Geological calibration of spallation production rates in the CRONUS-Earth project

Geological Calibration of Spallation Production Rates in the CRONUS-Earth Project

Brian Borchers[a], Shasta Marrero[b], Greg Balco[c], Marc Caffee[d], Brent Goehring[e], Nathaniel Lifton[f], Kunihiko Nishiizumi[g], Fred Phillips[h], Joerg Schaefer[i], John Stone[j]

[a]*New Mexico Tech, Department of Mathematics, Socorro, NM 87801, USA. borchers@nmt.edu*

[b] *Department of Earth and Environmental Science, New Mexico Tech. Currently at University of Edinburgh, School of Geosciences, Edinburgh, EH8 9XP, United Kingdom. Shasta.Marrero@ed.ac.uk*

[c]*Berkeley Geochronology Center, 2455 Ridge Road, Berkeley CA 94709, USA. gbalco@bgc.org*

[d]*Department of Physics, Purdue University, West Lafayette, IN 47907, USA. mcaffee@purdue.edu*

[e]*Department of Earth, Atmospheric, and Planetary Sciences, Purdue University, West Lafayette, IN 47907 USA. bgoehrin@purdue.edu*

[f]*Department of Earth, Atmospheric, and Planetary Sciences, Purdue University, West Lafayette, IN 47907, USA. nlifton@purdue.edu*

[g]*Space Sciences Laboratory, University of California-Berkeley, Berkeley CA 94720, USA. kuni@sunspot.ssl.berkeley.edu*

[h]*Department of Earth and Environmental Science, New Mexico Tech, Socorro, NM 87801, USA. phillips@nmt.edu*

[i]*Department of Earth and Environmental Sciences, Columbia University, Palisades, NY 10964, USA. schaefer@ldeo.columbia.edu*

[j]*Department of Earth and Space Sciences, University of Washington, Seattle, WA 98195, USA. stone@ess.washington.edu*

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-

viously published studies. We considered seven frameworks for elevation and latitude scaling and five commonly used cosmogenic nuclides, ^{3}He, ^{10}Be, ^{14}C, ^{26}Al, and ^{36}Cl. 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

1. Introduction

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

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

rate at sea level and high latitude. The production rate can be integrated to obtain predicted cosmogenic nuclide concentrations for samples at a particular site with known age. The production rate parameter is calibrated by finding the value that best fits measured concentrations from a collection of sites for which independent age measurements are available. In this paper we test the calibration of production rates for five commonly used cosmogenic nuclides using seven different scaling frameworks.

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

In some cases, a cosmogenic nuclide may be produced by spallation reactions involving different elements in a single sample. For example, ^{36}Cl is commonly produced by spallation of Fe, Ti, Ca, and K. In order to model production of ^{36}Cl in a sample it is necessary to know the chemical composition of the sample. Because several of these elements may be present in our calibration samples, it may be necessary to simultaneously estimate multiple production rates. For this study, we have estimated production rates for ^{36}Cl only from Ca and K. We have used previously published values for the spallation production rates of ^{36}Cl from Fe and Ti (Masarik, 2002; Stone, 2005). In practice, production of ^{36}Cl from Fe and Ti is typically small in comparison with production from Ca and K (i.e., production from Ti and Fe is probably no more than 7% and 3.5% of that from Ca by weight, respectively, and Ti 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.

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 the scaling frameworks. The oldest and most widely used of these scaling frameworks is the model of Lal (1991). We use a version of this scaling framework described in Stone (2000) that has been updated to use atmospheric pressure rather than elevation. This scaling framework is denoted by "St." Balco et al. (2008) adapted the St framework further to incorporate a time-dependent correction for long term changes in the magnetic field of the earth. This modified time-dependent version of the St framework is denoted by "Lm." We also consider the scaling frameworks of Desilets et al. (2006), denoted by "De", the framework of Dunai (2001), denoted by "Du", and the framework of Lifton et al. (2005, 2008), denoted by "Li."

