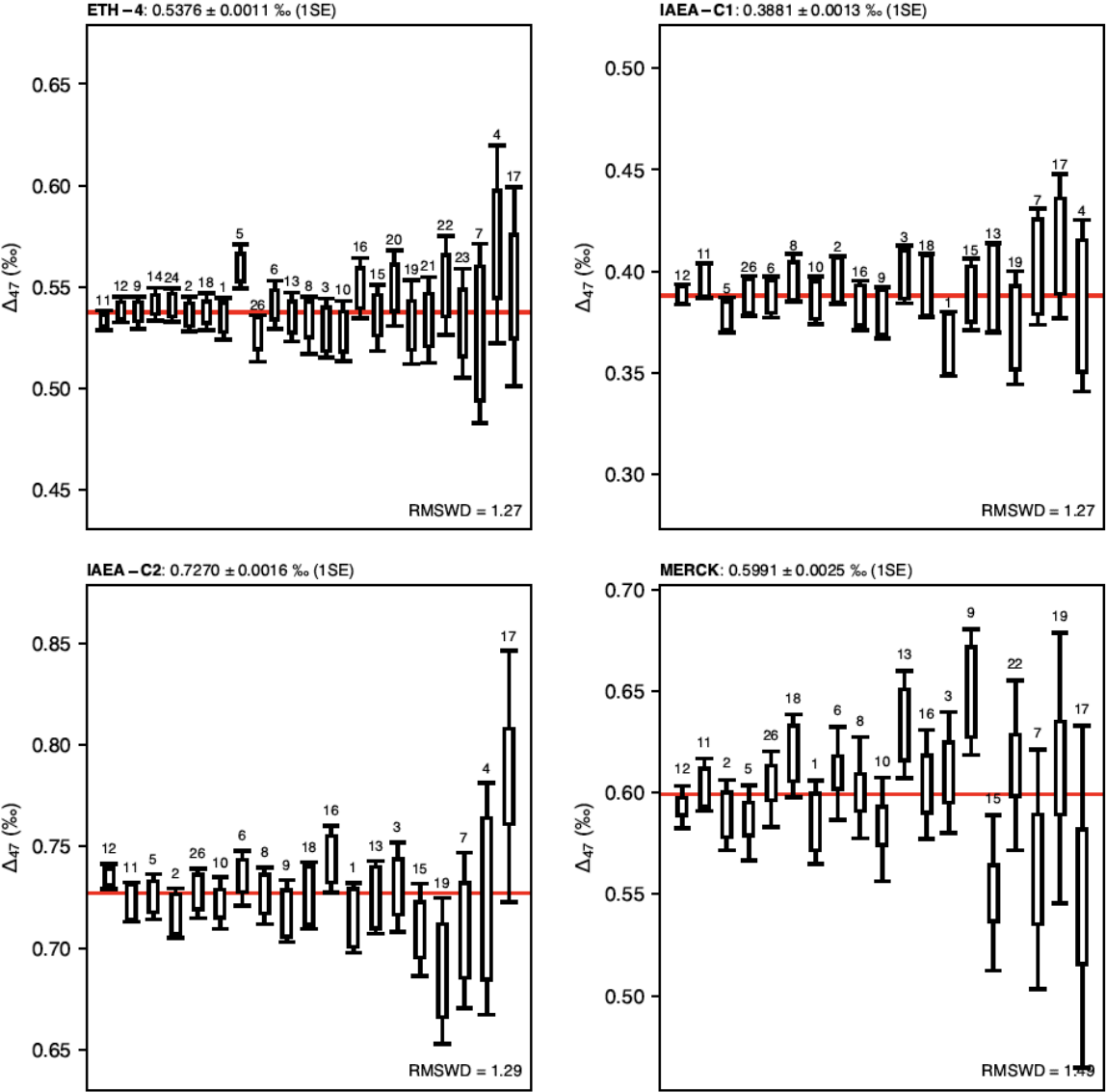
Maintaining a conversion to 25°C has the benefit that past reconstructed temperatures can be directly compared if samples and calibration data were treated equally. Reasons to change to projection to a different temperature, are that the values presented in this study were determined mostly with acid digestion at 90 °C and that the vast majority of laboratories currently active have moved to higher reaction temperatures of either 70 or 90 °C. Thus the conversion to 25°C reaction values could also be abandoned in favor of reporting data directly for acid reaction temperatures of 70 or 90 °C.

### 3.2 InterCarb results

We obtained data from 26 Mass spectrometers from 23 Laboratories. The  $\Delta_{47}$  of the 4 unknown samples normalized to the new community-derived values of the ETH reference materials averaged per individual analytical session and mass spectrometer are listed in Table 3. Sample mean values obtained in each mass spectrometer are shown in Figure 5. The details of each analytical session, including the number of samples and standards measured, the composition of the working standard, the scaling parameters and the reproducibilities of the individual sessions are given in the supplementary information (Table S1). Some laboratories reported data for only a subset of the unknown samples, and both replication levels and analytical repeatabilities are very different from laboratory to laboratory (Table S1).

41

568



569

570

21

42

43

571

572

573 *Figure 5. Final results by laboratory. Error bars correspond to fully propagated 95 %*  
 574 *confidence limits, taking into account errors in both unknown and anchor analyses.*  
 575 *Boxes correspond to 95 % confidence limits not accounting for normalization errors*  
 576 *(i.e. only taking into account errors in unknown analyses). Results are sorted by*  
 577 *increasing analytical errors and laboratories are identified by number. Overall*  
 578 *error weighted average  $\Delta_{47}$  values are displayed as solid red lines and reported in*  
 579 *top left label. Root mean squared weighted deviation values (RMSWD, equivalent to*  
 580 *reduced  $\chi^2$ ) are reported in bottom right labels. All plots have the same vertical*  
 581 *scale.*

582

583 Laboratories with stronger analytical constraints (better in-lab repeatability of  $\Delta_{47}$   
 584 measurements and/or greater number of analyses) generally converge towards the overall  
 585 mean value for each unknown sample. This observation suggests that inter-laboratory  
 586 variability observed here is largely due to random errors that can be alleviated by replication,  
 587 even for laboratories with relatively large analytical errors on individual measurements. It is  
 588 also notable that fully propagated analytical errors (i.e., taking into account uncertainties in  
 589 the standardization procedure) can be substantially larger than the errors based only on the  
 590 uncertainty associated with unknown sample analyses which is what is generally reported in  
 591 the literature. The increase in error bar is also related to in-lab repeatability and the number  
 592 of standards measured. In addition, this effect increases for unknown samples whose  
 593 compositions lie outside the “anchor triangle” defined by ETH-1/2/3 (and are the highest for  
 594 MERCK as it is farthest from the anchor triangle), consistent with the models of Daëron (this  
 595 volume) and Kocken et al., (2019).

596

597 Table 3. Results of individual sessions by laboratory/mass spectrometer.

| Lab | Session | ETH-4           | IAEA-C1         | IAEA-C2         | MERCK           | RMSWD |
|-----|---------|-----------------|-----------------|-----------------|-----------------|-------|
| 01  | 01      | 0.5367 ± 0.0160 |                 |                 |                 |       |
|     | 02      | 0.4957 ± 0.0262 |                 |                 |                 |       |
|     | 03      | 0.5360 ± 0.0057 | 0.3643 ± 0.0080 | 0.7149 ± 0.0088 | 0.5855 ± 0.0105 |       |
|     | all     | 0.5345 ± 0.0052 | 0.3643 ± 0.0080 | 0.7149 ± 0.0088 | 0.5855 ± 0.0105 | 1.07  |
| 02  | 01      | 0.5355 ± 0.0047 | 0.3976 ± 0.0069 | 0.7187 ± 0.0070 | 0.5888 ± 0.0099 |       |
|     | 02      | 0.5432 ± 0.0115 | 0.3893 ± 0.0126 | 0.7120 ± 0.0130 | 0.5897 ± 0.0200 |       |
|     |         |                 |                 |                 |                 |       |
|     | all     | 0.5367 ± 0.0044 | 0.3957 ± 0.0061 | 0.7172 ± 0.0061 | 0.5889 ± 0.0089 | 0.48  |
| 03  | 01      | 0.5382 ± 0.0120 |                 |                 |                 |       |
|     | 02      | 0.5246 ± 0.0095 | 0.3985 ± 0.0073 | 0.7301 ± 0.0112 | 0.6099 ± 0.0152 |       |
|     | all     | 0.5298 ± 0.0074 | 0.3985 ± 0.0073 | 0.7301 ± 0.0112 | 0.6099 ± 0.0152 | 0.89  |
| 04  | 01      | 0.5710 ± 0.0249 | 0.3830 ± 0.0215 | 0.7241 ± 0.0291 |                 |       |
| 05  | 01      | 0.5722 ± 0.0090 | 0.3727 ± 0.0064 | 0.7261 ± 0.0085 | 0.6159 ± 0.0148 |       |
|     | 02      | 0.5373 ± 0.0098 | 0.3915 ± 0.0081 | 0.7249 ± 0.0102 | 0.5455 ± 0.0168 |       |
|     | 03      | 0.5697 ± 0.0101 | 0.3737 ± 0.0094 | 0.7239 ± 0.0117 | 0.5848 ± 0.0180 |       |
|     | all     | 0.5603 ± 0.0055 | 0.3786 ± 0.0044 | 0.7252 ± 0.0057 | 0.5852 ± 0.0094 | 1.65  |
| 06  | 01      | 0.5445 ± 0.0189 | 0.3873 ± 0.0140 | 0.7102 ± 0.0179 | 0.5975 ± 0.0313 |       |
|     | 02      | 0.5375 ± 0.0090 |                 |                 |                 |       |
|     | 03      | 0.5444 ± 0.0091 | 0.3898 ± 0.0070 | 0.7456 ± 0.0095 | 0.6212 ± 0.0159 |       |
|     | 04      |                 | 0.3835 ± 0.0091 | 0.7275 ± 0.0120 | 0.5951 ± 0.0204 |       |
|     | all     | 0.5413 ± 0.0061 | 0.3874 ± 0.0052 | 0.7345 ± 0.0069 | 0.6094 ± 0.0117 | 0.82  |
| 07  | 01      | 0.5272 ± 0.0225 | 0.4023 ± 0.0146 | 0.7087 ± 0.0195 | 0.5621 ± 0.0301 |       |
| 08  | 01      | 0.5299 ± 0.0205 | 0.4241 ± 0.0169 | 0.7199 ± 0.0203 | 0.5724 ± 0.0344 |       |
|     | 02      | 0.5540 ± 0.0241 | 0.3793 ± 0.0204 | 0.7453 ± 0.0227 | 0.6383 ± 0.0455 |       |
|     | 03      | 0.5154 ± 0.0257 | 0.3743 ± 0.0235 | 0.7167 ± 0.0227 | 0.5896 ± 0.0453 |       |
|     | 04      | 0.5937 ± 0.0344 | 0.3711 ± 0.0288 | 0.7036 ± 0.0320 | 0.5857 ± 0.0590 |       |
|     | 05      | 0.6080 ± 0.0361 | 0.3817 ± 0.0294 | 0.7587 ± 0.0349 | 0.7097 ± 0.0659 |       |
|     | 06      | 0.5475 ± 0.0277 | 0.3829 ± 0.0215 | 0.7132 ± 0.0243 | 0.6036 ± 0.0530 |       |
|     | 07      | 0.4676 ± 0.0341 | 0.3885 ± 0.0296 | 0.6781 ± 0.0340 | 0.4750 ± 0.0650 |       |
|     | 08      | 0.5855 ± 0.0348 | 0.4378 ± 0.0257 | 0.7537 ± 0.0316 | 0.6144 ± 0.0603 |       |
|     | 09      | 0.5549 ± 0.0348 | 0.4109 ± 0.0282 | 0.7494 ± 0.0376 | 0.6143 ± 0.0585 |       |
|     | 10      | 0.5651 ± 0.0305 | 0.3919 ± 0.0242 | 0.7247 ± 0.0365 | 0.6552 ± 0.0509 |       |
|     | 11      | 0.5024 ± 0.0359 | 0.3579 ± 0.0302 | 0.7312 ± 0.0389 | 0.6606 ± 0.0642 |       |
|     | 12      | 0.5021 ± 0.0467 | 0.4121 ± 0.0361 | 0.6896 ± 0.0424 | 0.6147 ± 0.0742 |       |
|     | 13      | 0.4732 ± 0.0261 | 0.3457 ± 0.0273 | 0.7392 ± 0.0310 | 0.6180 ± 0.0525 |       |
|     | 14      | 0.5300 ± 0.0198 | 0.4251 ± 0.0165 | 0.7419 ± 0.0218 | 0.6197 ± 0.0338 |       |
|     | 15      | 0.4682 ± 0.0268 | 0.4031 ± 0.0213 | 0.7096 ± 0.0262 | 0.5346 ± 0.0462 |       |
|     | 16      | 0.5550 ± 0.0380 | 0.3896 ± 0.0289 | 0.7219 ± 0.0386 | 0.5557 ± 0.0724 |       |
|     | all     | 0.5310 ± 0.0072 | 0.3970 ± 0.0060 | 0.7257 ± 0.0071 | 0.6023 ± 0.0127 | 1.05  |
| 09  | 01      | 0.4809 ± 0.0202 |                 |                 |                 |       |
|     | 02      | 0.5439 ± 0.0067 |                 |                 |                 |       |
|     | 03      | 0.5579 ± 0.0090 |                 | 0.7427 ± 0.0239 |                 |       |
|     | 04      | 0.5259 ± 0.0066 | 0.3797 ± 0.0064 | 0.7153 ± 0.0083 | 0.6496 ± 0.0158 |       |
|     | all     | 0.5373 ± 0.0041 | 0.3797 ± 0.0064 | 0.7182 ± 0.0078 | 0.6496 ± 0.0158 | 2.13  |

