1	Geological Calibration of Spallation Production Rates
	in the CRONUS-Earth Project
2	III one Ontonob-Laron i roject
3	Brian Borchers ^a , Shasta Marrero ^b , Greg Balco ^c , Marc Caffee ^d , Brent
4	Goehring ^e , Nathaniel Lifton ^r , Kunihiko Nishiizumi ^g , Fred Phillips ⁿ , Joerg
5	Schaefer ⁱ , John Stone ^j
6 7	^a New Mexico Tech, Department of Mathematics, Socorro, NM 87801, USA. borchers@nmt.edu
8	^b Department of Earth and Environmental Science, New Mexico Tech. Currently at
9	University of Edinburgh, School of Geosciences, Edinburgh, EH8 9XP, United Kingdom.
10	Shasta.Marrero@ed.ac.uk
11	^c Berkeley Geochronology Center, 2455 Ridge Road, Berkeley CA 94709, USA.
12	gbalco@bgc.org
13	^d Department of Physics, Purdue University, West Lafayette, IN 47907, USA.
14	m caff ee@purdue.edu
15 16	^e Department of Earth, Atmospheric, and Planetary Sciences, Purdue University, West Lafavette, IN 47907 USA, baoehrin@nurdue.edu
17	^f Department of Earth. Atmospheric, and Planetary Sciences. Purdue University. West
18	Lafayette, IN 47907, USA. nlifton@purdue.edu
19	^g Space Sciences Laboratory, University of California-Berkeley, Berkeley CA 94720, USA.
20	kuni@sunspot.ssl.berkeley.edu
21	^h Department of Earth and Environmental Science, New Mexico Tech, Socorro, NM
22	$87801, \ USA. \ phillips@nmt.edu$
23	ⁱ Department of Earth and Environmental Sciences, Columbia University, Palisades, NY
24	10964, USA. schaefer@ldeo.columbia.edu
25	³ Department of Earth and Space Sciences, University of Washington, Seattle, WA 98195,
26	$USA.\ stone@ess.washington.edu$

27 Abstract

Models of the production of cosmogenic nuclides typically incorporate an adjustable production rate parameter that is scaled for variations in production with latitude and altitude. In practice, this production rate parameter is set by calibration of the model using cosmogenic nuclide data from sites with independent age constraints. In this paper, we describe a calibration procedure developed during the Cosmic-Ray Produced Nuclide Systematics on Earth (CRONUS-Earth) project and its application to an extensive data set that included both new CRONUS-Earth samples and samples from pre-

Preprint submitted to Quaternary Geochronology

January 28, 2015

viously published studies. We considered seven frameworks for elevation and latitude scaling and five commonly used cosmogenic nuclides, ³He, ¹⁰Be, ¹⁴C, ²⁶Al, and ³⁶Cl. In general, the results show that the calibrated production rates fail statistical tests of goodness-of-fit. One conclusion from the calibration results is that two newly developed scaling frameworks and the widely used Lal scaling framework provide qualitatively similar fits to the data, while neutron-monitor based scaling frameworks have much poorer fit to the data. To further test the fitted models, we computed site ages for a number of secondary sites not included in the primary calibration data set. The root-mean-square percent differences between the median computed ages for these secondary sites and independent ages range from 7.1% to 27.1%, differences that are much larger than the typical uncertainties in the site ages. The results indicate that there are substantial unresolved difficulties in modeling cosmogenic nuclide production and the calibration of production rates.

Keywords: cosmogenic nuclide, production rate, calibration, beryllium-10,
aluminum-26, carbon-14, helium-3, chlorine-36

30 1. Introduction

In modeling the production of cosmogenic nuclides by spallation reactions, 31 we consider the flux of cosmic-ray neutrons at the surface as well as the cross 32 sections of reactions that produce the nuclide of interest. In theory, we can 33 combine a model that predicts the cosmic ray flux together with measured 34 or modeled reaction cross sections to directly predict the production rates in 35 a sample at a particular location (Argento et al., 2014a,b; Kollar & Masarik, 36 1999; Masarik & Beer, 1999; Masarik & Reedy, 1995; Masarik et al., 2007). 37 The difficulty in accurately modeling the cosmic ray flux at a particular 38 location on the earth surface and the lack of precise measurements of the 39 required reaction cross sections has made it difficult to apply this approach 40 in practice. 41

The modeling of cosmogenic nuclide production has typically been simplified by using a scaling model to account for variation in production with elevation and latitude. The production rate is typically expressed in units of atoms produced per year per gram of target material. The target material is typically either quartz or a specific element. The production rate at a particular site is then determined by multiplying the scaling factor for that site by a nominal production rate which is typically chosen as the production ⁴⁹ rate at sea level and high latitude. The production rate can be integrated ⁵⁰ to obtain predicted cosmogenic nuclide concentrations for samples at a par-⁵¹ ticular site with known age. The production rate parameter is calibrated by ⁵² finding the value that best fits measured concentrations from a collection of ⁵³ sites for which independent age measurements are available. In this paper we ⁵⁴ test the calibration of production rates for five commonly used cosmogenic ⁵⁵ nuclides using seven different scaling frameworks.

⁵⁶ Cosmogenic nuclides can be produced by high-energy spallation reactions,
⁵⁷ interactions with muons, or capture of low-energy neutrons (Gosse & Phillips,
⁵⁸ 2001). Although our models incorporate production through all of these
⁵⁹ reaction pathways, we assume that production rates for production by muons
⁶⁰ and low-energy neutron capture have been separately calibrated. This paper
⁶¹ focuses on the calibration of production rates for spallation reactions only.

In some cases, a cosmogenic nuclide may be produced by spallation re-62 actions involving different elements in a single sample. For example, 36 Cl is 63 commonly produced by spallation of Fe, Ti, Ca, and K. In order to model 64 production of ${}^{36}Cl$ in a sample it is necessary to know the chemical composi-65 tion of the sample. Because several of these elements may be present in our 66 calibration samples, it may be necessary to simultaneously estimate multiple 67 production rates. For this study, we have estimated production rates for ³⁶Cl 68 only from Ca and K. We have used previously published values for the spalla-69 tion production rates of ³⁶Cl from Fe and Ti (Masarik, 2002; Stone, 2005). In 70 practice, production of ³⁶Cl from Fe and Ti is typically small in comparison 71 with production from Ca and K (i.e., production from Ti and Fe is probably 72 no more than 7% and 3.5% of that from Ca by weight, respectively, and Ti 73 and Fe concentrations in most rocks are much smaller than Ca.)

St	Lal (1991) ; Stone (2000)
Lm	Balco et al. (2008)
De	Desilets et al. (2006)
Du	Dunai (2001)
Li	Lifton et al. $(2005, 2008)$
Sf	Lifton et al. $(2014b)$
Sa	Lifton et al. $(2014b)$

Table 1: Summary of seven scaling frameworks.

⁷⁵ In this paper we consider seven scaling frameworks, summarized in Table

1. We adopt a shorthand notation introduced in Balco et al. (2008) to denote 76 the scaling frameworks. The oldest and most widely used of these scaling 77 frameworks is the model of Lal (1991). We use a version of this scaling 78 framework described in Stone (2000) that has been updated to use atmo-79 spheric pressure rather than elevation. This scaling framework is denoted by 80 "St." Balco et al. (2008) adapted the St framework further to incorporate a 81 time-dependent correction for long term changes in the magnetic field of the 82 earth. This modified time-dependent version of the St framework is denoted 83 by "Lm." We also consider the scaling frameworks of Desilets et al. (2006), 84 denoted by "De", the framework of Dunai (2001), denoted by "Du", and the 85 framework of Lifton et al. (2005, 2008), denoted by "Li." 86

In the CRONUS-Earth project, two new scaling frameworks based on the 87 model of cosmic-ray fluxes proposed in Sato & Niita (2006) and Sato et al. 88 (2008) have been developed. These new scaling frameworks are described 89 in Lifton et al. (2014b). The first of these scaling frameworks, denoted by 90 "Sf", simply integrates the Sato spectrum to produce a scaling factor that 91 depends only on the total flux of neutrons and protons at a given location. 92 The second of these scaling frameworks, denoted by "Sa", multiplies the 93 energy-dependent fluxes by the reaction cross sections to produce a nuclide-94 dependent scaling factor. 95

Note that the scaling frameworks considered in this paper are actually 96 new implementations described in Marrero et al. (2014a) and Lifton et al. 97 (2014b). These new implementations incorporate recent paleomagnetic his-98 tory reconstructions and are thus not exactly identical to the previously pub-99 lished scaling frameworks. Similarly, in this paper all elevations have been 100 reduced to atmospheric pressures using the ERA-40 reanalysis of Uppala 101 et al. (2005). It is effectively impossible to test a scaling model without ref-102 erence to a particular paleomagnetic history reconstruction and atmospheric 103 pressure model. Throughout this paper, the phrase "scaling framework" 104 refers to these scaling models together with the particular paleomagnetic 105 history reconstructions used and the ERA-40 reanalysis of atmospheric pres-106 sure (Marrero et al., 2014a; Lifton et al., 2014b). 107

¹⁰⁸ These seven scaling frameworks have been incorporated into a MATLAB ¹⁰⁹ program described in Marrero et al. (2014a). This code currently supports ¹¹⁰ five cosmogenic nuclides, namely ³He, ¹⁰Be, ¹⁴C, ²⁶Al, and ³⁶Cl. The code ¹¹¹ can be used to predict the concentration, $N_{\rm pred}$, of a cosmogenic nuclide in ¹¹² a sample given its exposure age. It can also solve for the exposure age corresponding to the measured concentration, N_{meas} , of a cosmogenic nuclide in a sample. In this paper we will use the N_{pred} function in the process of calibrating production rates for the various nuclides and scaling frameworks.