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

Note that the scaling frameworks considered in this paper are actually new implementations described in Marrero et al. (2014a) and Lifton et al. (2014b). These new implementations incorporate recent paleomagnetic history reconstructions and are thus not exactly identical to the previously published scaling frameworks. Similarly, in this paper all elevations have been reduced to atmospheric pressures using the ERA-40 reanalysis of Uppala et al. (2005). It is effectively impossible to test a scaling model without reference to a particular paleomagnetic history reconstruction and atmospheric pressure model. Throughout this paper, the phrase "scaling framework" refers to these scaling models together with the particular paleomagnetic history reconstructions used and the ERA-40 reanalysis of atmospheric pressure (Marrero et al., 2014a; Lifton et al., 2014b).

These seven scaling frameworks have been incorporated into a MATLAB program described in Marrero et al. (2014a). This code currently supports five cosmogenic nuclides, namely ^{3}He, ^{10}Be, ^{14}C, ^{26}Al, and ^{36}Cl. The code can be used to predict the concentration, N_{pred}, of a cosmogenic nuclide in a sample given its exposure age. It can also solve for the exposure age cor-

responding 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.

2. Methods

Our statistical model for the calibration of the production rates begins with the assumption we have samples from multiple sites $i = 1, 2, \ldots, m$. 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 for each site. In this paper, all ages are reported in years before 1950. The independently measured site ages, a_i, are uncertain with standard deviations of ϵ_i. In the calibration process we will obtain a fitted age $a_i + \delta_i$ for each site i. For example, if site i has a nominal age of $a_i = 10,000$ years plus or minus $\epsilon_i = 500$ years, and the fitted value is $\delta_i = 1.5$, then the fitted age is $10,750$ years. Since uncertainties in the independent age constraints are sometimes on the order of 5% of the exposure age, while uncertainties in the measured concentrations are sometimes as small as 1%, it would be extremely difficult to fit production rates exactly to nominal independent ages without substantial differences between the measured and predicted concentrations.

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 ^{14}C 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 $j = 1, 2, \ldots, n_i$. Let $n = n_1 + n_2 + \ldots + n_m$ 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 $N_{\text{pred},i,j}(x_{i,j}, a_i + \delta_i, P)$.

We assume that measured concentrations $N_{\text{meas},i,j}$ are unbiased and normally distributed with standard deviations $\sigma_{i,j}$. Under these assumptions we obtain a least squares problem

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

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

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

For the calibrations reported in this paper, the values of a_i, ϵ_i, $x_{i,j}$, and $N_{\text{meas},i,j}$ 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 nuclide in a sample, they report on the internal analytical uncertainty in the concentration measurement. These uncertainties could be used in the calibration. However, there is also considerable variability from batch to batch within a laboratory and between different laboratories that is not reflected in these internal analytical uncertainties. Thus the uncertainties reported with the measured concentrations may overstate the precision of these measurements.

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

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

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 fitted production rates are insensitive to the inclusion or exclusion of any particular calibration site. After finding the best-fitting production rate for a nuclide using the entire primary calibration data set, we construct subsets of the primary calibration data set in which one site at a time is removed from the data set. We then repeat the calibration process using each of these subsets of the data. For example, suppose that we have data from three calibration sites, A, B, and C. in the calibration of the production rate we first fit the production rate using data from the A, B, and C. We then repeat the calibration using data from the subsets (A, B), (A,C), and (B,C).

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 cosmogenic nuclides production by muons and thermal neutrons is significant. The models used for production by these pathways are discussed in Marrero et al. (2014a). These models also involve production rate parameters that can be fit to data. For production of ^{36}Cl by thermal neutrons, a fixed parameter of $P_{f,(0)} = 704$ is used (Marrero, 2012). The production rates for the muon production pathway are separately estimated using a process that will be described in a forthcoming paper. The specific values used for the various scaling frameworks are given in the online appendix. However, these production rates are weakly coupled with the spallation production rates in that estimates of the spallation production rates are used in the calibration of the muon and thermal neutron production rate parameters and vice versa. In practice we have used an iterative approach in which we alternate calibration of the spallation production rates with calibration of the muon production rates until the rates converge to values that are stable to at least four digits.