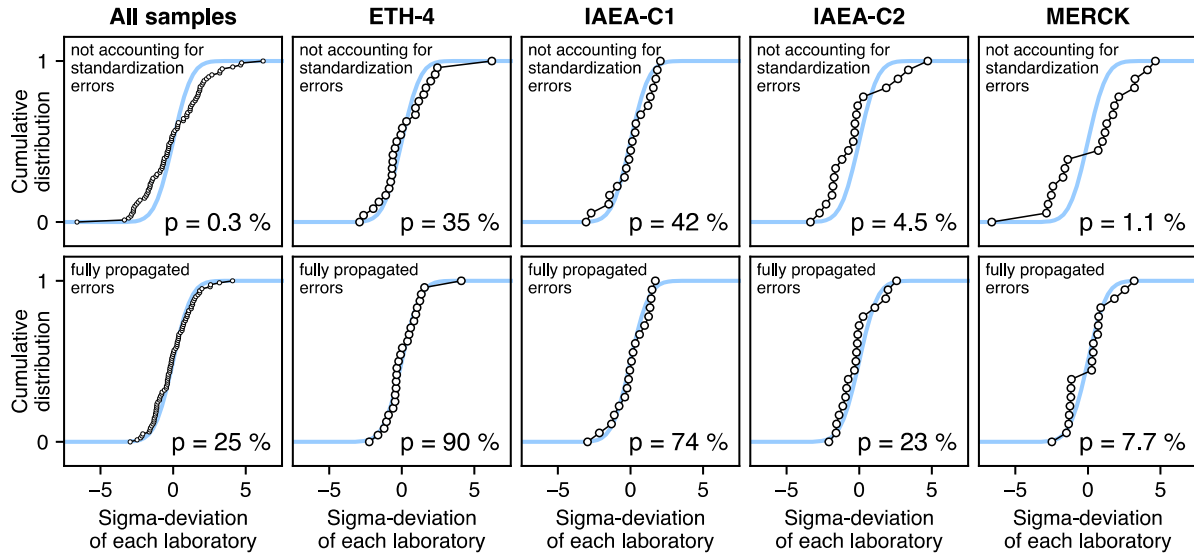
As seen in Table S1, there are stark differences in the total number of replicate analyses and the typical  $\Delta_{47}$  reproducibility achieved in different laboratories. As a result, final uncertainties in the average  $\Delta_{47}$  values of unknown samples vary considerably (Figure 5). It is thus not very useful to quantify inter-laboratory reproducibility in terms of a single, overall scatter of  $\Delta_{47}$  values. Inter-laboratory variability should be small among laboratories with small analytical uncertainties, and larger among laboratories with few replicate analyses and/or poor analytical repeatability. We may still, however, assess whether inter-lab discrepancies are significantly larger than expected from in-lab analytical uncertainties, i.e., whether we can detect the effects of hypothetical unrecognized sources of scatter beyond known analytical errors.

In order to do so, we compute the “number-of-sigma” deviation obtained by each laboratory for each unknown sample, relative to that sample’s overall weighted average value. For example, the sigma-deviation for sample ETH-4 and Lab01 is equal to

(0.5345 – 0.5376) / 0.0052 = –0.60 and that for MERCK and Lab13 is equal to  
 (0.6335 – 0.5991) / 0.0135 = +2.55. If the analytical errors reported in Table 4 are reasonably  
 accurate, we expect the population of sigma-deviations among all labs to be distributed as the  
 canonical Gaussian distribution ( $\mu = 0$ ;  $\sigma = 1$ ), and we can test this prediction using  
 established statistical methods such as a Kolmogorov–Smirnov test of normality (Massey,  
 1951). We carried out this test for the two cases: only considering the error of sample  
 replication (Fig. 6, lower row) and secondly including the normalization error, (i.e. the fully  
 propagated error (Fig. 6 upper row). If we neglect uncertainties arising from standardization  
 (the “allogenic” errors of Daëron et al., this volume), the sigma-deviations are no longer  
 normally distributed ( $p = 0.003$ , upper-left panel). As shown in the lower-left panel of Figure  
 6, the distribution of sigma-deviations for all labs and all samples is statistically  
 indistinguishable from the expected normal distribution ( $p = 0.25$ ) when considering fully  
 propagated analytical errors. Figure 6 also illustrates that neglecting standardization errors  
 does not strongly affect the normality of sigma-deviations for ETH-4 and IAEA-C1, both of  
 which have  $\delta_{47}$  and  $\Delta_{47}$  values within the range covered by the three anchor samples. By  
 contrast, sigma-deviations for unknowns with “exotic” isotopic compositions (IAEA-C2 and  
 especially MERCK) are only normally distributed if standardization uncertainties are  
 correctly accounted for.

Based on these tests, we conclude that the inter-lab scatter observed in the InterCarb data set  
 is neither smaller nor larger than expected from the analytical uncertainties computed within  
 each laboratory, as long as standardization errors are taken into account. This important  
 finding implies that, at least for the time being, we can rule out any systematic inter-  
 laboratory discrepancies in carbonate-standardized  $\Delta_{47}$  measurements, which arguably  
 constitutes an important milestone in the progress of clumped-isotope techniques.





**Figure 6:** Kolmogorov–Smirnov tests of normality for the sigma-deviations obtained in each laboratory (circular markers), either neglecting standardization uncertainty (upper row) or considering fully propagated analytical errors (lower row). Lower-right corner Kolmogorov–Smirnov p-values correspond to the null hypothesis that the sigma-deviations are normally distributed with a mean of zero and a standard deviation of 1. Blue lines correspond to the canonical Gaussian distribution ( $\mu = 0$ ;  $\sigma = 1$ ).

### 3.3 Effects of acid reaction temperature and IRMS models

To test for the possible effect of acid temperature, a commonly discussed cause for different slopes in the published temperature calibration curves (Came et al., 2014; Fernandez et al., 2014; Swart et al., 2019), we plot the values of the laboratories reacting at 70 vs. those at 90°C (Fig. 7, Tab 7). Out of 26 mass spectrometers/extraction systems 11 react samples at 70 °C and 15 at 90 °C.  $\Delta_{47}$  values averaged by acid temperature are statistically indistinguishable for all of the unknowns. Thus we can conclude that relative  $\Delta_{47}$  differences between  $\text{CO}_2$  evolved from different samples are independent of acid reaction temperature within the range of experimental conditions covered here, and for a very wide range of clumped isotope compositions spanning 0.392 (marbles) to 0.729 ‰ (carbonates formed at ambient temperatures).