116 2. Methods

Our statistical model for the calibration of the production rates begins with the assumption we have samples from multiple sites i = 1, 2, ..., m. Here, a site refers to a collection of samples from a location that have effectively the same exposure age. There are often cases where multiple exposure events have occurred in close geographic proximity but at different points in time. In these situations we treat each exposure event as a separate "site" for purposes of the calibration exercise.

We assume that an independently determined exposure age a_i is known 124 for each site. In this paper, all ages are reported in years before 1950. The 125 independently measured site ages, a_i , are uncertain with standard deviations 126 of ϵ_i . In the calibration process we will obtain a fitted age $a_i + \delta_i$ for each 127 site i. For example, if site i has a nominal age of $a_i = 10,000$ years plus or 128 minus $\epsilon_i = 500$ years, and the fitted value is $\delta_i = 1.5$, then the fitted age 129 is 10,750 years. Since uncertainties in the independent age constraints are 130 sometimes on the order of 5% of the exposure age, while uncertainties in the 131 measured concentrations are sometimes as small as 1%, it would be extremely 132 difficult to fit production rates exactly to nominal independent ages without 133 substantial differences between the measured and predicted concentrations. 134

We also need to be able to handle saturated samples, which are samples that have reached a maximum concentration determined by the balance of production and decay at a particular site. Several of the ¹⁴C calibration sites have such samples. For these saturated samples, the actual exposure age is irrelevant. Instead, we set the exposure age, a_i , to a very large value (e.g. one million years), and remove the uncertainty in the exposure age, δ_i , from the formula.

At each site *i*, there are sample measurements $j = 1, 2, ..., n_i$. Let $n = n_1 + n_2 + ... + n_m$ be the total number of measured concentrations. Let $N_{\text{meas},i,j}$ be the measured concentration of the cosmogenic nuclide for sample measurement *j* from site *i*. Note that we may include repeated measurements of the concentration in the same physical sample. It is assumed that any errors in these repeated measurements are independent. The predicted concentration of the cosmogenic nuclide depends on properties of the samples and sites such as the erosion rate, sample thickness, and density. The properties are encoded as a vector $x_{i,j}$. These parameters are assumed to be known precisely. This assumption is difficult to justify, but since good estimates of the uncertainty in these parameters are not available and since in any case it would be impossible to simultaneously fit all of these parameters using only one measured concentration per sample, the assumption is practically necessary.

Given the entire collection of sample parameters $x_{i,j}$, site ages, $a_i + \delta_i$, and a vector of production rates P, we can predict the concentration of the cosmogenic nuclide in each sample i, j, as $N_{\text{pred},i,j}(x_{i,j}, a_i + \delta_i, P)$.

¹⁵⁹ We assume that measured concentrations $N_{\text{meas},i,j}$ are unbiased and nor-¹⁶⁰ mally distributed with standard deviations $\sigma_{i,j}$. Under these assumptions we ¹⁶¹ obtain a least squares problem

$$\min_{P,\delta} \sum_{i=1}^{m} \sum_{j=1}^{n_i} \left(\frac{N_{\text{pred}}(x_{i,j}, a_i + \delta_i, P) - N_{\text{meas}, i, j}}{\sigma_{i,j}} \right)^2 + \sum_{i=1}^{m} \left(\frac{\delta_i}{\epsilon_i} \right)^2.$$
(1)

Here P and δ are the vectors of parameters that are adjusted to minimize the 162 objective function. The least squares problem is nonlinear due to the depen-163 dence on δ . This nonlinear least squares problem is solved by the Levenberg-164 Marquardt method (Aster et al., 2012; Levenberg, 1944; Marquardt, 1963). 165 Let \hat{P} and $\hat{\delta}$ be the optimal parameters that minimize (1). Let χ^2_{obs} be the 166 value of the objective function corresponding to these optimal parameters. 167 The χ^2_{obs} value can be divided by the number of degrees of freedom to obtain 168 a reduced χ^2 value, χ^2_{ν} . 169

Once we have fit the optimal production rate \hat{P} and age adjustments $\hat{\delta}$, we 170 can use the χ^2 test of goodness-of-fit. In Equation (1) there are n+m terms. 171 The number of parameters in the vector P will be denoted by length(P). We 172 are fitting m parameters δ_i , i = 1, 2, ..., m, and length(P) production rate 173 parameters. Thus the χ^2 test is performed with n + m - m - length(P) =174 n - length(P) degrees of freedom. The result of this goodness-of-fit test 175 is a *p*-value corresponding to the probability of having a misfit as large as 176 the observed misfit if the model and its parameters were correct. Following 177 standard practice, we reject the fit whenever the p-value is smaller than 5% 178 (Aster et al., 2012). 179

For the calibrations reported in this paper, the values of a_i , ϵ_i , $x_{i,j}$, and N_{meas,*i*,*j*} come from the CRONUS-Earth project and a variety of other published papers discussed in Section 3. An important issue in the calibration process is the determination of the uncertainties in the concentration measurements, $\sigma_{i,j}$.

In practice, when researchers measure the concentration of a cosmogenic 185 nuclide in a sample, they report on the internal analytical uncertainty in the 186 concentration measurement. These uncertainties could be used in the cali-187 bration. However, there is also considerable variability from batch to batch 188 within a laboratory and between different laboratories that is not reflected in 189 these internal analytical uncertainties. Thus the uncertainties reported with 190 the measured concentrations may overstate the precision of these measure-191 ments. 192

In the CRONUS-Earth project, an inter-laboratory comparison was per-193 formed to more broadly quantify the uncertainty in measurements of ¹⁰Be, 194 ²⁶Al, and ¹⁴C concentrations in samples from two reference materials (Jull 195 et al., 2013). Repeated measurements were taken from several laboratories. 196 The coefficient of variation (CV) of such a set of measurements is defined 197 to be the ratio of the standard deviation to the mean. The coefficients of 198 variation in the laboratory inter-comparison were higher than typical stated 199 analytical uncertainties. Furthermore, the coefficient of variation for ¹⁰Be 200 and ²⁶Al were larger for samples with lower concentrations of the cosmogenic 201 nuclides. 202

For ¹⁰Be, we compute an uncertainty in the measured concentration based 203 on interpolation between the CV for high–concentration samples (2.3%) at 204 a concentration of 3.47×10^7 atoms/gram) and low-concentration samples 205 (3.6% at a concentration of 2.13×10^5 atoms/gram.) We then use this com-206 puted uncertainty or the stated analytical uncertainty, whichever is larger. 207 Similarly, for ²⁶Al, we interpolate between the CV for high-concentration 208 samples (4.9% at 1.45×10^8 atoms/gram) and the CV for low-concentration 209 samples (10.1% at 1.06×10^6 atoms/gram.) We use the larger of this com-210 puted uncertainty and the stated analytical uncertainty. For ^{14}C , we use an 211 uncertainty of 7.3% of the measured concentration or the stated analytical 212 uncertainty, whichever is larger. For 36 Cl, we use an uncertainty of 5% of the 213 measured concentration or the stated uncertainty, whichever is larger (Mar-214 rero, 2012). For ³He, no inter-laboratory comparison results were available 215 and so we simply used the stated analytical uncertainties. 216

In all cases, the uncertainty, $\sigma_{i,j}$, used in our calculations is at least as large as the analytical uncertainty. In most cases, $\sigma_{i,j}$ is considerably larger than the analytical uncertainty. This has the effect of reducing χ^2 and makes it easier to pass the goodness-of-fit test.