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 parameter. For the CRONUS-Earth sites, every parameter was measured and documented in the field and lab with photos available as appendices to the papers documenting the sites, in addition to the original sample collection notes. For previously published studies, authors were contacted to gather any information that was not explicitly included in the publication. If missing information could not be obtained the study was not used in this calibration effort.

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

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 ^{26}Al calibration are simply the subset of the full ^{10}Be data set that also had ^{26}Al 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.

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

1. 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.
3. If density was not provided, the rock type was used to estimate a general 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 contacted to confirm the ^{10}Be AMS standard that was used. If necessary, concentrations were renormalized to the standard of Nishiizumi et al. (2007), using the procedure employed by the Balco et al. (2008) calculator. The calculations assume a ^{10}Be half-life of 1.387 Myr (Korschinek et al., 2010; Chmeleff et al., 2010) and an ^{26}Al half life of 7.05 Myr (Nishiizumi, 2004). To ensure consistency between measurements from different AMS laboratories, all ^{10}Be data used in the calibration are normalized to the Nishiizumi 01-5-4 standard with an assumed ^{10}Be/^{9}Be ratio equal to 2.851×10^{-12} (Nishiizumi et al., 2007). This is equivalent to the 07KNSTD normalization of the CRONUS calculator (Balco et al., 2008). Note therefore that production rates derived from this study should only be used with ^{10}Be data normalized to this same standard value. Likewise, all ^{26}Al data used in the calibration are normalized to the ^{26}Al/^{27}Al standard series described in Nishiizumi (2004), and production rates should only be applied to Al-26 data so normalized. Samples for which the analytical standard could not be identified were not used.
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).

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}C	^{36}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.

3.1. Primary Calibration Data Sets

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

Site	Age (yr)	Uncertainty (yr)	^{10}Be	^{26}Al	^{3}He	^{36}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 wave-polished shoreline for ^{10}Be, ^{26}Al, and ^{14}C calibration. Both sites are described in Lifton et al. (2014a). One Be laboratory's set of Promontory Point ^{10}Be 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 ^{10}Be, ^{26}Al, ^{14}C, and ^{36}Cl. One site was removed from the ^{36}Cl calibration due to evidence of possible inheritance. All chlorine samples were mineral separates.

Quelccaya, Peru (PERU). This is original CRONUS-Earth data. Samples 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 ^{14}C sites (ANT, NCHL, WMDV). The ANT samples come from sites in the Transantarctic Mountains and Prince Charles Mountains in Antarctica. They were collected from bedrock surfaces and large erratic boulders beyond the mapped limits of last glacial maximum (LGM) ice advance at each site. The NCHL samples are from Northern Chile. The WMDV samples are from the White Mountains and Death Valley in California. The Chilean and eastern California samples were collected from bedrock outcrops and boulders on alluvial surfaces with geomorphic evidence of long-term surficial stability and antiquity.

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.

Breque, Peru (BRQ). This glacial moraine data was previously published 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 ^{10}Be data from Nishiizumi (1989) (SN) and ^{36}Cl 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, PH8, PH9, PH10, PH11, PH12). These are data previously published in Phillips et al. (1996) and revised in Phillips et al. (2001). This data set contains many sites and landforms including basalt flows and glacial moraines. These are named PH1, PH2, PH3, etc. up to PH12. See Table 4 and the appendix to see specific ages and locations. Some sample sets were removed from the 1996 data set due to new information about the uncertainty in the independent age or other problems with the data set. All chlorine analyses were whole-rock samples.

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

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

Puerto Bandera Moraines, Patagonia (PAT). These data were previously published in Kaplan et al. (2011). Only the Puerto Bandera Moraines samples were included.

Titcomb Basin, USA (WY). These data were previously published in Gosse et al. (1995). Samples were collected from glacial landforms. These data were renormalized to the current ^{10}Be standard of Nishiizumi et al. (2007). Since the finalization of the data set for this paper, additional questions have been raised about the proper normalization of these ^{10}Be measurements (Gosse, 2014). However, removing these samples from the secondary data sets for ^{10}Be would result in a change of less than 2% in the RMSE and would not materially affect the conclusions of this paper.