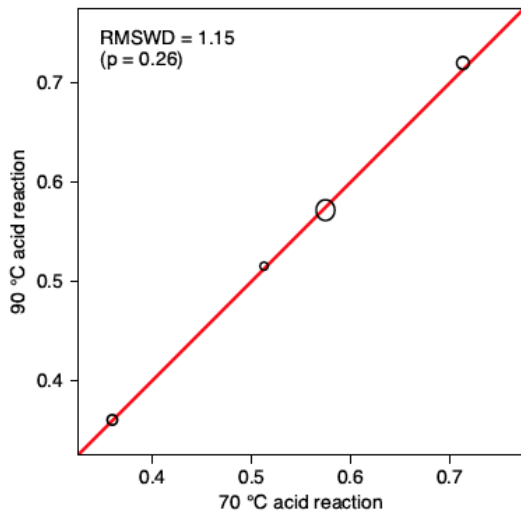


Figure 7. Results grouped by acid reaction temperature are statistically indistinguishable (see also Table 7).

|                          | $\Delta_{47}$ (‰, $\pm 1SE$ ) |                     | Difference                            |
|--------------------------|-------------------------------|---------------------|---------------------------------------|
|                          | 70 °C                         | 90 °C               |                                       |
| ETH-4                    | $0.5380 \pm 0.0016$           | $0.5396 \pm 0.0016$ | $0.0016 \pm 0.0023$                   |
| IAEA-C1                  | $0.3913 \pm 0.0020$           | $0.3921 \pm 0.0019$ | $0.0008 \pm 0.0027$                   |
| IAEA-C2                  | $0.7264 \pm 0.0024$           | $0.7324 \pm 0.0023$ | $0.0060 \pm 0.0033$                   |
| MERCK                    | $0.5985 \pm 0.0036$           | $0.5937 \pm 0.0038$ | $-0.0049 \pm 0.0052$                  |
| <b>Avg (all samples)</b> |                               |                     | <b><math>0.0009 \pm 0.0018</math></b> |

Table 7 Averages divided by acid temperature

The error weighted results separated by mass spectrometer type and design of associated preparation lines, another postulated source of differences for clumped isotope measurements (Swart et al., 2019), are shown in Figure 7 and Table 8. Out of the 25 participating laboratories, 8 groups are equipped with a Nu Perspective, 16 groups with different versions of a Thermo MAT253, and one laboratory uses an Isoprime 100. Most results are statistically indistinguishable across instruments. Only IAEA-C2 yielded a significantly (+.26 sigma) higher ( $\Delta\Delta_{47}$  of +0.0111 and -0.012 vs Nu perspective and MAT253, respectively ) mean  $\Delta_{47}$

value when analyzed with an Isoprime 100 versus other instrument models (but note that all of the Isoprime 100 data comes

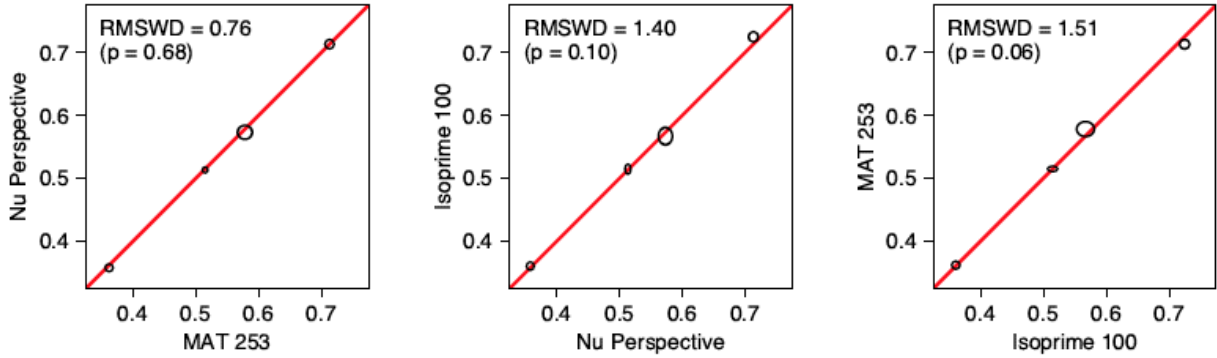


Figure 7. Effect of different mass spectrometer type and associated preparation lines for the 4 standards.

Table 8. Mass spectrometer effects, Errors-weighted by a factor of  $1/\sigma^2$

|                          | MAT 253 vs<br>Nu Perspective           | Nu Perspective<br>Isoprime 100        | Isoprime 100<br>vs MAT 253             |
|--------------------------|--|---------------------------------------|--|
| ETH-4                    | $-0.0002 \pm 0.0024$                   | $-0.0013 \pm 0.0036$                  | $0.0015 \pm 0.0035$                    |
| IAEA-C1                  | $0.0034 \pm 0.0032$                    | $-0.0025 \pm 0.0035$                  | $-0.0009 \pm 0.0033$                   |
| IAEA-C2                  | $-0.0002 \pm 0.0038$                   | $-0.0110 \pm 0.0043$                  | $0.0112 \pm 0.0040$                    |
| MERCK                    | $0.0043 \pm 0.0057$                    | $0.0056 \pm 0.0068$                   | $-0.0099 \pm 0.0066$                   |
| <b>Avg (all samples)</b> | <b><math>-0.0019 \pm 0.0020</math></b> | <b><math>0.0023 \pm 0.0024</math></b> | <b><math>-0.0004 \pm 0.0023</math></b> |

from a single laboratory). Inter-instrument differences averaged over all four samples (bottom row of Table 8) remain, however, indistinguishable from zero. Thus, any potential biases introduced by the use of different mass spectrometer models and/or the design of the preparation line are

$$\Delta_{47}^{\text{old}} = u + v \delta_{47} + w^{\text{raw}} \Delta_{47}$$

$$\Delta_{47}^{\text{new}} = x + y \delta_{47} + z^{\text{raw}} \Delta_{47}$$

we can rearrange the above equations to express  $\Delta_{47}^{\text{new}}$  as an affine function of  $(\delta_{47}, \Delta_{47}^{\text{old}})$ :

$$\Delta_{47}^{\text{new}} = a + b \delta_{47} + c \Delta_{47}^{\text{old}} \quad (\text{eq. X})$$

55

699 with

700

701 
$$a = x - u/w$$

702 
$$b = y - v/w$$

703 
$$c = z/w$$

704

705 The above reasoning also applies if we choose to define  $\delta_{47}$  relative to a fixed CO<sub>2</sub>  
 706 composition, e.g., relative to stochastic VPDB-CO<sub>2</sub> ( $\delta^{13}\text{C}_{\text{VPDB}} = 0$ ;  $\delta^{18}\text{O}_{\text{VSMOW}} \approx 41.5 \text{ ‰}$ ;  $\Delta_{47} =$   
 707  $0$ ,  $R_{47} = 4.834 \times 10^{-5}$ ). In that case, numerical values of the parameters ( $a, b, c$ ) may be used to  
 708 compute  $^{\text{new}}\Delta_{47}$  for any unknown sample based only on its  $\delta_{47}$  and  $^{\text{old}}\Delta_{47}$  values (the former  
 709 being defined relative to VPDB-CO<sub>2</sub> and computed assuming an acid  $^{18}\text{O}/^{16}\text{O}$  fractionation  
 710 factor of 1.01025).

711

712 Computing the numerical values of ( $a, b, c$ ) requires knowing  $\delta_{47}$ ,  $^{\text{old}}\Delta_{47}$ , and  $^{\text{new}}\Delta_{47}$  for three  
 713 anchor samples. For instance, considering ETH-1, ETH-2, and ETH-3:

714

715

| 716 | <b>Anchor</b> | <b><math>\delta_{47}</math></b> | <b><math>^{\text{old}}\Delta_{47}</math></b> | <b><math>^{\text{new}}\Delta_{47}</math></b> |
|-----|---------------|---------------------------------|--|--|
| 717 | ETH-1         | 0.010                           | 0.258  | 0.2921                                       |
| 718 | ETH-2         | -28.375                         | 0.256  | 0.2949                                       |
| 719 | ETH-3         | 0.538                           | 0.691  | 0.7008                                       |

720

721

722 Writing equation X in matrix form yields:

723

$$724 \begin{pmatrix} 1 & \delta_{47}(\text{ETH } 1) & ^{\text{old}}\Delta_{47}(\text{ETH } 1) \\ 1 & \delta_{47}(\text{ETH } 2) & ^{\text{old}}\Delta_{47}(\text{ETH } 2) \\ 1 & \delta_{47}(\text{ETH } 3) & ^{\text{old}}\Delta_{47}(\text{ETH } 3) \end{pmatrix} \begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} ^{\text{new}}\Delta_{47}(\text{ETH } 1) \\ ^{\text{new}}\Delta_{47}(\text{ETH } 2) \\ ^{\text{new}}\Delta_{47}(\text{ETH } 3) \end{pmatrix}$$

725

726 thus:

727

$$\begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} 1 & \delta_{47}(\text{ETH } 1) & {}^{\text{old}}\Delta_{47}(\text{ETH } 1) \\ 1 & \delta_{47}(\text{ETH } 2) & {}^{\text{old}}\Delta_{47}(\text{ETH } 2) \\ 1 & \delta_{47}(\text{ETH } 3) & {}^{\text{old}}\Delta_{47}(\text{ETH } 3) \end{pmatrix}^{-1} \begin{pmatrix} {}^{\text{new}}\Delta_{47}(\text{ETH } 1) \\ {}^{\text{new}}\Delta_{47}(\text{ETH } 2) \\ {}^{\text{new}}\Delta_{47}(\text{ETH } 3) \end{pmatrix}$$

729

730

731 i.e.

732

$${}^{\text{new}}\Delta_{47} = 0.048529 - 0.000165 \times \delta_{47} + 0.944081 \times {}^{\text{old}}\Delta_{47} \quad (\text{eq. Y})$$

734

735 Equation (Y) is the unique affine function linking the  ${}^{\text{old}}\Delta_{47}$  and  ${}^{\text{new}}\Delta_{47}$  values of ETH-1/2/3. In  
 736 this case, it is clear that the conversion from the old reference frame to the new one is not  
 737 very sensitive to  $\delta_{47}$  values: for unknown samples with  $\delta_{47}$  values within  $\pm 6$  ‰ of ETH-1, the  
 738 effect of the second term in eq. Y is less than 1 ppm, and less than 3 ppm for unknowns  
 739 within  $\pm 18$  ‰ of ETH-1. In such cases, the conversion may be simplified as a simpler affine  
 740 transformation of  $\Delta_{47}$ , akin to a more traditional two-point normalization (e.g. VSMOW-  
 741 VSLAP undetectable when using carbonate standardization. Sample sizes used for individual  
 742 measurements ranged from 90-120  $\mu\text{g}$  for Kiel IV preparation system,  $\sim 500$   $\mu\text{g}$  for NuCarb  
 743 individual acid bath systems to a range of 3 to 12 mg for samples reacted in common acid  
 744 bath custom-built extraction lines. As all small sample measurements are carried out at 70°C  
 745 and large ones at 90°C, we can also conclude that there is no significant effect of sample sizes  
 746 and variations in sample to acid ratios in the final results and likely a result of random  
 747 uncertainty.

748

749 4. Guidelines for laboratory optimization and improving data quality in the future.

750

751 The results of this inter-laboratory exercise support the use of carbonate standardization for  
 752 clumped isotope measurements. When considering all laboratories, the standard deviation of  
 753 laboratory averages for the 4 unknowns range from 11 to 18 ppm for the samples requiring no  
 754 to moderate extrapolation, to 25 ppm for MERCK, an extreme case of extrapolation. This  
 755 spread across laboratories is still relatively large, and on the same magnitude to those  
 756 obtained by HG-EG normalization if we consider what has been reported for 4 laboratories in  
 757 (Dennis et al., 2011) and the scatter in the values reported by the 9 laboratories that provided  
 758 data for the redetermination of the accepted values of ETH-1 to 3 in this study. However,

based on this study, we can clearly state that the large scatter is dominated by random errors, and there are strategies to improve the repeatability within each laboratory.