In this study we have used cross-validation as a way to check that the 221 fitted production rates are insensitive to the inclusion or exclusion of any 222 particular calibration site. After finding the best-fitting production rate for 223 a nuclide using the entire primary calibration data set, we construct subsets 224 of the primary calibration data set in which one site at a time is removed 225 from the data set. We then repeat the calibration process using each of these 226 subsets of the data. For example, suppose that we have data from three 227 calibration sites, A, B, and C. in the calibration of the production rate we 228 first fit the production rate using data from the A, B, and C. We then repeat 220 the calibration using data from the subsets (A, B), (A,C), and (B,C). 230

In theory, if the model fits the data well, then fitted production rates should be similar for each subset of the data. However, if the best-fitting production rate varies substantially over the different subsets of the calibration data, then this is indication that one or more of the calibration sites is having a very large influence on the fitted production rate.

A minor complication in the calibration process is that for some cosmo-236 genic nuclides production by muons and thermal neutrons is significant. The 237 models used for production by these pathways are discussed in Marrero et al. 238 (2014a). These models also involve production rate parameters that can be 239 fit to data. For production of ³⁶Cl by thermal neutrons, a fixed parameter 240 of $P_{f,(0)} = 704$ is used (Marrero, 2012). The production rates for the muon 241 production pathway are separately estimated using a process that will be 242 described in a forthcoming paper. The specific values used for the various 243 scaling frameworks are given in the online appendix. However, these produc-244 tion rates are weakly coupled with the spallation production rates in that 245 estimates of the spallation production rates are used in the calibration of the 246 muon and thermal neutron production rate parameters and vice versa. In 247 practice we have used an iterative approach in which we alternate calibration 248 of the spallation production rates with calibration of the muon production 249 rates until the rates converge to values that are stable to at least four digits. 250

251 3. Data Sets

The CRONUS-Earth Project was funded, in part, to identify, sample, and analyze nuclides from calibration sites that would improve on prior calibration efforts. In the summer of 2010 a suite of primary calibration sites was agreed upon by consensus of the CRONUS-Earth participants. These consisted partly of sites identified and sampled by CRONUS and partly of sites from previous studies that were considered especially reliable. Since that time a number of new calibration studies have been published, but the procedure did not permit them to be added into the calibration in an ad-hoc fashion. The calibration data set in this paper is therefore limited to those highest quality sites agreed upon in 2010.

The data sets were carefully scrutinized to provide accurate values for each 262 parameter. For the CRONUS-Earth sites, every parameter was measured and 263 documented in the field and lab with photos available as appendices to the 264 papers documenting the sites, in addition to the original sample collection 265 notes. For previously published studies, authors were contacted to gather any 266 information that was not explicitly included in the publication. If missing 267 information could not be obtained the study was not used in this calibration 268 effort. 269

The data sets were divided into categories based on the quality or com-270 pleteness data from the site. Primary calibration sites have little uncertainty 271 in the parameters (such as location, independent age constraints, and erosion 272 rate) and have an internally consistent data set. All samples in the secondary 273 data set have independent age constraints, but do not meet one or more of 274 the strict criteria for the primary data sets. For example, sites with uncer-275 tainty in the erosion rate or the possibility of snow cover were categorized as 276 secondary sites. These decisions were based on the authors' interpretation of 277 the geological evidence and different interpretations of the available evidence 278 could well have led to different results. The primary and secondary data sets 279 are summarized in Tables 2 through 4. Data from the primary calibration 280 sites were used in the actual calibration of the production rates. Data from 281 the secondary sites was used only to check the fitted model. 282

For the CRONUS-Earth data sets, the description includes a discussion on any samples that were removed. For the previously published studies, most of the information is taken directly from the original papers. The samples used for ²⁶Al calibration are simply the subset of the full ¹⁰Be data set that also had ²⁶Al measurements made. For that reason they are not explicitly discussed in this section.

The primary and secondary calibration sites are summarized in Tables 2 through 4. In the appendix, available at http://euler.nmt.edu/~brian/appendix.zip, there is a spreadsheet including all of the data. In the spreadsheet, data sets are color-coded to indicate which parameters are directly from the paper and which parameters were calculated or estimated as part of the CRONUS-Earth project. Although more recent calibration papers may have been published, this paper is based on data that were available at the time that data set for this paper was finalized in late 2012.

²⁹⁷ Several general procedures were used for all samples of all nuclides, unless²⁹⁸ we had site-specific information for the parameters.

- Atmospheric pressure was calculated for all samples based on the latitude, longitude, and elevation of the sample using a geographically variable elevation-pressure relationship derived from the ERA-40 reanalysis (Uppala et al., 2005) as implemented in the CRONUScalc program.
- 2. If thickness was not provided or was listed as 0 in the publication, a
 thickness of 0.1 cm was used because a non-zero sample thickness is
 required in the program.
- 306 3. If density was not provided, the rock type was used to estimate a general
 307 lithology-specific density.
- 4. Collection years were assumed to be two years prior to the publication
 date unless the date was known by other means.
- 5. Unless already explicitly stated in the publication, authors were con-310 tacted to confirm the ¹⁰Be AMS standard that was used. If necessary, 311 concentrations were renormalized to the standard of Nishiizumi et al. 312 (2007), using the procedure employed by the Balco et al. (2008) calcula-313 tor. The calculations assume a ¹⁰Be half-life of 1.387 Myr (Korschinek 314 et al., 2010; Chmeleff et al., 2010) and an ²⁶Al half life of 7.05 Myr 315 (Nishiizumi, 2004). To ensure consistency between measurements from 316 different AMS laboratories, all ¹⁰Be data used in the calibration are nor-317 malized to the Nishiizumi 01-5-4 standard with an assumed ${}^{10}\text{Be}/{}^{9}\text{Be}$ 318 ratio equal to 2.851×10^{-12} (Nishiizumi et al., 2007). This is equiva-319 lent to the 07KNSTD normalization of the CRONUS calculator (Balco 320 et al., 2008). Note therefore that production rates derived from this 321 study should only be used with ¹⁰Be data normalized to this same 322 standard value. Likewise, all ²⁶Al data used in the calibration are nor-323 malized to the ²⁶Al/²⁷Al standard series described in Nishiizumi (2004), 324 and production rates should only be applied to Al-26 data so normal-325 ized. Samples for which the analytical standard could not be identified 326 were not used. 327
- 328 329

6. Uncertainties on concentrations were rounded to two significant figures. Concentrations were then rounded to conform with the uncertainties.

7. If horizon values were present, as they were for all CRONUS-Earth
 data sets, the attenuation length has been calculated to include the

topographic effect. In previously published papers, shielding information was typically not available and the standard attenuation length is calculated based on latitude, longitude, elevation, and pressure using methods discussed in Gosse & Phillips (2001).

335 336

332

333

334

337

8. Independent ages based on radiocarbon measurements were recalculated using CALIB 6.0 (Stuiver et al., 2005; Stuiver & Reimer, 1993).

Site	Age (yr)	Uncertainty (yr)	¹⁰ Be	²⁶ Al	³ He	$^{14}\mathrm{C}$	$^{36}\mathrm{Cl}$
ANT	sat	NA				14	
ARG-O	108700	2800			9		
ARG-Y	67800	3000			4		
ICE-MO	8060	120			8		
ICE-MY	5210	110			6		
ICE-O	10330	80			4		
ICE-Y	4040	250			4		
ID	18240	300			3		
NCHL	sat	NA				11	
NZ	9632	50	7			4	
OR-Y	7666	50			3		
OR-O	8571	409			1		
PERU	12260	110	27	10			10
PPT	18240	300	39	25		19	
SCOT	11640	300	29	18		16	4
TAB	18140	300			20		10
WMDV	sat	NA				25	

Table 2: Summary of primary calibration sites. The number of sample concentration measurements is given for each nuclide at each site. The number of sample concentration measurements includes repeated measurements of some samples. See the online appendix for details on the individual samples. Site ages are in years before 1950. "sat" indicates saturated samples and "NA" indicates that uncertainty in the site age is not applicable to the ages of saturated samples.

338 3.1. Primary Calibration Data Sets

Lake Bonneville, Utah, USA (TAB and PPT). Samples were collected from the Tabernacle Hill basalt flow (TAB) for ³He and ³⁶Cl calibration.

Site	Age (yr)	Uncertainty (yr)	¹⁰ Be	²⁶ Al	³ He	$^{36}\mathrm{Cl}$
BL	13040	85	3	3		16
BRQ	13000	100	7			
CA-O	12701	59			1	
CA-Y	3247	84			1	
CAN-O	281000	19000			4	
CAN-Y	152000	26000			3	
CL	2848	69			6	
EV-QTZ	9940	300				8
HAW-M	8230	80			3	
HAW-O	149000	23000			1	
HAW-Y	1470	50			1	
LB	7091	130			2	
NE	13840	250	14	14		7
NZM	18202	200	10			
OL	6012	111	7			

Table 3: Summary of secondary calibration sites. The number of sample concentration measurements is given for each nuclide at each site. The number of sample concentration measurements includes repeated measurements of some samples. See the online appendix for details on the individual samples. Site ages are in years before 1950. Part 1 of 2.