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

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

Site	Age (yr)	Uncertainty (yr)	^{10}Be	^{26}Al	^{3}He	^{36}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.

4. Results and Discussion

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

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}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}C	^{36}Cl
Sa	5.12×10^{-5}	$\mathbf{7.15 \times 10^{-1}}$	$< 1.00 \times 10^{-14}$	3.70×10^{-9}	$\mathbf{3.66 \times 10^{-1}}$
St	4.61×10^{-4}	$\mathbf{3.57 \times 10^{-1}}$	$< 1.00 \times 10^{-14}$	2.65×10^{-9}	$\mathbf{1.05 \times 10^{-1}}$
Sf	4.11×10^{-4}	$\mathbf{6.25 \times 10^{-1}}$	$< 1.00 \times 10^{-14}$	1.37×10^{-8}	$\mathbf{3.20 \times 10^{-1}}$
Lm	1.00×10^{-3}	$\mathbf{3.98 \times 10^{-1}}$	$< 1.00 \times 10^{-14}$	2.57×10^{-9}	$\mathbf{1.39 \times 10^{-1}}$
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.

4.1. ^{10}Be Spallation Production Rate

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

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

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 ^{10}Be 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 ^{26}Al, ^{3}He, and ^{14}C by neutron spallation in units of atoms per gram of quartz per year. $P_{s,Cl,Ca}$ is the production rate of ^{36}Cl by neutron spallation of Ca in units of atmoms per gram of Ca per year. $P_{s,Cl,K}$ is the production rate of ^{36}Cl by neutron spallation of K in units of atmoms 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 a time out of the computation. The resulting fitted values of $P_{s,Be}$ were 3.83 (leaving out PPT), 3.89 (leaving out SCOT), 3.93 (leaving out NZ), and 4.02 (leaving out PERU). Since the individual sample measurements are precise to 3% or better and averaging over multiple samples further reduces the uncertainty, the differences between these best-fitting production rates cannot easily be explained by random variation in individual sample measurements. This is a further indication of some inconsistency between the sites or an error in the scaling framework.

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

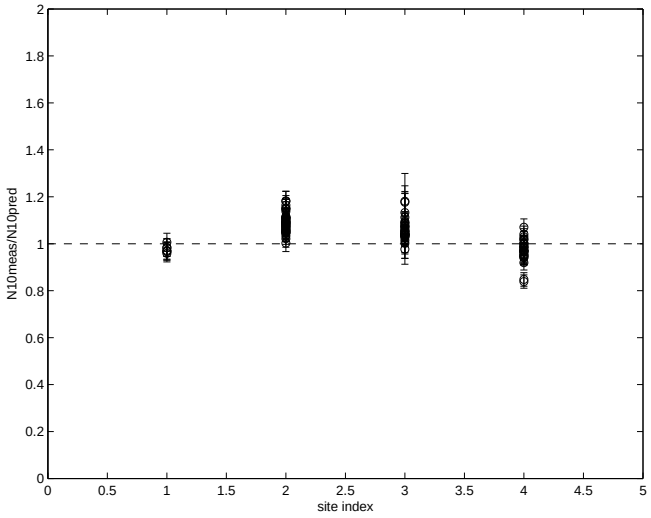

Figure 1: Ratios of measured concentrations to predicted concentrations for ^{10}Be calibration sites. Site 1=NZ, Site 2=PPT, Site 3=SCOT, Site 4=PERU.

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%.

4.2. ^{26}Al Spallation Production Rate

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

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.