The scatter in the InterCarb data, (Fig. 5) is especially influenced by laboratories that show the largest errors in the individual sample reproducibility and have a significantly magnified normalization error induced by a small number of replicates of anchors. This observation underscores the necessity of increased replication and of the measurement of a sufficient number of standards to produce data of the quality that is required for meaningful interpretations. Special care needs to be taken when reconstructing small temperature changes, such as for the reconstruction of climate change and/or the study of high temperature processes where the small temperature sensitivity of clumped isotope thermometer requires high analytical precision.

We can, for example, define a target laboratory internal repeatability of 10 ppm at the 95% CI as a desirable goal with the currently available analytical systems. This arbitrary target number is within the shot-noise limits of modern IRMS instruments, and is necessary for applications in paleoclimate reconstructions, one of the main applications of clumped isotopes, as it corresponds to an uncertainty of approximately  $\pm 3^{\circ}\text{C}$ . If for each sample we select only the results from laboratories that have provided data with a combined error of less than 10 ppm (1SE), inter-laboratory standard deviation ( $1\sigma$ ) becomes  $\leq 10$  ppm for ETH-4, IAEA-C1 and IAEA-C2 and  $\leq 15$  ppm for MERCK. It can be noted that this does not significantly change the average value of the reference materials.

Based on this example, it appears that with modern instrumentation from all manufacturers and with both custom built and commercially available systems used by the laboratories involved in this study, it is possible to reach this data quality and inter-laboratory consistency. The main factor to be taken into consideration is sufficient replication of both sample and standards (see Daëron, this volume, Bonifacie et al., 2017; Fernandez et al., 2017; Kocken et al., 2019). Based on this example, it appears that with modern instrumentation from all manufacturers and with both custom built and commercially available systems used by the laboratories involved in this study, it is possible to reach this data quality and inter-laboratory consistency. The main factor to be taken into consideration is sufficient replication of both sample and standards (see Daëron, this volume for a quantitative evaluation of this previously

792 suggested recommendation (Bonifacie et al., 2017; Fernandez et al., 2017; Kocken et al.,  
793 2019).

794

795 With standardization with carbonates, the principle of identical treatment of samples and  
796 standards (Werner & Brand, 2001b) is fulfilled, because in contrast to the use of HG-EG  
797 standardization, the anchors are analytically treated exactly like the samples. With this  
798 approach, we can exclude a number of factors possibly causing differences among  
799 laboratories. Two important outcomes of this study are that acid reaction temperature and  
800 instrument and preparation line design are not a cause for differences among laboratories  
801 when standardization is based on carbonates, because we do not observe a difference between  
802 reaction temperatures (at least between 70 and 90°C) and with a wide variety of custom built  
803 and commercial preparation systems. Swart et al. (2019) proposed that preparation on  
804 different lines can produce different slopes of a calibration curve although the samples,  
805 normalized to heated and equilibrated gases, were measured on the same mass spectrometer.  
806 They attributed the differences to partial equilibration of the produced CO<sub>2</sub> with the acid and/  
807 or heated metal surfaces. The conclusion of that work was that the degree of re-equilibration  
808 would depend on the design of the line, including the size of tubing, volume and quality of  
809 acid, size and shape of the reaction vessel and the presence of hot metal surfaces. The lack of  
810 resolvable differences observed in our dataset, however, indicates that while preparation line  
811 differences may affect  $\Delta_{47}$  measurements, standardization with carbonates eliminates any such  
812 effects. We note that HG-EG standardization may fail to do so (Swart et al. 2019), although  
813 we do not have evidence for this in the data reported here. Thus, when using carbonate  
814 standardization, these factors can be ignored provided the carbonate standards cover a large  
815 range in  $\Delta_{47}$ . As the acid digestion conditions (e.g., reaction times), vary with the carbonate  
816 mineralogy, possible effects on  $\Delta_{47}$  might be mineral-specific. For this reason, it would be  
817 highly desirable to produce reference materials for other minerals such as dolomite,  
818 aragonite, magnesite and siderite. For dolomite, three samples were proposed by Müller et al.,  
819 (2019) as possible reference materials for this common mineral and are available upon  
820 request.

821

822 Standardization errors could be reduced by extending the range of bulk composition of the  
823 anchor samples (e.g., as illustrated by Fig. 1 of *Daëron*, this volume), especially when  
824 samples are measured that require significant extrapolation. A sample with an extreme bulk

composition like MERCK would be a useful addition as an anchor, regardless of its  $\Delta_{47}$  value. While two heated standards for normalization are not strictly necessary, a “heated MERCK” anchor in combination with ETH-1 would allow verifying the PBL correction with greater confidence and with less replication than when using ETH-2 (keeping in mind that small quadratic components to PBL correction might introduce a significant bias over a  $\delta^{47}$  range of 60 ‰, e.g., Fig. 7 from He et al., [2012]).

### **3.4 Reporting data normalized to carbonates: definition of the “InterCarb” Carbon Dioxide Equilibration scale (I-CDES)**

An important consideration with this exercise, that there is a need for a community consensus on how to report clumped isotope measurements to reduce the confusion in the literature related to the different scales used in this rapidly evolving field. Currently  $\Delta_{47}$  data are reported for different temperatures of phosphoric acid digestion, mostly projected to 25°C (CDES25) but also to 70 °C (CDES70) or 90 °C (CDES90) reaction. In the literature, phosphoric acid correction factors used by different research groups to convert results from 90 to 25°C reaction have varied between 0.069 ‰ (Wacker et al., 2014) to 0.092 ‰ e.g. (Bonifacie et al., 2017) thus representing a significant source of confusion and uncertainty. With the introduction of carbonate standardization, as long as standards and samples share the same mineralogy, direct standardization to accepted values of the solid phases removes the need for a phosphoric acid correction, independently of the temperature at which the samples were reacted. This eliminates different phosphoric acid correction factors as a source of uncertainty and provides a consistent framework to report  $\Delta_{47}$  without uncertainties related to the reaction temperature.

For these reasons, we recommend that in the future carbonate clumped isotope data should be normalized to the three-carbonate reference materials ETH-1, ETH-2 and ETH-3 using their nominal values reported in table 1. These nominal values have been determined at 90°C and at 70 °C and were projected to 25°C using correction factors of 0.088 ‰ and 0.066 ‰ for 90 °C and 70 °C reactions, respectively (Petersen et al. 2019). To clearly distinguish these data normalization scheme from previous ones we propose the denomination I-CDES (InterCarb - Carbon Dioxide Equilibration Scale) to indicate that the data were normalized with the scheme based on the three carbonates described in this paper. This procedure is



analogous to the change from the PDB to the VPDB scale, which was accomplished by defining a consensus offset of +1.95 ‰ between the original PDB reference material and the NBS19 carbonate. This was subsequently improved by defining a second anchor point with the L-SVEC Lithium carbonate standard (Coplen et al., 2006). We note that because the carbonate values are firmly anchored to the CDES scale via HG/EG measurements in multiple laboratories, the two scales are equivalent. However, I-CDES has the advantage that it follows the principle of equal treatment of sample and standards and thus is preferred because it removes inter-laboratory uncertainties and is based on traceable solid standards.

In line with the tradition we keep a projection to 25 °C acid reaction temperature. However, we emphasize that previously published  $\Delta_{47}$  data, either normalized with heated and equilibrated gases or with carbonates, cannot be directly compared to the new ones. All previous carbonate-normalized data have to be recalculated in order to fully take into account the new values of the ETH standards reported here. For HG/EG normalized data it may not be possible to establish the compatibility to the new scale, unless an adequate number of ETH standards had been measured during the same period of time in the same laboratory. In this case, as suggested by Petersen et al. (2019), it may be more reasonable to directly compare reconstructed temperatures with a laboratory-specific calibration, rather than  $\Delta_{47}$  values, although the uncertainties are difficult to estimate for such comparisons.

### 3.5 Converting $\Delta_{47}$ values from the previous carbonate-based values to the I-CDES

Redefining the nominal  $\Delta_{47}$  values for a set of three carbonate reference materials is equivalent to defining a new reference frame: there is a one-to-one relationship between the  $\Delta_{47}$  values of unknown samples standardized using the old nominal values and the new ones. In this section we summarize how to directly convert  $\Delta_{47}$  from an arbitrary “old” set of standard values (e.g., Bernasconi et al. 2018) ( $^{\text{old}}\Delta_{47}$ ) to a “new” one ( $^{\text{new}}\Delta_{47}$ , e.g., expressed in the I-CDES) without fully reprocessing the original raw data, by computing this one-to-one relationship explicitly.

Considering that  $^{\text{old}}\Delta_{47}$  and  $^{\text{new}}\Delta_{47}$  are both derived from the same underlying raw data ( $\delta_{47}$ ,  $^{\text{raw}}\Delta_{47}$ ) by way of affine functions of the form: standardization):

891

$$\Delta_{47}^{\text{new}} \approx 0.048529 + 0.944081 \times \Delta_{47}^{\text{old}} \quad (\text{eq. Z})$$

893

894 For instance, to convert the  $\Delta_{47}^{\text{old}}$  value of ETH-4 reported by Bernasconi et al. (2018) to the  
 895 I-CDES defined in this study, we only need to know that  $\delta_{47}(\text{ETH-4}) = -28.8 \text{ ‰}$  and  
 896  $\Delta_{47}^{\text{old}}(\text{ETH-4}) = 0.507 \pm 0.004 \text{ ‰}$ . The  $\Delta_{47}^{\text{new}}$  value predicted by eq. Y is then  $0.5318 \pm$   
 897  $0.004 \text{ ‰}$ , to be compared with the independently constrained values reported here in Table 1  
 898 ( $0.5369 \pm 0.0020 \text{ ‰}$ ) and Fig. 5 ( $0.5376 \pm 0.0011 \text{ ‰}$ ). This approximation can also be  
 899 described as a simple two-point correction between old and new values, constrained by ETH-  
 900 1 and ETH-3 (or, alternatively, by ETH-3 and the average of ETH-1 and ETH-2). This  
 901 assumption is valid as long as the PBL correction has been done appropriately and has a very  
 902 low residual slope.

903

## 904 5. Outlook for future improvements

905

906 InterCarb has shown that with carbonate standardization it is possible to reach an inter-  
 907 laboratory reproducibility that is similar to what can be reached within a single laboratory.  
 908 The use of carbonate standards allows for a better monitoring of the performance of the entire  
 909 preparation system, including acid reaction and sample purification. Regular and systematic  
 910 measurement of carbonate reference material distributed within a run or measurement  
 911 interval is a prerequisite to correct for short- and long-term variations in the analytical  
 912 systems. In this study, analyses were grouped in measurement intervals, and all data were  
 913 processed assuming no short-term variation in the instrumentation. However, especially with  
 914 "small sample approaches" relying on short (~30 min) measurements of many replicates, one  
 915 can observe short-term variations (Bernasconi et al. 2018, Fig. 4). Thus a moving window  
 916 correction with a variable window size may be desirable.