Quartzite samples were collected from Promontory Point (PPT) from a wavepolished shoreline for ¹⁰Be, ²⁶Al, and ¹⁴C calibration. Both sites are described in Lifton et al. (2014a). One Be laboratory's set of Promontory Point ¹⁰Be samples were removed due to a laboratory error. Chlorine samples are feldspar mineral samples. Additional Ca-feldspar separates data from TAB were included from Stone et al. (1996).

Isle of Skye and Highlands, Scotland, UK (SCOT). This is primarily new
CRONUS-Earth data Marrero et al. (2014b). Additional samples were previously measured by John Stone (Evans et al., 1997). These samples were
collected from glacially related rock falls and moraines and contain samples
appropriate for ¹⁰Be, ²⁶Al, ¹⁴C, and ³⁶Cl. One site was removed from the
³⁶Cl calibration due to evidence of possible inheritance. All chlorine samples
were mineral separates.

³⁵⁴ *Quelccaya*, *Peru (PERU)*. This is original CRONUS-Earth data. Sam-³⁵⁵ ples are from a set of well-dated moraines formed by ice cap fluctuations. Radiocarbon age constraints are taken from Kelly et al. (2012). Chlorine
samples are feldspar mineral samples.

Iceland (ICE-Y, ICE-O, ICE-MY, ICE-MO). These are all samples from
 previously published studies. Helium samples are described in Licciardi et al.
 (2006). This includes samples collected from various flows (older, middle
 older, middle younger, younger.)

New Zealand (NZ). This data was previously published in Putnam et al. (2010). The samples are from a rock fall deposit.

Helium Calibration Sites (ARG-O, ARG-Y, OR-Y, OR-O, ID). This is primarily a compilation of previously published data summarized in Goehring et al. (2010). Argentina sites ARG-O and ARG-Y are described in Ackert Jr. et al. (2003). Oregon sites OR-Y and OR-O are described in Cerling & Craig (1994). Idaho site ID is described in Poreda & Cerling (1992).

Saturated ¹⁴C sites (ANT, NCHL, WMDV). The ANT samples come 369 from sites in the Transantarctic Mountains and Prince Charles Mountains in 370 Antarctica. They were collected from bedrock surfaces and large erratic boul-371 ders bevond the mapped limits of last glacial maximum (LGM) ice advance 372 at each site. The NCHL samples are from Northern Chile. The WMDV 373 samples are from the White Mountains and Death Valley in California. The 374 Chilean and eastern California samples were collected from bedrock outcrops 375 and boulders on alluvial surfaces with geomorphic evidence of long-term sur-376 ficial stability and antiquity. 377

378 3.2. Secondary Data Sets

Puget Lowlands, Washington, USA (PUG). This is CRONUS-Earth data described in Marrero et al. (2014c). Radiocarbon age constraints on the deglaciation age of the area were taken from Swanson & Caffee (2001). The chlorine samples from this site include both mineral separates and whole-rock samples.

³⁸⁴ Breque, Peru (BRQ). This glacial moraine data was previously published ³⁸⁵ in Farber et al. (2005). We included only the Quebrada Rurec samples.

Sierra Nevada Sites, California, USA (BL, SN, SNE-K, SNE-CL, SNP-O, SNP-M, SNP-Y). These are primarily samples from previous studies
(Evans et al., 1997; Phillips et al., 2009, 2014; Nishiizumi et al., 1989).
The ¹⁰Be data from Nishiizumi (1989) (SN) and ³⁶Cl data from Evans et al. (1997) (SNE-K, SNE-CL) are from glacial moraines at the same location.
The Phillips (2009) (SNP-O, SNP-M, SNP-Y) data all includes samples from glacial moraines, but from a different location in the Sierra Nevada. New

CRONUS-Earth samples were collected from erratics at the Baboon Lakes (BL) site. Chlorine samples from the Baboon Lakes site include both mineral separates (feldspar and biotite) as well as whole-rock samples. The Evans et al. (1997) study used K-feldspar separates except for one set (SNE-K) that consisted of high-Cl feldspars that were finely ground and from which Cl was separated by leaching.

Littleton-Bethlehem Moraine, New England, USA (NE). This is CRONUS-Earth data described in Balco et al. (2009). Samples are from moraines dated using varve chronology. The age constraints are taken from Balco et al. (2009). This site is treated as a secondary calibration site due to concerns about erosion and cover. All chlorine samples were K-feldspar separates.

Phillips legacy calibration sites (PH1, PH2, PH3, PH4, PH5, PH6, PH7, 404 PH8, PH9, PH10, PH11, PH12). These are data previously published in 405 Phillips et al. (1996) and revised in Phillips et al. (2001). This data set con-406 tains many sites and landforms including basalt flows and glacial moraines. 407 These are named PH1, PH2, PH3, etc. up to PH12. See Table 4 and the 408 appendix to see specific ages and locations. Some sample sets were removed 409 from the 1996 data set due to new information about the uncertainty in the 410 independent age or other problems with the data set. All chlorine analyses 411 were whole-rock samples. 412

New Zealand (NZM). This data was previously published in Putnam et al.
 (2010). These samples are from a glacial moraine near the NZ site.

⁴¹⁵ Norway (OL and YDC). These data were previously published in Goehring ⁴¹⁶ et al. (2012b,a).

⁴¹⁷ Puerto Bandera Moraines, Patagonia (PAT). These data were previously ⁴¹⁸ published in Kaplan et al. (2011). Only the Puerto Bandera Moraines sam-⁴¹⁹ ples were included.

Titcomb Basin, USA (WY). These data were previously published in 420 Gosse et al. (1995). Samples were collected from glacial landforms. These 421 data were renormalized to the current ¹⁰Be standard of Nishiizumi et al. 422 (2007). Since the finalization of the data set for this paper, additional ques-423 tions have been raised about the proper normalization of these ¹⁰Be measure-424 ments (Gosse, 2014). However, removing these samples from the secondary 425 data sets for 10 Be would result in a change of less than 2% in the RMSE and 426 would not materially affect the conclusions of this paper. 427

428 Scotland, UK (EV-QTZ). These data were previously published in Evans 429 (2001). The samples were collected from glacial landforms and are quartz 430 mineral separates.

Helium Calibration Sites (CA-O, CA-Y, ID, CAN-Y, CAN-O, SCLY-431 O, SCLY-Y, HAW-O, HAW-M, HAW-Y, YAP, SBLK, CL, LB). This is 432 primarily a compilation of previously published data summarized in Goehring 433 et al. (2010). California sites CA-O, and CA-Y are described in Cerling &434 Craig (1994). Idaho site ID is described in Poreda & Cerling (1992). Canary 435 Islands sites CAN-Y and CAN-O are described in Dunai & Wijbrans (2000). 436 Sicily sites SCLY-O and SCLY-Y and Hawaii sites HAW-O, HAW-M, and 437 HAW-Y are described in Blard et al. (2006). Site YAP is described in Cerling 438 & Craig (1994); Licciardi et al. (1999). Sites SBLK, CL, and LB are described 439 in Licciardi et al. (1999). 440

Site	Age (yr)	Uncertainty (yr)	¹⁰ Be	²⁶ Al	³ He	$^{36}\mathrm{Cl}$
PAT	12830	240	8			
PH1	1980	60				3
PH10	15310	180				1
PH11	17230	260				2
PH12	18990	170				1
PH2	3130	80				1
PH3	5910	160				3
PH4	8640	160				3
PH5	8870	160				3
PH6	9940	1000				2
PH7	11170	50				1
PH8	11770	470				3
PH9	14940	270				3
PUG	15500	500	3	3		
SBLK	2752	17			7	
SCLY-O	41000	3000			2	
SCLY-Y	33000	2000			2	
SN	15750	500	10	10		
SNE-K	15750	500				8
SNE-CL	15750	500				4
SNP-M	15750	500				5
SNP-O	16000	500				4
SNP-Y	13250	300				4
WY	12040	700	9	9		
YAP	2453	780			7	
YDC	11592	100	8			

Table 4: Summary of secondary calibration sites. The number of sample concentration measurements is given for each nuclide at each site. The number of sample concentration measurements includes repeated measurements of some samples. See the online appendix for details on the individual samples. Site ages are in years before 1950. Part 2 of 2.

441 4. Results and Discussion

Using the seven scaling frameworks discussed in Section 1 and the data 442 sets described in Section 3, calibrations were performed for the spallation 443 production rates for each of the cosmogenic nuclides ³He, ¹⁰Be, ¹⁴C, ²⁶Al, 444 and ³⁶Cl. The resulting reduced χ^2 values are shown in Table 5. The cor-445 responding *p*-values for each calibration are shown in Table 6. Most of the 446 calibrations fail the χ^2 goodness-of-fit test with large reduced χ^2 values and 447 small corresponding p-values. The only calibrations that pass the χ^2 test are 448 the calibrations for ²⁶Al and ³⁶Cl, and these calibrations only pass the test 449 when using the Sa, Sf, St, and Lm scaling frameworks. 450

Although failure to pass the goodness-of-fit test is strong evidence that the data and model are inconsistent, passing the goodness-of-fit test does not prove that the model and observations are correct. Further examination of the primary calibration results and evaluation of the secondary calibration data shows that there are significant issues with all of the fits described here.