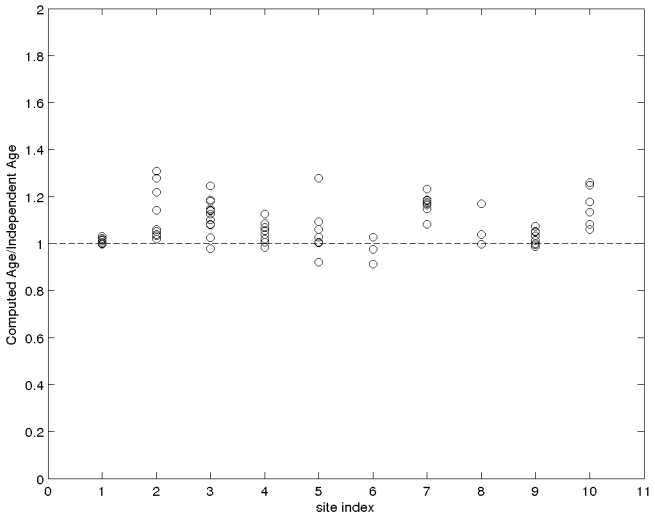

Figure 2: Ratios of Computed Age to Independent Age for secondary 10-Be calibration sites. Site 1=NZM, Site 2=SN, Site 3=NE, Site 4=YDC, Site 5=OL, Site 6=BL, Site 7=WY, Site 8=PUG, Site 9=PAT, Site 10=BRQ.

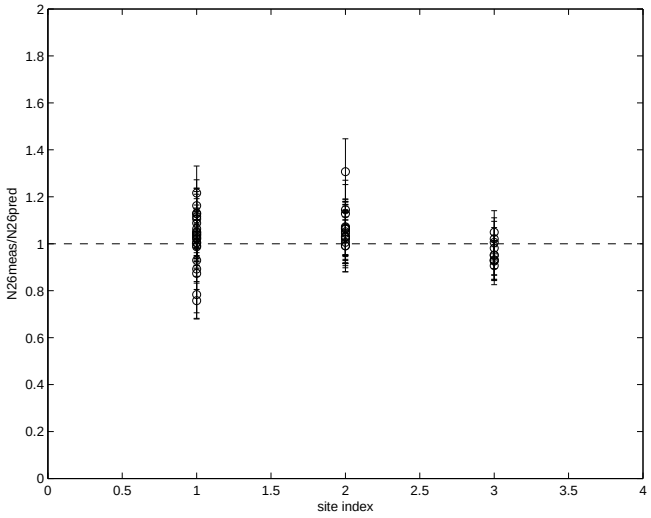

Figure 3: Ratios of measured concentration to predicted concentration for ^{26}Al calibration sites. Site 1=PPT, Site 2=SCOT, Site 3=PERU.

Figure 3 shows the ratios of the measured ^{26}Al concentrations to predicted ^{26}Al 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 ^{10}Be there appear to be systematic offsets from the model at different sites, although the ^{26}Al data is somewhat more scattered. Unlike ^{10}Be, these data do not seem to show a bias toward old ages. The RMSE of the median ages for the sites was 7.1%.

4.3. ^{3}He 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 ^{3}He concentrations to predicted ^{3}He concentrations for samples from the primary calibration sites. Here, there appear to be systematic site offsets as well as scattered measurements

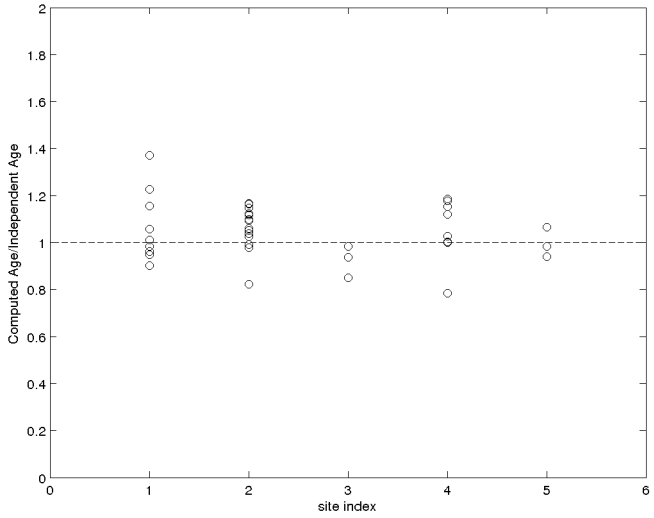

Figure 4: Ratios of Computed Age to Independent Age for secondary ^{26}Al calibration sites. Site 1=SN, Site 2=NE, Site 3=BL, Site 4=WY, Site 5=PUG.

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 ^{3}He 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 site-to-site variations in the data.