917

918 Replicate sample measurements should be spread in time as much as possible. With a moving  
 919 window approach, this even allows the incorporation of standardization errors into observed  
 920 reproducibility, if corrections of replicate measurements are based on independent standard  
 921 measurements. The number and distribution of standards in a run are also important  
 922 parameters that can improve reproducibility and reduce errors. This has been discussed in  
 923 detail by (Kocken et al., 2019) who concluded that carbonate standards with bulk and

clumped-isotope compositions similar to those of unknowns should be analyzed with greater frequency than the other anchors, while preserving a minimal level of replication for each anchor.

For InterCarb all aliquots of IAEA C1, IAEA C2 and MERCK that were distributed stemmed from a single bottle. The results demonstrate that these carbonates were homogeneous at a 100  $\mu\text{g}$  level within this aliquot, but we strongly recommend verifying that additional aliquots purchased from IAEA and MERCK are identical to the ones tested here. Finally, as there are postulated differences in phosphoric acid fractionation for different mineralogies, there should be a community effort to identify and characterize reference materials for other carbonates such as aragonite and siderite, similarly to the dolomite reference materials proposed by (Müller et al., 2019).

## 6. Conclusions

- This study has shown that carbonate-based standardization is robust in spite of the smaller range of isotopic composition, compared to heated/equilibrated gas compositions.
- The smaller range of clumped and bulk isotope compositions of anchor samples compared to HG-EG standardization scheme requires a sufficient number of replicate measurements of standards in order to minimize error, particularly when extrapolating to “exotic” compositions, in ( $\delta^{47}$ ,  $\Delta_{47}$ ) space.
- Standardization with ETH-1, ETH-2 and ETH-3 provides a robust framework for converting carbonate clumped isotope data to the CDES. Data standardized with this method should be reported as I-CDES.
- We propose new community accepted values for the 4 ETH standards, two samples distributed by the IAEA: C1 and C2 and for MERCK.
- We note that only the 4 ETH standards are fully tested for homogeneity over multiple batches. When a laboratory purchases samples of IAEA-C1, C2 and MERK we recommend checking the composition against the aliquots distributed through InterCarb.
- Laboratories can now carefully use the seven carbonate standards studied here to

move into the Carbon Dioxide Equilibration Scale (CDES), ensuring robust inter-lab comparability within a consistent reference frame, we term this approach the I-CDES.

- Carbonate standardization removes uncertainties due to poorly known acid fractionation factors and different preparation systems and thus reduces differences between laboratories.
- It is recommended to switch to carbonate standardization in order to follow the principle of equal treatment between samples and standards and improve inter-laboratory data comparability.
- IAEA C2 can be used as a substitute for ETH-3 to correct for scale compression, or as independent reference material to monitor the long-term reproducibility of an instrument and the correction procedures.
- The use of a reference material with an extreme bulk isotope composition such as MERCK is recommended, as the large range of compositions reduces standardization errors for samples with “exotic” bulk compositions (as previously suggested by some authors but precisely quantified in Daëron et al., this volume).
- Error should include both replication error and reference frame error in proper statistical data treatment and reporting of uncertainties.

## References

- Affek, H. P., & Eiler, J. M. (2006). Abundance of mass 47 CO<sub>2</sub> in urban air, car exhaust, and human breath. *Geochimica et Cosmochimica Acta*, 70(1), 1–12.  
<https://doi.org/10.1016/j.gca.2005.08.021>
- Bernasconi, S. M., Hu, B., Wacker, U., Fiebig, J., Breitenbach, S. F. M., & Rutz, T. (2013). Background effects on Faraday collectors in gas-source mass spectrometry and implications for clumped isotope measurements. *Rapid Communications in Mass Spectrometry*, 27(5), 603–612. <https://doi.org/10.1002/rcm.6490>
- Bernasconi, S. M., Müller, I. A., Bergmann, K. D., Breitenbach, S. F. M., Fernandez, A., Hodell, D. A., et al. (2018). Reducing uncertainties in carbonate clumped isotope analysis through consistent carbonate-based standardization. *Geochemistry, Geophysics, Geosystems*, 19(9), 1–20. <https://doi.org/10.1029/2017GC007385>
- Bonifacie, M., Calmels, D., Eiler, J. M., Horita, J., Chaduteau, C., Vasconcelos, C., et al.

- (2017). 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. *Geochimica et Cosmochimica Acta*, 200, 255–279. <https://doi.org/10.1016/j.gca.2016.11.028>
- Brand, W. a., Assonov, S. S., & Coplen, T. B. (2010). Correction for the 17O interference in δ(13C) measurements when analyzing CO<sub>2</sub> with stable isotope mass spectrometry (IUPAC Technical Report). *Pure and Applied Chemistry*, 82(8), 1719–1733. <https://doi.org/10.1351/PAC-REP-09-01-05>
- Came, R. E., Brand, U., & Affek, H. P. (2014). Clumped isotope signatures in modern brachiopod carbonate. *Chemical Geology*, 377, 20–30. <https://doi.org/10.1016/j.chemgeo.2014.04.004>
- Carter, J. F., & Fry, B. (2013). Ensuring the reliability of stable isotope ratio data - Beyond the principle of identical treatment. *Analytical and Bioanalytical Chemistry*, 405(9), 2799–2814. <https://doi.org/10.1007/s00216-012-6551-0>
- Daëron, M., Blamart, D., Peral, M., & Affek, H. P. P. (2016a). Absolute isotopic abundance ratios and the accuracy of  $\delta^{47}$  measurements. *Chemical Geology*, 442, 83–96. <https://doi.org/10.1016/j.chemgeo.2016.08.014>
- Daëron, M., Blamart, D., Peral, M., & Affek, H. P. (2016b). Absolute isotopic abundance ratios and the accuracy of  $\Delta^{47}$  measurements. *Chemical Geology*, 442, 83–96. <https://doi.org/10.1016/j.chemgeo.2016.08.014>
- Defliese, W. F., & Lohmann, K. C. (2015). Non-linear mixing effects on mass-47 CO<sub>2</sub> clumped isotope thermometry: Patterns and implications. *Rapid Communications in Mass Spectrometry*, 29(9), 901–909. <https://doi.org/10.1002/rcm.7175>
- Dennis, K. J., Affek, H. P., Passey, B. H., Schrag, D. P., & Eiler, J. M. (2011). Defining an absolute reference frame for “clumped” isotope studies of CO<sub>2</sub>. *Geochimica et Cosmochimica Acta*, 75(22), 7117–7131. <https://doi.org/10.1016/j.gca.2011.09.025>
- Eagle, R. a, Schauble, E. a, Tripathi, A. K., Tütken, T., Hulbert, R. C., & Eiler, J. M. (2010). Body temperatures of modern and extinct vertebrates from (13)C-(18)O bond abundances in bioapatite. *Proceedings of the National Academy of Sciences of the United States of America*, 107(23), 10377–10382. <https://doi.org/10.1073/pnas.0911115107>
- Fernandez, A., Tang, J., & Rosenheim, B. E. (2014). Siderite ‘clumped’ isotope thermometry:

- 1023 A new paleoclimate proxy for humid continental environments. *Geochimica et*  
 1024 *Cosmochimica Acta*, 126, 411–421. <https://doi.org/10.1016/j.gca.2013.11.006>
- 1025 Fernandez, A., Müller, I. A., Rodríguez-Sanz, L., van Dijk, J., Looser, N., & Bernasconi, S.  
 1026 M. (2017). A Reassessment of the Precision of Carbonate Clumped Isotope  
 1027 Measurements: Implications for Calibrations and Paleoclimate Reconstructions.  
 1028 *Geochemistry, Geophysics, Geosystems*, 18(12). <https://doi.org/10.1002/2017GC007106>
- 1029 Ferry, J. M., Passey, B. H., Vasconcelos, C., & Eiler, J. M. (2011). Formation of dolomite at  
 1030 40–80 °C in the Latemar carbonate buildup, Dolomites, Italy, from clumped isotope  
 1031 thermometry. *Geology*, 39(6), 571–574. <https://doi.org/10.1130/G31845.1>
- 1032 Fiebig, J., Bajnai, D., Löffler, N., Methner, K., Krsnik, E., Mulch, A., & Hofmann, S. (2019).  
 1033 Combined high-precision  $\Delta 48$  and  $\Delta 47$  analysis of carbonates. *Chemical Geology*,  
 1034 522(February), 186–191. <https://doi.org/10.1016/j.chemgeo.2019.05.019>
- 1035 Ghosh, P., Adkins, J., Affek, H., Balta, B., Guo, W., Schauble, E. A., et al. (2006). 13C–18O  
 1036 bonds in carbonate minerals: A new kind of paleothermometer. *Geochimica et*  
 1037 *Cosmochimica Acta*, 70(6), 1439–1456. <https://doi.org/10.1016/j.gca.2005.11.014>
- 1038 Ghosh, P., Garzzone, C., & Eiler, J. (2006). Rapid Uplift of the Altiplano Revealed Through  
 1039 13C–18O Bonds in Paleosol Carbonates. *Science*, 311(January), 511–515.  
 1040 <https://doi.org/10.1126/science.1119365>
- 1041 Grauel, A. L., Schmid, T. W., Hu, B., Bergami, C., Capotondi, L., Zhou, L., & Bernasconi, S.  
 1042 M. (2013). Calibration and application of the “clumped isotope” thermometer to  
 1043 foraminifera for high-resolution climate reconstructions. *Geochimica et Cosmochimica*  
 1044 *Acta*, 108, 125–140. <https://doi.org/10.1016/j.gca.2012.12.049>
- 1045 Guo, W., & Eiler, J. M. (2007). Temperatures of aqueous alteration and evidence for methane  
 1046 generation on the parent bodies of the CM chondrites. *Geochimica et Cosmochimica*  
 1047 *Acta*, 71(22), 5565–5575. <https://doi.org/10.1016/j.gca.2007.07.029>
- 1048 He, B., Olack, G. A., & Colman, A. S. (2012). Pressure baseline correction and high-  
 1049 precision CO<sub>2</sub> clumped-isotope ( $\Delta 47$ ) measurements in  
 1050 bellows and micro-volume modes. *Rapid Communications in Mass Spectrometry*,  
 1051 26(24), 2837–2853. <https://doi.org/10.1002/rcm.6436>
- 1052 Hu, B., Radke, J., Schlüter, H. J., Heine, F. T., Zhou, L., Bernasconi, S. M., et al. (2014). A  
 1053 modified procedure for gas-source isotope ratio mass spectrometry: The long-integration  
 1054 dual-inlet (LIDI) methodology and implications for clumped isotope measurements.  
 1055 *Rapid Communications in Mass Spectrometry*, 28(13), 1413–1425.