For reference, we have also given the values of the best-fitting production rates in Table 7. As solutions to the minimization problem in Equation (1), these values are precise to at least 4 digits. However, because of the failure of the goodness-of-fit tests described above we cannot associate any statistical uncertainty with these production rates. Using other calibration data we might obtain very different production rates.

	¹⁰ Be	²⁶ Al	³ He	$^{14}\mathrm{C}$	³⁶ Cl
Degrees of Freedom	101	52	61	88	22
Sa	1.64	0.88	4.37	2.13	1.07
St	1.53	1.06	4.58	2.14	1.39
Sf	1.54	0.93	4.38	2.07	1.12
Lm	1.49	1.04	4.39	2.14	1.33
De	4.59	2.62	4.29	2.31	4.21
Du	4.40	2.59	4.09	2.25	4.24
Li	3.69	2.20	4.27	2.18	3.59

Table 5: Reduced χ^2_{ν} values for the calibrations. The seven scaling frameworks are denoted by the two-letter abbreviations described in Section 1. As solutions to the minimization problem in Equation (1), these values are precise to at least 4 digits. However, because of the failure of the goodness-of-fit tests described above we cannot associate any statistical uncertainty with these production rates.

	¹⁰ Be	²⁶ Al	³ He	$^{14}\mathrm{C}$	$^{36}\mathrm{Cl}$
Sa	5.12×10^{-5}	$7.15 imes10^{-1}$	$< 1.00 \times 10^{-14}$	3.70×10^{-9}	$3.66 imes10^{-1}$
St	4.61×10^{-4}	$3.57 imes10^{-1}$	$< 1.00 \times 10^{-14}$	2.65×10^{-9}	$1.05 imes10^{-1}$
Sf	4.11×10^{-4}	$6.25 imes10^{-1}$	$< 1.00 \times 10^{-14}$	1.37×10^{-8}	$3.20 imes10^{-1}$
Lm	1.00×10^{-3}	$3.98 imes10^{-1}$	$< 1.00 \times 10^{-14}$	2.57×10^{-9}	$1.39 imes10^{-1}$
De	$< 1.00 \times 10^{-14}$	1.64×10^{-9}	$< 1.00 \times 10^{-14}$	3.78×10^{-11}	1.20×10^{-10}
Du	$< 1.00 \times 10^{-14}$	3.17×10^{-9}	$< 1.00 \times 10^{-14}$	1.89×10^{-10}	9.54×10^{-11}
Li	$< 1.00 \times 10^{-14}$	1.42×10^{-6}	$< 1.00 \times 10^{-14}$	1.01×10^{-9}	2.41×10^{-8}

Table 6: *p*-values for calibrations. Values of less than 5.0×10^{-2} indicate a failed goodness-of-fit test. Values of larger than 5.0×10^{-2} are shown in bold face.

In the remainder of this section we will present detailed results for the Sa scaling framework. Results for the other scaling frameworks are presented in the online appendix to the paper. The results for the St, Sf, and Lm scaling frameworks are generally qualitatively similar to the results for the Sa framework. The results for the De, Du, and Li scaling frameworks have much poorer fit to the data as shown by the χ^2_{ν} values in Tables 5 and 6.

468 4.1. ¹⁰Be Spallation Production Rate

Using the Sa scaling framework, the best-fitting production rate for ¹⁰Be 469 from quartz was $P_{s,Be} = 3.92$ atoms/g/year. The reduced χ^2 value was 1.64 470 with 101 degrees of freedom. The corresponding *p*-value was 5.12×10^{-5} . 471 Thus this fit fails the goodness-of-fit test. Obtaining this fit required ex-472 tremely large adjustments to the site ages. For example, the nominal age for 473 the PPT site of $18,240 \pm 300$ years was adjusted by 4.3 standard deviations 474 to 19,540 years. Such an extremely large deviation from the nominal age 475 seems implausible. 476

Figure 1 shows the ratios of the measured ¹⁰Be concentrations to pre-477 dicted ¹⁰Be concentrations for the calibration samples at the four calibration 478 sites, NZ, PPT, SCOT, and PERU. The measured concentrations have been 470 normalized by dividing by the predicted concentrations because sample to 480 sample variations in thickness, density, assumed erosion rate and altitude can 481 lead to substantial differences in the measured and predicted concentrations. 482 Note that the individual samples at each site have normalized concentrations 483 that cluster reasonably well, although there is more spread than we might 484 expect from the laboratory inter-comparison (Jull et al., 2013). Furthermore, 485

Nuclide	$P_{s,Be}$	$P_{s,Al}$	$P_{s,He}$	$P_{s,C}$	$P_{s,Cl,Ca}$	$P_{s,Cl,K}$
Sa	3.92	28.54	114.55	12.76	56.27	156.09
St	4.01	27.93	118.20	12.24	52.34	150.72
Sf	4.09	28.61	118.64	12.72	56.61	153.95
Lm	4.00	27.93	117.23	12.22	51.83	151.64
De	3.69	26.26	122.47	12.49	55.90	128.25
Du	3.70	26.29	122.75	12.44	55.27	128.89
Li	4.06	28.72	131.32	13.42	60.66	142.24

Table 7: Best-fitting production rates for the various scaling frameworks. $P_{s,Be}$ is the production rate of ¹⁰Be by neutron spallation in atoms per gram of quartz per year. Similarly, $P_{s,Al}$, $P_{s,He}$, and $P_{s,C}$ are production rates for ²⁶Al, ³He, and ¹⁴C by neutron spallation in units of atoms per gram of quartz per year. $P_{s,Cl,Ca}$ is the production rate of ³⁶Cl by neutron spallation of Ca in units of atoms per gram of Ca per year. $P_{s,Cl,K}$ is the production rate of ³⁶Cl by neutron spallation of K in units of atoms per gram of K per year.

there are significant site-to-site deviations from the model. These deviations
are on the order 10%, which is large compared with the independent age
uncertainties and the concentration uncertainties.

We also performed cross-validation of the calibration, leaving one site at 489 a time out of the computation. The resulting fitted values of $P_{s,Be}$ were 3.83 490 (leaving out PPT), 3.89 (leaving out SCOT), 3.93 (leaving out NZ), and 4.02 491 (leaving out PERU). Since the individual sample measurements are precise 492 to 3% or better and averaging over multiple samples further reduces the un-493 certainty, the differences between these best-fitting production rates cannot 494 easily be explained by random variation in individual sample measurements. 495 This is a further indication of some inconsistency between the sites or an 496 error in the scaling framework. 497

We next used the fitted production rate to compute ages for samples 498 from ten secondary sites. Figure 2 shows the ratios of computed ages to 499 independent ages for the samples from these secondary sites. No uncertainties 500 have been attached to these ratios since there is no way to compute such an 501 uncertainty without detailed knowledge of the uncertainty in the individual 502 sample measurements and a properly statistically calibrated production rate. 503 Note that nearly all of the computed ages are older than the independent 504 ages for the sites. This suggests that the fitted production rate is biased too 505 low. Due to the possibility of outliers, we took the median of the computed 506



Figure 1: Ratios of measured concentrations to predicted concentrations for ¹⁰Be calibration sites. Site 1=NZ, Site 2=PPT, Site 3=SCOT, Site 4=PERU.

ages for each secondary site, and then computed the root mean square error
(RMSE) for each of the ten sites. The RMSE of the median site ages was
8.3%.

510 4.2. ²⁶Al Spallation Production Rate

The calibration of the ²⁶Al production rate was done using data from only 511 three sites, PPT, SCOT, and PERU. From the point of view of experimental 512 design, using only three primary calibration sites to fit the ²⁶Al production 513 rate results in a very limited test of the scaling frameworks. It would have 514 been desirable to have several more primary calibration sites. As described 515 in Section 2, the measurements of ²⁶Al concentrations were given an assumed 516 uncertainty of approximately 10%, which is larger than the analytical uncer-517 tainties supplied with the measurements. These measures are considerably 518 less precise than the measurements of ¹⁰Be concentrations. For these two 519 reasons, obtaining a fit that passed a goodness-of-fit test was easier in the 520 case of 26 Al than in the case of 10 Be. 521

The resulting fitted production rate was $P_{s,Al} = 28.54$ atoms/g quartz/year. The calibration resulted in a reduced χ^2 value of 0.88 with 52 degrees of freedom, for a *p*-value of 0.71. Although this fit passes the goodness-of-fit test, there are other reasons to be concerned about the fit.