^{14}C 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

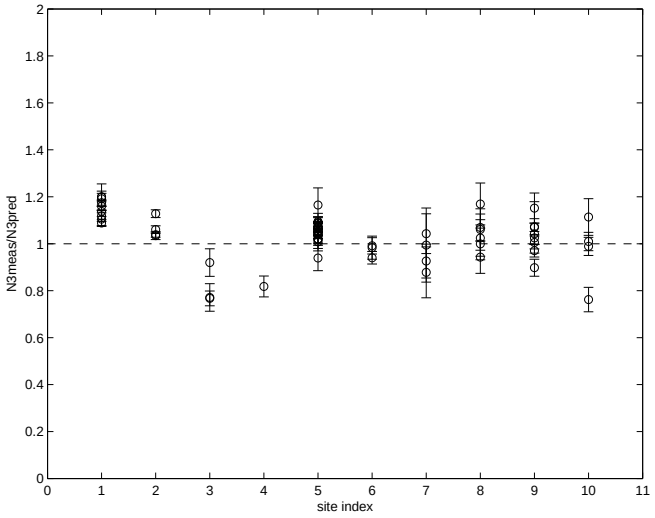

Figure 5: Ratios of measured concentration to predicted concentration for ^{3}He samples from primary calibration sites. Site 1=ARG-O, Site 2=ARG-Y, Site 3=OR-Y, Site 4=OR-O, Site 5=TAB, Site 6=ID, Site 7=ICE-Y, Site 8=ICE=MY, Site 9=ICE-MO, Site 10=ICE-O.

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

4.4. ^{36}Cl Spallation Production Rates

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

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

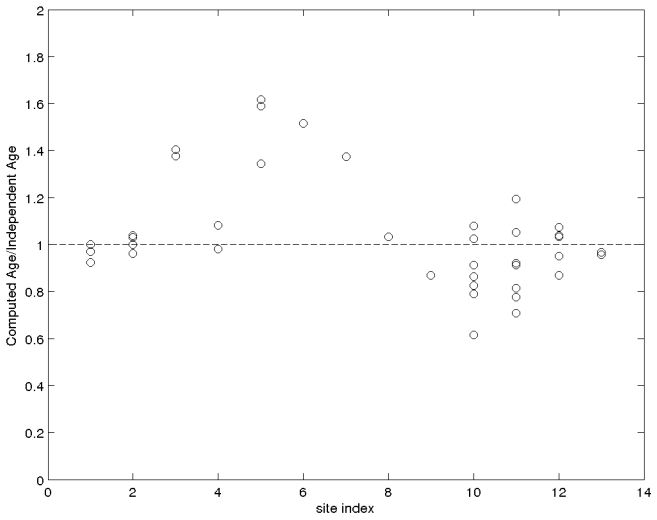

Figure 6: Ratios of Computed Age to Independent Age for secondary ^{3}He calibration sites. Site 1=CAN-Y, Site 2=CAN-O, Site 3=SCLY-O, Site 4=SCLY-Y, Site 5=HAW-M, Site 6=HAW-Y, Site 7=HAW-O, Site 8=CA-Y, Site 9=CA-O, Site 10=YAP, Site 11=SBLK, Site 12=CL, Site 13=LB.

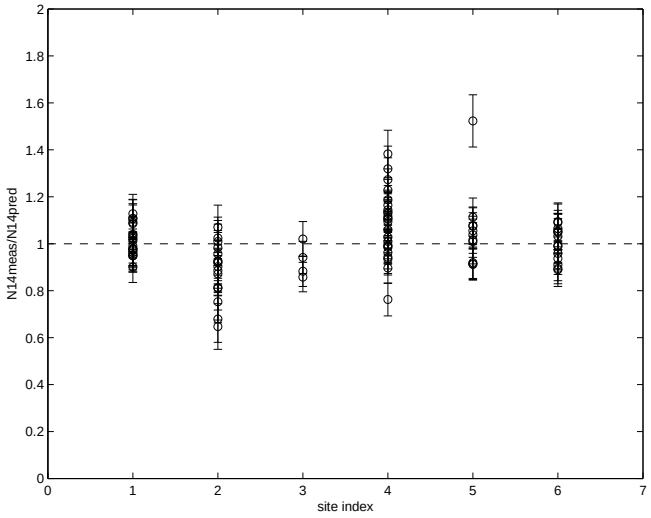

Figure 7: Ratios of measured concentrations to predicted concentrations for ^{14}C calibration sites. Site 1=PPT, Site 2=SCOT, Site 3=NZ, Site 4=WMDV, Site 5=NCHL, Site 6=ANT.