- 1056 <https://doi.org/10.1002/rcm.6909>
- 1057 Huntington, K. W., Eiler, J. M., Affek, H. P., Guo, W., Bonifacie, M., Yeung, L. Y., et al.  
 1058 (2009). Methods and limitations of “clumped” CO<sub>2</sub> isotope ( $\Delta_{47}$ ) analysis by gas-  
 1059 source isotope ratiomass spectrometry. *Journal of Mass Spectrometry*, 44(9), 1318–  
 1060 1329. <https://doi.org/10.1002/jms.1614>
- 1061 Huntington, K. W., Budd, D. A., Wernicke, B. P., & Eiler, J. M. (2011). Use of Clumped-  
 1062 Isotope Thermometry To Constrain the Crystallization Temperature of Diagenetic  
 1063 Calcite. *Journal of Sedimentary Research*, 81(9), 656–669.  
 1064 <https://doi.org/10.2110/jsr.2011.51>
- 1065 Jautzy, J. J., Savard, M. ., Dhillon, R. ., Bernasconi, S. ., Lavoie, D., & Smirnov, A. (2021).  
 1066 Clumped isotope temperature calibration for calcite: Bridging theory and  
 1067 experimentation. *Geochemical Perspective Letters*, (2020), 36–41.  
 1068 <https://doi.org/10.7185/geochemlet.2021>
- 1069 Katz, A., Bonifacie M., Hermoso M., Cartigny P., Calmels D. (2017) Laboratory-grown  
 1070 coccoliths exhibit no vital effect in clumped isotope ( $\Delta_{47}$ ) composition on a range of  
 1071 geologically relevant temperatures in *GCA*, 208, 335–353. DOI :10.1016/j.gca.2017.02.025
- 1072 Kelson, J. R., Huntington, K. W., Schauer, A. J., Saenger, C., & Lechler, A. R. (2017).  
 1073 Toward a universal carbonate clumped isotope calibration: Diverse synthesis and  
 1074 preparatory methods suggest a single temperature relationship. *Geochimica et*  
 1075 *Cosmochimica Acta*, 197, 104–131. <https://doi.org/10.1016/j.gca.2016.10.010>
- 1076 Kim, S. T., Coplen, T. B., & Horita, J. (2015). Normalization of stable isotope data for  
 1077 carbonate minerals: Implementation of IUPAC guidelines. *Geochimica et*  
 1078 *Cosmochimica Acta*, 158, 276–289. <https://doi.org/10.1016/j.gca.2015.02.011>
- 1079 Kocken, I. J., Müller, I. A., & Ziegler, M. (2019). Optimizing the Use of Carbonate Standards  
 1080 to Minimize Uncertainties in Clumped Isotope Data. *Geochemistry, Geophysics,*  
 1081 *Geosystems*, 20(11), 5565–5577. <https://doi.org/10.1029/2019GC008545>
- 1082 Massey, F. J. (1951). The Kolmogorov-Smirnov Test for Goodness of Fit. *Journal of the*  
 1083 *American Statistical Association*, 46(253), 68–78.  
 1084 <https://doi.org/10.1080/01621459.1951.10500769>
- 1085 Meckler, A. N., Ziegler, M., Millán, M. I., Breitenbach, S. F. M., Bernasconi, S. M., Millán,  
 1086 M. I., et al. (2014). Long-term performance of the Kiel carbonate device with a new  
 1087 correction scheme for clumped isotope measurements. *Rapid Communications in Mass*  
 1088 *Spectrometry*, 28(15), 1705–1715. <https://doi.org/10.1002/rcm.6949>

- 1089 Meier-Augenstein, W., & Schimmelmann, A. (2019). A guide for proper utilisation of stable  
1090 isotope reference materials \*. *Isotopes in Environmental and Health Studies*, 55(2), 113–  
1091 128. <https://doi.org/10.1080/10256016.2018.1538137>
- 1092 Meinicke, N., Ho, S. L., Hannisdal, B., Nürnberg, D., Tripathi, A., Schiebel, R., & Meckler, A.  
1093 N. (2020). A robust calibration of the clumped isotopes to temperature relationship for  
1094 foraminifers. *Geochimica et Cosmochimica Acta*, 270, 160–183. [https://doi.org/10.1016/](https://doi.org/10.1016/j.gca.2019.11.022)  
1095 [j.gca.2019.11.022](https://doi.org/10.1016/j.gca.2019.11.022)
- 1096 Müller, I. A., Fernandez, A., Radke, J., Dijk, J. Van, Bowen, D., Schwieters, J., et al. (2017).  
1097 Carbonate clumped isotope analyses with the long-integration dual-inlet (LIDI)  
1098 workflow: scratching at the lower sample weight boundaries. *Rapid Communications in*  
1099 *Mass Spectrometry*, 31(12), 1057–1066. <https://doi.org/10.1002/rcm.7878>
- 1100 Müller, I. A., Violay, M. E. S., Storck, J.-C. C., Fernandez, A., van Dijk, J., Madonna, C., &  
1101 Bernasconi, S. M. (2017). Clumped isotope fractionation during phosphoric acid  
1102 digestion of carbonates at 70 °C. *Chemical Geology*, 449, 1–14.  
1103 <https://doi.org/10.1016/j.chemgeo.2016.11.030>
- 1104 Müller, I. A., Rodriguez-Blanco, J. D., Storck, J.-C., do Nascimento, G. S., Bontognali, T. R.  
1105 R., Vasconcelos, C., et al. (2019). Calibration of the oxygen and clumped isotope  
1106 thermometers for (proto-)dolomite based on synthetic and natural carbonates. *Chemical*  
1107 *Geology*, 525(December 2018), 1–17. <https://doi.org/10.1016/j.chemgeo.2019.07.014>
- 1108 Passey, B. H., & Henkes, G. A. (2012). Carbonate clumped isotope bond reordering and  
1109 geospeedometry. *Earth and Planetary Science Letters*, 351–352, 223–236.  
1110 <https://doi.org/10.1016/j.epsl.2012.07.021>
- 1111 Passey, B. H., Levin, N. E., Cerling, T. E., BROWN, F. H., Eiler, J. M., & Turekian, K. K.  
1112 (2010). High-temperature environments of human evolution in East Africa based on  
1113 bond ordering in paleosol carbonates. *Proceedings of the National Academy of Sciences*  
1114 *of the United States of America*, 107(25), 11245–11249.  
1115 <https://doi.org/10.1073/pnas.1001824107>
- 1116 Peral, M., Daëron, M., Blamart, D., Bassinot, F., Dewilde, F., Smialkowski, N., et al. (2018).  
1117 Updated calibration of the clumped isotope thermometer in planktonic and benthic  
1118 foraminifera. *Geochimica et Cosmochimica Acta*, 239, 1–16.  
1119 <https://doi.org/10.1016/j.gca.2018.07.016>
- 1120 Petersen, S. V., Defliese, W. F. F., Saenger, C., Daëron, M., Huntington, K. W. W., John, C.  
1121 M. M., et al. (2019). Effects of Improved <sup>17</sup>O Correction on Inter-laboratory



- 1122 Agreement in Clumped Isotope Calibrations, Estimates of Mineral-Specific Offsets, and  
 1123 Temperature Dependence of Acid Digestion Fractionation. *Geochemistry, Geophysics,*  
 1124 *Geosystems*, 20, 3495–3519. <https://doi.org/10.1029/2018GC008127>
- 1125 Petersen, Sierra V., Winkelstern, I. Z., Lohmann, K. C., & Meyer, K. W. (2016). The effects  
 1126 of Porapak™ trap temperature on  $\delta$  18 O,  $\delta$  13 C, and  $\Delta$  47 values in preparing samples  
 1127 for clumped isotope analysis. *Rapid Communications in Mass Spectrometry*, 30(1), 199–  
 1128 208. <https://doi.org/10.1002/rcm.7438>
- 1129 Piasecki, A., Bernasconi, S. M., Grauel, A., Hannisdal, B., Ho, S. L., Leutert, T. J., et al.  
 1130 (2019). Application of Clumped Isotope Thermometry to Benthic Foraminifera.  
 1131 *Geochemistry, Geophysics, Geosystems*, 2018GC007961.  
 1132 <https://doi.org/10.1029/2018GC007961>
- 1133 Schauer, A. J., Kelson, J., Saenger, C., & Huntington, K. W. (2016). Choice of 17O  
 1134 correction affects clumped isotope ( $\Delta$ 47) values of CO<sub>2</sub> measured with mass  
 1135 spectrometry. *Rapid Communications in Mass Spectrometry*, 30(24), 2607–2616. <https://doi.org/10.1002/rcm.7743>
- 1137 Schmid, T. W., & Bernasconi, S. M. (2010). An automated method for “clumped-isotope”  
 1138 measurements on small carbonate samples. *Rapid Communications in Mass*  
 1139 *Spectrometry*, 24(14), 1955–1963. <https://doi.org/10.1002/rcm.4598>
- 1140 Spooner, P. T., Guo, W., Robinson, L. F., Thiagarajan, N., Hendry, K. R., Rosenheim, B. E.,  
 1141 & Leng, M. J. (2016). Clumped isotope composition of cold-water corals: A role for  
 1142 vital effects? *Geochimica et Cosmochimica Acta*, 179, 123–141. <https://doi.org/10.1016/j.gca.2016.01.023>
- 1144 Swart, P. K., Murray, S. T., Staudigel, P. T., & Hodell, D. A. (2019). Oxygen Isotopic  
 1145 Exchange between CO<sub>2</sub> and Phosphoric Acid: Implications for the Measurement of  
 1146 Clumped Isotopes in Carbonates. *Geochemistry, Geophysics, Geosystems*, (1950),  
 1147 2019GC008209. <https://doi.org/10.1029/2019GC008209>
- 1148 Thaler C., Katz A., Bonifacie M., Ménez B., Ader M. Oxygen isotope composition of waters  
 1149 recorded in carbonates in strong clumped and oxygen isotopic disequilibrium (2020).  
 1150 *Biogeosciences*, 17, 1–14, 2020. DOI 10.5194/bg-17-1-2020
- 1151 Wacker, U., Fiebig, J., Tödter, J., Schöne, B. R., Bahr, A., Friedrich, O., et al. (2014).  
 1152 Empirical calibration of the clumped isotope paleothermometer using calcites of various  
 1153 origins. *Geochimica et Cosmochimica Acta*, 141, 127–144.  
 1154 <https://doi.org/10.1016/j.gca.2014.06.004>