Figure 2: Ratios of Computed Age to Independent Age for secondary 10-Be calibration sites. Site 1=NZM, Site 2=SN, Site 3=NE, Site 4=YDC, Site 5=OL, Site 6=BL, Site 7=WY, Site 8=PUG, Site 9=PAT, Site 10=BRQ.



Figure 3: Ratios of measured concentration to predicted concentration for ²⁶Al calibration sites. Site 1=PPT, Site 2=SCOT, Site 3=PERU.

Figure 3 shows the ratios of the measured ²⁶Al concentrations to predicted ²⁶Al concentrations for the calibration samples at the three calibration sites. Note that the data for each site is scattered over a range of 20% or more with some apparent outlier values. Under cross-validation, fitted production rates were $P_{s,Al} = 28.07$ (leaving out SCOT), $P_{s,Al} = 28.48$ (leaving out PPT), and $P_{s,Al} = 29.14$ (leaving out PERU).

Figure 4 shows the ratios of computed ages to independent ages for samples from five secondary calibration sites. As with ¹⁰Be there appear to be systematic offsets from the model at different sites, although the ²⁶Al data is somewhat more scattered. Unlike ¹⁰Be, these data do not seem to show a bias toward old ages. The RMSE of the median ages for the sites was 7.1%.

⁵³⁷ 4.3. ³He Spallation Production Rate

The fitted production rate was $P_{s,He} = 114.55$ atoms/g quartz/year. The calibration resulted in a reduced χ^2 value of 4.37 with 61 degrees of freedom, for a *p*-value of less than 1.00×10^{-14} . Thus the fit fails the goodness-of-fit test. Figure 5 shows the ratios of measured ³He concentrations to predicted ³He concentrations for samples from the primary calibration sites. Here, there appear to be systematic site offsets as well as scattered measurements



Figure 4: Ratios of Computed Age to Independent Age for secondary ²⁶Al calibration sites. Site 1=SN, Site 2=NE, Site 3=BL, Site 4=WY, Site 5=PUG.

and outliers at some sites. Under cross-validation, production rates from $P_{s,He} = 112.46$ (minus ARG-O) to $P_{s,He} = 117.47$ (minus OR-Y) were obtained.

Figure 6 shows the ratios of computed ages to independent ages for samples from thirteen secondary sites. Again, there is evidence of significant site-to-site effects, as well as some outlier samples. Computed ages for multiple samples at the same site are quite scattered, an indication of possible problems with the concentration measurements. The RMSE of the median site ages is 27.1%.

The ³He calibration data sets contain samples from sites with a much larger range of ages than the other calibration data sets. Issues with the time-dependent scaling factors may have contributed to the very large siteto-site variations in the data.

⁵⁵⁷ ¹⁴C Spallation Production Rate

The fitted production rate was $P_{s,C} = 12.76$ atoms/g quartz/year. The calibration resulted in a reduced χ^2 value of 2.13 with 88 degrees of freedom, for a *p*-value of 3.70×10^{-9} . Thus the fit fails the goodness-of-fit test. Figure



Figure 5: Ratios of measured concentration to predicted concentration for ³He samples from primary calibration sites. Site 1=ARG-O, Site 2=ARG-Y, Site 3=OR-Y, Site 4=OR-O, Site 5=TAB, Site 6=ID, Site 7=ICE-Y, Site 8=ICE=MY, Site 9=ICE-MO, Site 10=ICE-O.

⁵⁶¹ 7 shows the ratios of the measured concentrations to the predicted concen-⁵⁶² trations for samples from the six calibration sites. For ¹⁴C, we assumed ⁵⁶³ an uncertainty of 7.3%. It appears that the data are too scattered to be ⁵⁶⁴ consistent with this assumption. For ¹⁴C, issues with sample concentration ⁵⁶⁵ measurements appear to be more significant than any site-to-site variability.

⁵⁶⁶ 4.4. ³⁶Cl Spallation Production Rates

The fitted production rates were $P_{s,Cl,K} = 156.09$ atoms/gram K/year 567 and $P_{s,Cl,Ca} = 56.27$ atoms/gram Ca/year. The reduced χ^2 value was 1.07 568 on 22 degrees of freedom, for a p-value of 0.366. Although this fit passes the 569 goodness-of-fit test, it is based on data from only 3 calibration sites. From 570 an experimental design point of view, using only three sites to calibrate 571 two production rates provides a very poor test of the scaling frameworks. 572 Furthermore, of these three sites, only TAB had any substantial contribution 573 to ³⁶Cl by spallation of calcium. As a result, this production rate is effectively 574 determined by the TAB site alone. 575

⁵⁷⁶ Under cross-validation, the fitted production rates were extremely unsta-⁵⁷⁷ ble. Production rates from calcium from $P_{s,Cl,Ca} = 56.19$ (leaving out PERU)



Figure 6: Ratios of Computed Age to Independent Age for secondary ³He calibration sites. Site 1=CAN-Y, Site 2=CAN-O, Site 3=SCLY-O, Site 4=SCLY-Y, Site 5=HAW-M, Site 6=HAW-Y, Site 7=HAW-O, Site 8=CA-Y, Site 9=CA-O, Site 10=YAP, Site 11=SBLK, Site 12=CL, Site 13=LB.



Figure 7: Ratios of measured concentrations to predicted concentrations for 14 C calibration sites. Site 1=PPT, Site 2=SCOT, Site 3=NZ, Site 4=WMDV, Site 5=NCHL, Site 6=ANT.

to $P_{s,Cl,Ca} = 1144.70$ (leaving out TAB) were obtained. For production from potassium, we obtained production rates from $P_{s,Cl,K} = 132.98$ (leaving out TAB) to $P_{s,Cl,K} = 166.93$ (leaving out PERU).

Figure 9 shows the ratios of computed ages to independent ages for 20 secondary calibration sites. As with other nuclides, there is considerable scatter in the data from some sites, and there appear to be systematic offsets at certain sites. The RMSE of the median ages for the sites is 17.7%.

585 5. Conclusions

The results of the fitting exercise clearly show that the Sa, Sf, St, and Lm 586 scaling frameworks performed much better than the neutron monitor based 587 scaling frameworks (De, Du, Li) in fitting the primary calibration data sets. 588 In all cases, χ^2 values are much lower for the Sa, Sf, St, and Lm frameworks. 589 This result is consistent with the conclusions of Lifton et al. (2014b), who 590 showed that the neutron monitor based scaling frameworks most likely over-591 estimate the altitude dependence of cosmogenic-nuclide production because 592 of unrecognized multiplicity effects in the neutron monitor data on which 593 they are based. Thus, both physical arguments and fitting to calibration 594



Figure 8: Ratios of measured concentration to predicted concentration for ³⁶Cl calibration sites. Site 1=TAB, Site 2=PERU, Site 3=SCOT.

⁵⁹⁵ data indicate that these scaling frameworks are not, in general, expected to ⁵⁹⁶ yield accurate exposure-dating results.

We observed very little difference in χ^2 values between the Sa, Sf, St, 597 and Lm scaling frameworks. Thus, despite the significant difference in com-598 plexity between these scaling frameworks, available data are not sufficient 599 to show whether one performs better than the other. The Sa and Sf scaling 600 frameworks include many physical aspects of cosmogenic-nuclide production 601 that are not included in the St and Lm scaling frameworks, specifically, a 602 full spectral representation of the neutron flux and the ability to incorporate 603 direct laboratory cross-section measurements. Thus, given best-fitting refer-604 ence production rates fit to our calibration data set for both these scaling 605 frameworks, the scaling frameworks predict different production rates, and 606 therefore exposure ages, for some locations and ages. The Sa and Sf scaling 607 frameworks, in particular the nuclide-dependent Sa scaling framework, are 608 more closely linked to the physical processes involved in cosmogenic-nuclide 609 production, whereas the St and Lm scaling frameworks are primarily em-610 pirical. Thus, arguments based on physical principles give strong reason to 611 believe that the Sa and Sf frameworks will yield more accurate predictions 612 for locations and ages that are very different from those represented in the 613



Figure 9: Ratios of Computed Age to Independent Age for quantitative secondary ³⁶Cl calibration sites. Site 1=NE, Site 2=PH11, Site 3=PH 3, Site 4=PH 4, Site 5=PH9, Site 6=PH8, Site 7=PH1, Site 8=PH5, Site 9=PH6, Site 10=PH10, Site 11=PH2, Site 12=PH12, Site 13=PH7, Site 14=BL, Site 15=SNE-K, Site 16=SNE-Cl, Site 17=SNP-M, Site 18=SNP-O, Site 19=SNP-Y, Site 20=EV-QTZ.

calibration data set. However, at present we cannot verify this conclusionwith the available data.