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%.

5. Conclusions

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

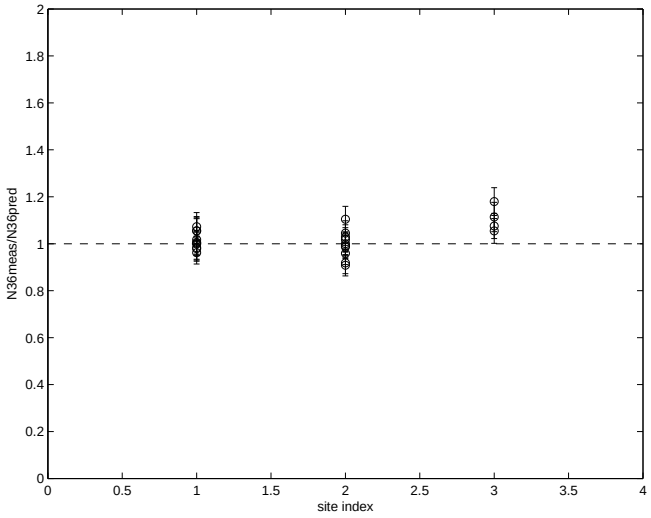

Figure 8: Ratios of measured concentration to predicted concentration for ^{36}Cl calibration sites. Site 1=TAB, Site 2=PERU, Site 3=SCOT.

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

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

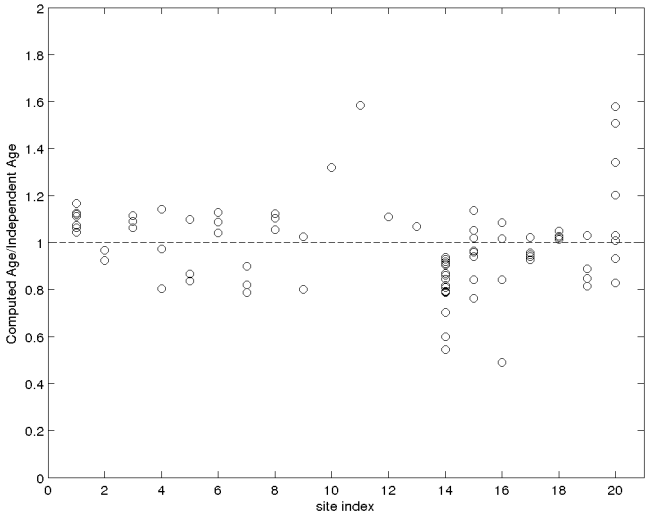

Figure 9: Ratios of Computed Age to Independent Age for quantitative secondary ^{36}Cl calibration sites. Site 1=NE, Site 2=PH11, Site 3=PH 3, Site 4=PH 4, Site 5=PH9, Site 6=PH8, Site 7=PH1, Site 8=PH5, Site 9=PH6, Site 10=PH10, Site 11=PH2, Site 12=PH12, Site 13=PH7, Site 14=BL, Site 15=SNE-K, Site 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 conclusion with the available data.

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

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

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 ^{3}He and ^{14}C, and to improve our understanding of scaling frameworks and site characterization in order to understand the underlying cause of the site-to-site variability.

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.

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.

References

Ackert Jr., R. P., Singer, B. S., Guillou, H., Kaplan, M. R., & Kurz, M. D. (2003). Long-term cosmogenic ^{3}He production rates from ^{40}Ar/^{39}Ar 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). Physics-based 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 late-glacial 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 ^{10}Be and ^{26}Al measurements. *Quaternary Geochronology*, *3*, 174–195.