- 1155 Wang, Z., Schauble, E. A., & Eiler, J. M. (2004). Equilibrium thermodynamics of multiply  
1156 substituted isotopologues of molecular gases. *Geochimica et Cosmochimica Acta*,  
1157 68(23), 4779–4797. <https://doi.org/10.1016/j.gca.2004.05.039>
- 1158 Werner, R. A., & Brand, W. A. (2001a). Referencing strategies and techniques in stable  
1159 isotope ratio analysis. *Rapid Communications in Mass Spectrometry*, 15(7), 501–519.  
1160 <https://doi.org/10.1002/rcm.258>
- 1161 Werner, R. A., & Brand, W. A. (2001b). Referencing strategies and techniques in stable  
1162 isotope ratio analysis. *Rapid Communications in Mass Spectrometry*, 15(7), 501–519.  
1163 <https://doi.org/10.1002/rcm.258>
- 1164
- 1165

## 1166 Supplementary information

## 1167 Table S1 analytical session details for all mass spectrometers/Laboratories

| Lab | Session | Number of analytes |     |     |     |    |    |    | Working gas    |                                   | Scaling parameters                |                           |                            | Reproducibility (ppm) |                                   |                                   |                 |
|-----|---------|--------------------|-----|-----|-----|----|----|----|----------------|-----------------------------------|-----------------------------------|---------------------------|----------------------------|-----------------------|-----------------------------------|-----------------------------------|-----------------|
|     |         | E1                 | E2  | E3  | E4  | C1 | C2 | M  | N <sub>t</sub> | δ <sup>13</sup> C <sub>VYON</sub> | δ <sup>18</sup> O <sub>KNOV</sub> | a                         | b                          | c                     | δ <sup>13</sup> C <sub>VYON</sub> | δ <sup>18</sup> O <sub>VYON</sub> | Δ <sub>47</sub> |
| 01  | 01      | 16                 | 17  | 10  | 7   | 0  | 0  | 0  | 46             | -3.58                             | 25.38                             | 0.91                      | (6.0 × 10 <sup>-4</sup> )  | -0.893                | 41                                | 91                                | 31.5            |
|     | 02      | 6                  | 5   | 3   | 1   | 0  | 0  | 0  | 11             | -3.52                             | 25.58                             | 0.89                      | (-2.1 × 10 <sup>-2</sup> ) | -0.765                | 34                                | 64                                | 22.8            |
|     | 03      | 150                | 146 | 65  | 72  | 19 | 21 | 22 | 488            | -3.63                             | 25.22                             | 0.98                      | (-2.9 × 10 <sup>-3</sup> ) | -0.965                | 33                                | 74                                | 33.5            |
| 02  | 01      | 19                 | 24  | 20  | 18  | 4  | 5  | 4  | 87             | -36.89                            | 8.76                              | 0.99                      | (-5.6 × 10 <sup>-4</sup> ) | -0.955                | 17                                | 92                                | 13.0            |
|     | 02      | 6                  | 8   | 5   | 4   | 2  | 3  | 2  | 23             | -36.88                            | 8.83                              | 0.98                      | (-5.5 × 10 <sup>-4</sup> ) | -0.931                | 25                                | 77                                | 16.1            |
| 03  | 01      | 37                 | 24  | 17  | 9   | 0  | 0  | 0  | 83             | -10.44                            | 31.64                             | 0.98                      | (-1.6 × 10 <sup>-4</sup> ) | -0.917                | 22                                | 56                                | 27.9            |
|     | 02      | 29                 | 32  | 12  | 14  | 17 | 13 | 11 | 121            | -3.65                             | 25.28                             | 1.00                      | (-1.7 × 10 <sup>-4</sup> ) | -0.917                | 46                                | 93                                | 25.2            |
| 04  | 01      | 6                  | 9   | 7   | 9   | 6  | 4  | 0  | 35             | -6.57                             | 27.18                             | 0.97                      | 5.0 × 10 <sup>-3</sup>     | -1.022                | 259                               | 562                               | 40.6            |
| 05  | 01      | 3                  | 3   | 5   | 2   | 3  | 2  | 2  | 13             | -10.43                            | 31.31                             | 0.95                      | 1.7 × 10 <sup>-3</sup>     | -0.970                | 15                                | 27                                | 8.6             |
|     | 02      | 13                 | 13  | 13  | 12  | 10 | 11 | 8  | 73             | -3.62                             | 25.05                             | 0.99                      | (3.8 × 10 <sup>-4</sup> )  | -0.968                | 15                                | 24                                | 20.9            |
|     | 03      | 7                  | 10  | 10  | 8   | 5  | 4  | 4  | 41             | -3.63                             | 25.06                             | 0.90                      | 1.1 × 10 <sup>-3</sup>     | -0.901                | 42                                | 113                               | 17.3            |
| 06  | 01      | 6                  | 3   | 5   | 3   | 3  | 3  | 3  | 19             | -2.95                             | 25.52                             | 0.83                      | (-3.8 × 10 <sup>-4</sup> ) | -0.920                | 22                                | 25                                | 21.0            |
|     | 02      | 6                  | 6   | 6   | 6   | 0  | 0  | 0  | 20             | -2.98                             | 24.93                             | 0.92                      | (-9.9 × 10 <sup>-4</sup> ) | -0.920                | 14                                | 71                                | 13.3            |
|     | 03      | 3                  | 3   | 3   | 3   | 3  | 3  | 3  | 14             | -3.01                             | 24.90                             | 0.88                      | (3.6 × 10 <sup>-4</sup> )  | -0.932                | 10                                | 43                                | 9.4             |
|     | 04      | 6                  | 6   | 6   | 6   | 0  | 6  | 6  | 30             | -2.95                             | 25.28                             | 0.90                      | (-1.4 × 10 <sup>-4</sup> ) | -0.926                | 18                                | 61                                | 17.3            |
| 07  | 01      | 5                  | 4   | 4   | 2   | 4  | 4  | 3  | 19             | -11.64                            | 35.75                             | 0.87                      | 3.5 × 10 <sup>-3</sup>     | -0.836                | 91                                | 303                               | 23.9            |
| 08  | 01      | 5                  | 6   | 9   | 4   | 4  | 4  | 4  | 29             | -2.68                             | 25.86                             | 0.94                      | (-9.2 × 10 <sup>-4</sup> ) | -0.686                | 13                                | 25                                | 28.4            |
|     | 02      | 5                  | 4   | 14  | 6   | 4  | 5  | 4  | 35             | -2.64                             | 25.96                             | 0.94                      | (8.6 × 10 <sup>-4</sup> )  | -0.741                | 83                                | 88                                | 33.2            |
|     | 03      | 4                  | 4   | 13  | 4   | 3  | 5  | 6  | 32             | -2.64                             | 25.91                             | 0.93                      | (-1.7 × 10 <sup>-4</sup> ) | -0.728                | 15                                | 33                                | 33.2            |
|     | 04      | 4                  | 5   | 9   | 5   | 4  | 4  | 4  | 28             | -2.67                             | 25.85                             | 0.85                      | (1.3 × 10 <sup>-4</sup> )  | -0.629                | 17                                | 51                                | 44.5            |
|     | 05      | 3                  | 6   | 8   | 4   | 4  | 4  | 4  | 26             | -2.70                             | 25.79                             | 0.87                      | (1.3 × 10 <sup>-3</sup> )  | -0.660                | 16                                | 56                                | 43.3            |
|     | 06      | 4                  | 4   | 16  | 6   | 6  | 6  | 4  | 39             | -2.63                             | 25.90                             | 0.92                      | (3.9 × 10 <sup>-4</sup> )  | -0.693                | 85                                | 54                                | 37.8            |
|     | 07      | 3                  | 4   | 16  | 6   | 6  | 4  | 6  | 38             | -2.66                             | 25.90                             | 0.96                      | (-1.9 × 10 <sup>-3</sup> ) | -0.709                | 19                                | 52                                | 48.8            |
|     | 08      | 4                  | 4   | 16  | 4   | 4  | 4  | 4  | 33             | -2.66                             | 25.89                             | 1.03                      | (3.9 × 10 <sup>-4</sup> )  | -0.806                | 12                                | 46                                | 42.7            |
|     | 09      | 6                  | 6   | 8   | 4   | 4  | 4  | 4  | 27             | -2.67                             | 25.84                             | 0.92                      | (1.6 × 10 <sup>-4</sup> )  | -0.722                | 19                                | 25                                | 46.7            |
|     | 10      | 6                  | 6   | 6   | 4   | 4  | 2  | 4  | 25             | -2.63                             | 25.91                             | 0.97                      | (4.4 × 10 <sup>-4</sup> )  | -0.767                | 36                                | 39                                | 40.5            |
|     | 11      | 6                  | 5   | 8   | 4   | 4  | 3  | 4  | 27             | -2.67                             | 25.87                             | 0.97                      | (2.5 × 10 <sup>-4</sup> )  | -0.760                | 11                                | 31                                | 49.5            |
|     | 12      | 6                  | 6   | 8   | 3   | 4  | 4  | 4  | 28             | -2.66                             | 25.86                             | 1.02                      | (7.9 × 10 <sup>-4</sup> )  | -0.767                | 58                                | 40                                | 61.3            |
|     | 13      | 4                  | 6   | 8   | 6   | 4  | 4  | 6  | 31             | -2.63                             | 25.93                             | 0.89                      | (1.3 × 10 <sup>-3</sup> )  | -0.685                | 19                                | 38                                | 41.0            |
|     | 14      | 6                  | 7   | 5   | 4   | 4  | 4  | 4  | 26             | -2.59                             | 25.90                             | 0.90                      | (-3.6 × 10 <sup>-4</sup> ) | -0.665                | 76                                | 104                               | 27.4            |
|     | 15      | 6                  | 4   | 8   | 4   | 4  | 4  | 4  | 27             | -2.68                             | 25.79                             | 0.95                      | (-2.0 × 10 <sup>-3</sup> ) | -0.685                | 21                                | 52                                | 36.0            |
|     | 16      | 2                  | 2   | 10  | 5   | 4  | 2  | 4  | 22             | -2.63                             | 25.89                             | 0.96                      | (-5.4 × 10 <sup>-4</sup> ) | -0.765                | 40                                | 39                                | 38.8            |
| 09  | 01      | 4                  | 4   | 5   | 6   | 0  | 0  | 0  | 15             | -3.60                             | 25.36                             | 0.89                      | 3.8 × 10 <sup>-3</sup>     | -0.856                | 22                                | 74                                | 28.3            |
|     | 02      | 26                 | 19  | 16  | 24  | 0  | 0  | 0  | 81             | -3.36                             | 19.94                             | 0.90                      | 5.2 × 10 <sup>-3</sup>     | -0.928                | 46                                | 98                                | 18.4            |
|     | 03      | 21                 | 17  | 13  | 19  | 0  | 1  | 0  | 66             | -3.53                             | 24.49                             | 0.92                      | (-1.0 × 10 <sup>-2</sup> ) | -0.968                | 72                                | 1667                              | 22.4            |