It is clear from the results that measured concentrations of cosmogenic 616 nuclides samples collected at the various calibration sites were sometimes 617 much more variable than could be expected given the stated uncertainties in 618 these concentration measurements. This is clearly shown in Figures 5 and 7. 619 It is possible that these measurements are simply much less precise in practice 620 than expected. The comparison of measurements from separate samples also 621 depends on aspects of the individual samples such as the erosion rate, sample 622 thickness, and density. It is possible that errors in these parameters may have 623 contributed to the scatter seen in the calibration data. 624

It is also clear from the results that there are significant unexplained variations from site to site. This apparent bias could be due to problems with the elevation and latitude scaling frameworks, or it could be due to problems with the characterization of the sites, including incorrect assumptions about parameters such as erosion rates and atmospheric pressure. It is also possible that incorrect independent age constraints are a factor.

One of the main goals of the CRONUS-Earth project was to provide the 631 most accurate tools available for geochronological applications of cosmogenic 632 nuclides. As part of that goal, we collected and processed many samples 633 from new geological calibration sites. The goals of this paper are to i) make 634 a quantitative and minimally biased assessment of how well the production 635 rate scaling frameworks that we believe to be the best available are able 636 to reconcile what we believe to be the best available geological calibration 637 data, and ii) use this information to assess the accuracy of exposure-dating 638 applications using these scaling frameworks at unknown sites. The result of 639 this assessment is that the production models could not be statistically fit to 640 the data. Because of this, we cannot infer statistically justifiable production 641 rate uncertainties from the fitting exercise. 642

Although the calibration did not perform as originally expected, this large-scale calibration effort has provided clear directions for future projects. Further research is needed to address the issue of variability in concentration measurements, especially for ³He and ¹⁴C, and to improve our understanding of scaling frameworks and site characterization in order to understand the underlying cause of the site-to-site variability.

649 6. Statement of Contributions

Brian Borchers developed the statistical methodology and wrote the software that was used to find the best-fitting production rates. Shasta Marrero assembled the data sets described in section 3. Borchers and Marrero wrote the paper. Other coauthors collected and measured samples and selected calibration data for inclusion in the paper. All of the authors have reviewed the final draft of this paper.

656 7. Acknowledgements

This work was funded as part of the CRONUS-Earth program by the US National Science Foundation grants EAR-0345150, EAR-0345820, EAR-0345949, EAR-0345574, EAR-0345835, and EAR-0345817.

660 References

- Ackert Jr., R. P., Singer, B. S., Guillou, H., Kaplan, M. R., & Kurz, M. D.
 (2003). Long-term cosmogenic ³He production rates from ⁴⁰Ar/³⁹Ar and
 K-Ar dated Patagonian lava flows at 47 S. *Earth and Planetary Science Letters*, 210, 119–136.
- Argento, D. C., Stone, J. O., Reedy, R. C., & O'Brien, K. (2014a). Physics based modeling of cosmogenic nuclides part I-radiation transport methods
 and new insights. *Quaternary Geochronology*, . In Press.
- Argento, D. C., Stone, J. O., Reedy, R. C., & O'Brien, K. (2014b). Physicsbased modeling of cosmogenic nuclides part II-key aspects of; *i*; in-situ;/*i*;
 cosmogenic nuclide production. *Quaternary Geochronology*, . In Press.
- Aster, R. C., Borchers, B., & Thurber, C. H. (2012). Parameter Estimation
 and Inverse Problems. (2nd ed.). Amsterdam: Academic Press.

Balco, G., Briner, J., Finkel, R. C., Rayburn, J. A., Ridge, J. C., & Schaefer,
J. M. (2009). Regional Beryllium-10 production rate calibration for lateglacial northeastern North America. *Quaternary Geochronology*, 4, 93–
107.

- Balco, G., Stone, J. O., Lifton, N., & Dunai, T. J. (2008). A complete and
 easily accessible means of calculating surface exposure ages or erosion rates
 from ¹⁰Be and ²⁶Al measurements. *Quaternary Geochronology*, 3, 174–195.
- Blard, P.-H., Pik, R., Lave, J., Marty, B., & Trusdell, F. (2006). Cosmogenic
 ³He production rates revisited from evidences of grain size dependent re lease of matrix-sited helium. *Earth and Planetary Science Letters*, 247, 222–234.
- Cerling, T. E., & Craig, H. (1994). Cosmogenic ³He production rates from 39°
 N to 46° N latitude, western USA and France. *Geochimica et Cosmochimica Acta*, 58, 249–255.
- ⁶⁸⁷ Chmeleff, J., von Blanckenburg, F., Kossert, K., & Jakob, D. (2010). Determination of the ¹⁰be half-life by multicollector icp-ms and liquid scintillation counting. Nuclear Instruments and Methods in Physics Research Section B: Beam Interactions with Materials and Atoms, 268, 192–199.

Desilets, D., Zreda, M. G., & Prabu, T. (2006). Extended scaling factors for
 in situ cosmogenic nuclides: New measurements at low latitude. *Earth and Planetary Science Letters*, 246, 265–276.

⁶⁹⁴ Dunai, T. (2001). Influence of secular variation of the geomagnetic field ⁶⁹⁵ on production rates of in situ produced cosmogenic nuclides. *Earth and* ⁶⁹⁶ *Planetary Science Letters*, 193, 197–212.

Dunai, T. J., & Wijbrans, J. R. (2000). Long-term cosmogenic ³He production rates (152 ka - 1.35 ma) from ⁴⁰Ar/³⁹Ar dated basalt flows at 29° N latitude. *Earth and Planetary Science Letters*, 176, 147–156.

Evans, J. M. (2001). Calibration of the production rates of cosmogenic ³⁶Cl
 from potassium. Ph.D. thesis The Australian National University, Can berra. Doctorate of Philosophy.

 Evans, J. M., Stone, J. O. H., Fifield, L. K., & Cresswell, R. G. (1997).
 Cosmogenic chlorine-36 production in K-feldspar. Nuclear Instruments and Methods in Physics Research B, 123, 334–340.

Farber, D. L., Hancock, G. S., Finkel, R. C., & Rodbell, D. T. (2005). The age and extent of tropical alpine glaciation in the Cordillera Blanca, Peru. *Journal of Quaternary Science*, 20, 759–776.

Goehring, B. M., Kurz, M. D., Balco, G., Schaefer, J. M., Licciardi, J., &
Lifton, N. (2010). A reevaluation of in situ cosmogenic ³He production
rates. *Quaternary Geochronology*, 5, 410–418.

Goehring, B. M., Lohne, O. S., Mangerud, J., Svendsen, J. I., Gyllencreutz,
R., Schaeffer, J., & Finkel, R. (2012a). Erratum: Late glacial and Holocene
¹⁰Be production rates for western Norway. *Journal of Quaternary Science*,
27, 544.

Goehring, B. M., Lohne, O. S., Mangerud, J., Svendsen, J. I., Gyllencreutz,
R., Schaeffer, J., & Finkel, R. (2012b). Late Glacial and Holocene ¹⁰Be
production rates for western Norway. *Journal of Quaternary Science*, 27,
89–96.

Gosse, J. C. (2014). Concerns about normalization of the WY samples.
 Personal Communication.

- Gosse, J. C., Evenson, E. B., Klein, J., Lawn, B., & Middleton, R. (1995).
 Precise cosmogenic ¹⁰Be measurements in western North America: Support for a global Younger Dryas cooling event. *Geology*, 23, 877–880.
- Gosse, J. C., & Phillips, F. M. (2001). Terrestrial in situ cosmogenic nuclides:
 theory and application. *Quaternary Science Reviews*, 20, 1475–1560.
- Jull, A., Scott, E., & Bierman, P. (2013). The CRONUS-Earth intercomparison for cosmogenic isotope analysis. *Quaternary Geochronology*, .
 In Press.
- Kaplan, M. R., Strelin, J. A., Schaeffer, J., Denton, G. H., Finkel, R.,
 Schwartz, R., Putnam, A. E., Goehring, B. M., & Travis, S. G. (2011).
 In-situ cosmogenic ¹⁰Be production rate at Lago Argentino, Patagonia:
 Implications for late-glacial climate chronology. *Earth and Planetary Science Letters*, 309, 21–32.
- Kelly, M. A., Lowell, T. V., Applegate, P. J., Smith, C. A., Phillips, F. M.,
 & Hudson, A. M. (2012). Late glacial fluctuations of Quelccaya Ice Cap,
 southeastern Peru. *Geology*, 40, 991–994.
- Kollar, D., & Masarik, J. (1999). Numerical simulation of particle fluxes and
 cosmogenic nuclide production rates in earth atmosphere. In Acta Physica
 Universitatis Comenianae.
- Korschinek, G., Bergmaier, A., Faestermann, T., Gerstmann, U., Knie, K.,
 Rugel, G., Wallner, A., Dillmann, I., Dollinger, G., Von Gostomski, C. L.
 et al. (2010). A new value for the half-life of ¹⁰be by heavy-ion elastic
 recoil detection and liquid scintillation counting. Nuclear Instruments and
 Methods in Physics Research Section B: Beam Interactions with Materials
 and Atoms, 268, 187–191.
- Lal, D. (1991). Cosmic ray labeling of erosion surfaces: In situ nuclide
 production rates and erosion models. *Earth and Planetary Science Letters*,
 104, 424–439.
- Levenberg, K. (1944). A method for the solution of certain problems in least
 squares. Quarterly of Applied Mathematics, 2, 164–168.
- Licciardi, J. M., Kurz, M. D., Clark, P. U., & Brook, E. J. (1999). Calibration
 of cosmogenic ³He production rates from Holocene lava flows in Oregon,