|     | 04      | 19                 | 16  | 13  | 16  | 8  | 7  | 2  | 74             | -3.60                             | 25.27                             | 0.98                      | (-9.6 × 10 <sup>-3</sup> ) | -0.994                | 44                                | 56                                | 16.0            |
| 10  | 01      | 7                  | 7   | 8   | 2   | 0  | 11 | 0  | 30             | -7.43                             | 32.38                             | 0.98                      | 1.9 × 10 <sup>-3</sup>     | -1.077                | 24                                | 38                                | 35.1            |
|     | 02      | 15                 | 15  | 21  | 15  | 11 | 20 | 11 | 101            | -7.41                             | 32.42                             | 0.93                      | (-2.0 × 10 <sup>-4</sup> ) | -0.877                | 25                                | 44                                | 23.0            |
|     | 03      | 17                 | 18  | 25  | 9   | 22 | 31 | 20 | 135            | -7.43                             | 32.37                             | 0.96                      | (-2.8 × 10 <sup>-4</sup> ) | -0.900                | 31                                | 92                                | 30.0            |
| 11  | 01      | 24                 | 24  | 28  | 28  | 0  | 0  | 0  | 100            | -3.63                             | 25.37                             | 0.99                      | (-8.1 × 10 <sup>-5</sup> ) | -0.974                | 23                                | 91                                | 19.1            |
|     | 02      | 20                 | 18  | 15  | 15  | 0  | 0  | 0  | 64             | -3.60                             | 25.53                             | 0.98                      | (3.5 × 10 <sup>-4</sup> )  | -0.996                | 35                                | 270                               | 28.9            |
|     | 03      | 69                 | 62  | 74  | 66  | 13 | 13 | 8  | 298            | -3.02                             | 24.99                             | 0.91                      | (-2.2 × 10 <sup>-4</sup> ) | -1.065                | 34                                | 89                                | 25.0            |
|     | 04      | 36                 | 34  | 34  | 35  | 6  | 4  | 8  | 150            | -3.01                             | 25.06                             | 1.00                      | (-3.1 × 10 <sup>-4</sup> ) | -1.068                | 76                                | 210                               | 33.7            |
|     | 05      | 90                 | 83  | 92  | 78  | 12 | 10 | 9  | 367            | -2.76                             | 25.78                             | 0.98                      | (-5.0 × 10 <sup>-5</sup> ) | -1.088                | 97                                | 317                               | 19.3            |
| 12  | 01      | 7                  | 7   | 9   | 5   | 5  | 6  | 5  | 37             | -3.75                             | 25.15                             | 0.89                      | 3.7 × 10 <sup>-3</sup>     | -0.904                | 7                                 | 41                                | 10.2            |
|     | 02      | 7                  | 6   | 6   | 6   | 5  | 5  | 33 | -3.74          | 25.18                             | 0.87                              | 4.6 × 10 <sup>-3</sup>    | -0.897                     | 8                     | 50                                | 9.3                               |                 |
|     | 03      | 8                  | 7   | 12  | 5   | 5  | 5  | 40 | -3.74          | 25.17                             | 0.88                              | 5.5 × 10 <sup>-3</sup>    | -0.909                     | 9                     | 51                                | 9.7                               |                 |
|     | 04      | 6                  | 7   | 6   | 5   | 5  | 5  | 4  | 31             | -3.74                             | 25.17                             | 0.88                      | 5.3 × 10 <sup>-3</sup>     | -0.908                | 7                                 | 51                                | 8.7             |
| 13  | 01      | 58                 | 51  | 59  | 47  | 6  | 12 | 9  | 235            | -10.29                            | 33.18                             | 0.98                      | (-3.7 × 10 <sup>-4</sup> ) | -0.993                | 176                               | 239                               | 26.8            |
| 14  | 01      | 4                  | 7   | 10  | 10  | 0  | 0  | 0  | 27             | -3.63                             | 24.95                             | 0.93                      | (1.3 × 10 <sup>-4</sup> )  | -0.972                | 42                                | 159                               | 19.3            |
|     | 02      | 10                 | 11  | 8   | 7   | 0  | 0  | 0  | 32             | -3.61                             | 25.04                             | 0.97                      | (5.8 × 10 <sup>-4</sup> )  | -1.021                | 40                                | 128                               | 30.0            |
|     | 03      | 6                  | 4   | 4   | 3   | 0  | 0  | 0  | 13             | -10.38                            | 31.93                             | 0.84                      | (-1.7 × 10 <sup>-2</sup> ) | -0.747                | 39                                | 59                                | 20.5            |
|     | 04      | 2                  | 2   | 2   | 2   | 0  | 0  | 0  | 4              | -10.40                            | 31.92                             | 0.86                      | (-1.0 × 10 <sup>-3</sup> ) | -0.794                | 20                                | 29                                | 9.2             |
|     | 05      | 4                  | 4   | 3   | 4   | 0  | 0  | 0  | 11             | -10.40                            | 31.92                             | 0.91                      | (-1.6 × 10 <sup>-3</sup> ) | -0.807                | 27                                | 60                                | 11.0            |
|     | 06      | 5                  | 6   | 6   | 7   | 0  | 0  | 0  | 20             | -10.43                            | 31.84                             | 0.99                      | (1.3 × 10 <sup>-4</sup> )  | -0.908                | 39                                | 53                                | 22.4            |
|     | 07      | 3                  | 5   | 2   | 1   | 0  | 0  | 0  | 7              | -10.41                            | 31.85                             | 0.97                      | (-1.7 × 10 <sup>-4</sup> ) | -0.877                | 51                                | 43                                | 12.8            |
|     | 08      | 11                 | 7   | 3   | 5   | 0  | 0  | 0  | 22             | -10.47                            | 31.66                             | 0.94                      | (-7.8 × 10 <sup>-4</sup> ) | -0.920                | 61                                | 84                                | 23.4            |
|     | 09      | 4                  | 2   | 3   | 4   | 0  | 0  | 0  | 9              | -10.43                            | 31.82                             | 0.95                      | (-4.8 × 10 <sup>-4</sup> ) | -0.907                | 55                                | 83                                | 12.0            |
| 10  | 4       | 4                  | 1   | 3   | 0   | 0  | 0  | 8  | -10.49         | 31.73                             | 0.99                              | (1.7 × 10 <sup>-4</sup> ) | -0.926                     | 40                    | 71                                | 13.3                              |                 |
| 15  | 01      | 4                  | 4   | 4   | 4   | 0  | 0  | 0  | 12             | -32.89                            | 36.92                             | 0.96                      | (-2.5 × 10 <sup>-3</sup> ) | -0.887                | 87                                | 70                                | 14.6            |
|     | 02      | 4                  | 4   | 4   | 4   | 4  | 4  | 4  | 21             | -3.72                             | 24.98                             | 1.02                      | 4.6 × 10 <sup>-3</sup>     | -1.027                | 59                                | 41                                | 14.0            |
| 16  | 01      | 5                  | 6   | 6   | 6   | 4  | 3  | 2  | 23             | -10.49                            | 31.56                             | 0.99                      | (-4.1 × 10 <sup>-3</sup> ) | -0.979                | 47                                | 109                               | 10.1            |
| 17  | 01      | 3                  | 5   | 2   | 5   | 6  | 6  | 3  | 23             | -9.73                             | 23.81                             | 0.81                      | (6.3 × 10 <sup>-4</sup> )  | -0.940                | 65                                | 204                               | 29.3            |
| 18  | 01      | 168                | 147 | 172 | 169 | 20 | 20 | 25 | 714            | -3.45                             | 25.25                             | 0.81                      | (1.5 × 10 <sup>-4</sup> )  | -0.722                | 65                                | 110                               | 37.7            |
|     | 02      | 17                 | 14  | 17  | 13  | 4  | 4  | 4  | 66             | -3.41                             | 25.42                             | 0.83                      | (1.6 × 10 <sup>-3</sup> )  | -0.761                | 21                                | 52                                | 45.7            |
|     | 03      | 11                 | 12  | 13  | 14  | 2  | 4  | 2  | 51             | -3.52                             | 25.12                             | 0.96                      | (6.0 × 10 <sup>-4</sup> )  | -0.835                | 23                                | 45                                | 40.5            |
| 19  | 01      | 4                  | 4   | 5   | 7   | 5  | 4  | 4  | 26             | -24.48                            | 25.66                             | 0.99                      | (2.0 × 10 <sup>-4</sup> )  | -0.970                | 69                                | 193                               | 23.4            |
|     | 02      | 7                  | 8   | 10  | 7   | 0  | 0  | 0  | 28             | 5.03                              | 38.66                             | 0.99                      | (2.0 × 10 <sup>-4</sup> )  | -0.962                | 164                               | 416                               | 22.5            |
| 20  | 01      | 9                  | 6   | 6   | 6   | 0  | 0  | 0  | 23             | -3.63                             | 28.89                             | 0.93                      | (-2.1 × 10 <sup>-3</sup> ) | -0.921                | 11                                | 50                                | 14.3            |
| 21  | 01      | 3                  | 3   | 3   | 3   | 0  | 0  | 0  | 8              | -3.62                             | 25.20                             | 0.90                      | 1.0 × 10 <sup>-3</sup>     | -0.886                | 65                                | 139                               | 11.4            |
| 22  | 01      | 8                  | 8   | 8   | 7   | 0  | 0  | 7  | 33             | -3.54                             | 25.37                             | 0.98                      | 9.9 × 10 <sup>-3</sup>     | -0.951                | 155                               | 443                               | 20.5            |
| 23  | 01      | 6                  | 6   | 6   | 6   | 0  | 0  | 0  | 20             | -10.77                            | 31.02                             | 1.00                      | 4.4 × 10 <sup>-3</sup>     | -0.948                | 47                                | 91                                | 20.5            |
| 24  | 01      | 19                 | 18  | 15  | 12  | 0  | 0  | 0  | 60             | -4.40                             | 25.32                             | 0.98                      | (2.1 × 10 <sup>-4</sup> )  | -0.955                | 42                                | 107                               | 9.9             |
| 26  | 01      | 4                  | 4   | 4   | 5   | 3  | 3  | 3  | 19             | -40.04                            | 5.51                              | 0.89                      | (2.2 × 10 <sup>-4</sup> )  | -0.998                | 96                                | 145                               | 15.0            |
|     | 02      | 6                  | 7   | 6   | 3   | 3  | 3  | 3  | 24             | -40.03                            | 5.40                              | 0.92                      | (-1.1 × 10 <sup>-2</sup> ) | -1.014                | 50                                | 88                                | 8.7             |

1168