- USA, and effects of the earth's magnetic field. Earth and Planetary Science
 Letters, 172, 261–271.
- Licciardi, J. M., Kurz, M. D., & Curtice, J. M. (2006). Cosmogenic ³He
 production rates from Holocene lava flows in Iceland. *Earth and Planetary Science Letters*, 246, 251–264.
- Lifton, N., Caffee, M., Finkel, R., Marrero, S., Nishiizumi, K., Phillips, F. M.,
 Goehring, B., Gosse, J., Stone, J., Schaefer, J. et al. (2014a). ¡ i¿ in
 situ¡/i¿ cosmogenic nuclide production rate calibration for the CRONUSEarth project from Lake Bonneville, Utah, shoreline features. *Quaternary Geochronology*, . In Press.
- Lifton, N., Sato, T., & Dunai, T. J. (2014b). Scaling *in situ* cosmogenic
 nuclide production rates using analytical approximations to atmospheric
 cosmic-ray fluxes. *Earth and Planetary Science Letters*, 386, 149–160.
- Lifton, N., Smart, D. F., & Shea, M. A. (2008). Scaling time-integrated in
 situ cosmogenic nuclide production rates using a continuous geomagnetic
 model. *Earth and Planetary Science Letters*, 268, 190–201.
- Lifton, N. A., Bieber, J. W., Clem, J. M., Duldig, M. L., Evenson, P., Humble, J. E., & Pyle, R. (2005). Addressing solar modulation and long-term
 uncertainties in scaling in situ cosmogenic nuclide production rates. *Earth*and Planetary Science Letters, 239, 140–161.
- Marquardt, D. W. (1963). An algorithm for least-squares estimation of non linear parameters. Journal of the Society for Industrial & Applied Mathematics, 11, 431–441.
- Marrero, S. (2012). Calibration of Cosmogenic Chlorine-36. Ph.D. thesis
 New Mexico Institute of Mining and Technology.
- Marrero, S., Phillips, F., Borchers, B., Lifton, N., Aumer, R., & Balco, G.
 (2014a). Cosmogenic nuclide systematics and the CRONUScalc program.
 Submitted for Publication in this issue of *Quaternary Geochronology*.
- Marrero, S. M., Caffee, M. W., Lifton, N., Theirault, B., Phillips, F., Stone,
 J., Owen, L., L.K., F., Ballantyne, C. K., & Sugden, D. E. (2014b). Cosmogenic nuclides from the CRONUS-Earth calibration sites in scotland.
 Submitted for Publication in this issue of *Quaternary Geochronology*.

Marrero, S. M., Phillips, F. M., Caffee, M., Swanson, T., & Hinz, M.
(2014c). Resampling of Puget lowlands yields lower discrepancy in cosmogenic chlorine-36 production rates. Submitted for Publication in this
issue of *Quaternary Geochronology*.

Masarik, J. (2002). Numerical simulation of in-situ produced cosmogenic
nuclides. *Geochimica et Cosmochimica Acta*, 66, A491.

Masarik, J., & Beer, J. (1999). Simulation of particle fluxes and cosmogenic
 nuclide production in the earth's atmosphere. Journal of Geophysical Re search: Atmospheres (1984–2012), 104, 12099–12111.

Masarik, J., Kim, K. J., & Reedy, R. C. (2007). Numerical simulations of in
 situ production of terrestrial cosmogenic nuclides. Nuclear Instruments and
 Methods in Physics Research Section B: Beam Interactions with Materials
 and Atoms, 259, 642–645.

Masarik, J., & Reedy, R. C. (1995). Terrestrial cosmogenic-nuclide production systematics calculated from numerical simulations. *Earth and Plane- tary Science Letters*, 136, 381–396.

Nishiizumi, K. (2004). Preparation of ²⁶al ams standards. Nuclear Instru ments and Methods in Physics Research Section B: Beam Interactions with
 Materials and Atoms, 223, 388–392.

Nishiizumi, K., Imamura, M., Caffee, M. W., Southon, J. R., Finkel, R. C., &
 McAninch, J. (2007). Absolute calibration of ¹⁰be ams standards. Nuclear
 Instruments and Methods in Physics Research Section B: Beam Interac tions with Materials and Atoms, 258, 403–413.

Nishiizumi, K., Winterer, E. L., Kohl, C. P., Klein, J., Middleton, R., Lal,
D., & Arnold, J. R. (1989). Cosmic ray production rates of ¹⁰Be and ²⁶Al
in quartz from glacially polished rocks. *Journal of Geophysical Research*,
94, 9.

Phillips, F., Stone, W. D., & Fabryka-Martin, J. (2001). An improved approach to calculating low-energy cosmic-ray neutron fluxes near the land/atmosphere interface. *Chemical Geology*, 175, 689–701.

Phillips, F., Zreda, M. G., Flinsch, M. R., Elmore, D., & Sharma, P. (1996).
A reevaluation of cosmogenic ³⁶Cl production rates in terrestrial rocks. *Geophysical Research Letters*, 23, 949–952.

Phillips, F. M., Hinz, M., Marrero, S. M., Nishiizumi, K., & Stone, J. O.
(2014). CRONUS-Earth cosmogenic-nuclide calibration sites in the Sierra Nevada, California. Submitted for Publication in this issue of *Quaternary Geochronology*.

Phillips, F. M., Zreda, M. G., Plummer, M. A., Elmore, D., & Clark, D.
(2009). Glacial geology and chronology of Bishop Creek and vicinity, eastern Sierra Nevada, California. *Geological Society of America Bulletin*, 121, 1013–1033.

Poreda, R. J., & Cerling, T. E. (1992). Cosmogenic neon in recent lavas from
the western United States. *Geophysical Research Letters*, 19, 1863–1866.

Putnam, A. E., Schaefer, J. M., Barrell, D. J. A., Vandergoes, M., Denton,
G. H., Kaplan, M. R., Finkel, R. C., Schwartz, R., Goehring, B. M., &
Kelley, S. E. (2010). In situ cosmogenic ¹⁰Be production-rate calibration
from the Southern Alps, New Zealand. *Quaternary Geochronology*, 5,
392–409.

Sato, T., & Niita, K. (2006). Analytical functions to predict cosmic-ray
neutron spectra in the atmosphere. *Radiation Research*, 166, 544–555.

Sato, T., Yasuda, H., Niita, K., Endo, A., & Sihver, L. (2008). Development
of PARMA: PHITS-based analytical radiation model in the atmosphere. *Radiation Research*, 170, 244–259.

Stone, J. (2000). Air pressure and cosmogenic isotope production. Journal
of Geophysical Research, 105, 23753-23760.

Stone, J. O. (2005). Terrestrial chlorine-36 production from spallation of
iron. In 10th International Conference on Accelerator Mass Spectrometry.
Berkeley, California.

Stone, J. O., Allan, G. L., Fifield, L. K., & Cresswell, R. G. (1996). Cosmogenic chlorine-36 from calcium spallation. *Geochimica et Cosmochimica*Acta, 60, 679–692.

- Stuiver, M., & Reimer, P. J. (1993). Extended database and revised CALIB
 radiocarbon dating program. *Radiocarbon*, 35, 215–230.
- Stuiver, M., Reimer, P. J., & Reimer, R. (2005). CALIB radiocarbon calibration. Http://radiocarbon.pa.qub.ac.uk/calib.
- Swanson, T. W., & Caffee, M. L. (2001). Determination of ³⁶Cl production
 rates derived from the well-dated deglaciation surfaces of Whidbey and
 Fidalgo Islands, Washington. *Quaternary Research*, 56, 366–382.
- ⁸⁵⁴ Uppala, S. M., Kållberg, P., Simmons, A., Andrae, U., Bechtold, V., Fiorino,
- M., Gibson, J., Haseler, J., Hernandez, A., Kelly, G. et al. (2005). The
- ERA-40 re-analysis. Quarterly Journal of the Royal Meteorological Society,
- ⁸⁵⁷ *131*, 2961–3012.