Blard, P.-H., Pik, R., Lave, J., Marty, B., & Trusdell, F. (2006). Cosmogenic ^{3}He production rates revisited from evidences of grain size dependent release of matrix-sited helium. *Earth and Planetary Science Letters*, *247*, 222–234.

Cerling, T. E., & Craig, H. (1994). Cosmogenic ^{3}He production rates from 39° N to 46° N latitude, western USA and France. *Geochimica et Cosmochimica Acta*, *58*, 249–255.

Chmeleff, J., von Blanckenburg, F., Kossert, K., & Jakob, D. (2010). Determination of the ^{10}be 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.

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

Dunai, T. J., & Wijbrans, J. R. (2000). Long-term cosmogenic ^{3}He production rates (152 ka - 1.35 ma) from ^{40}Ar/^{39}Ar 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 ^{36}Cl from potassium*. Ph.D. thesis The Australian National University, Canberra. 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 ^{3}He 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 ^{10}Be 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 ^{10}Be 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 ^{10}Be 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 inter-comparison 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 ^{10}Be 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 ^{10}be 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 ^{3}He production rates from Holocene lava flows in Oregon,

754 USA, and effects of the earth's magnetic field. *Earth and Planetary Science*
755 *Letters*, *172*, 261–271.

756 Licciardi, J. M., Kurz, M. D., & Curtice, J. M. (2006). Cosmogenic ^{3}He
757 production rates from Holocene lava flows in Iceland. *Earth and Planetary*
758 *Science Letters*, *246*, 251–264.

759 Lifton, N., Caffee, M., Finkel, R., Marrero, S., Nishiizumi, K., Phillips, F. M.,
760 Goehring, B., Gosse, J., Stone, J., Schaefer, J. et al. (2014a). ¡ i¿ in
761 situ¡/i¿ cosmogenic nuclide production rate calibration for the CRONUS-
762 Earth project from Lake Bonneville, Utah, shoreline features. *Quaternary*
763 *Geochronology*, . In Press.

764 Lifton, N., Sato, T., & Dunai, T. J. (2014b). Scaling *in situ* cosmogenic
765 nuclide production rates using analytical approximations to atmospheric
766 cosmic-ray fluxes. *Earth and Planetary Science Letters*, *386*, 149–160.

767 Lifton, N., Smart, D. F., & Shea, M. A. (2008). Scaling time-integrated in
768 situ cosmogenic nuclide production rates using a continuous geomagnetic
769 model. *Earth and Planetary Science Letters*, *268*, 190–201.

770 Lifton, N. A., Bieber, J. W., Clem, J. M., Duldig, M. L., Evenson, P., Hum-
771 ble, J. E., & Pyle, R. (2005). Addressing solar modulation and long-term
772 uncertainties in scaling in situ cosmogenic nuclide production rates. *Earth*
773 *and Planetary Science Letters*, *239*, 140–161.

774 Marquardt, D. W. (1963). An algorithm for least-squares estimation of non-
775 linear parameters. *Journal of the Society for Industrial & Applied Mathe-*
776 *matics*, *11*, 431–441.

777 Marrero, S. (2012). *Calibration of Cosmogenic Chlorine-36*. Ph.D. thesis
778 New Mexico Institute of Mining and Technology.

779 Marrero, S., Phillips, F., Borchers, B., Lifton, N., Aumer, R., & Balco, G.
780 (2014a). Cosmogenic nuclide systematics and the CRONUScalc program.
781 Submitted for Publication in this issue of *Quaternary Geochronology*.

782 Marrero, S. M., Caffee, M. W., Lifton, N., Theirault, B., Phillips, F., Stone,
783 J., Owen, L., L.K., F., Ballantyne, C. K., & Sugden, D. E. (2014b). Cos-
784 mogenic nuclides from the CRONUS-Earth calibration sites in scotland.
785 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 Research: 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 Planetary Science Letters*, *136*, 381–396.

Nishiizumi, K. (2004). Preparation of ^{26}al ams standards. *Nuclear Instruments 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 ^{10}be ams standards. *Nuclear Instruments and Methods in Physics Research Section B: Beam Interactions 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 ^{10}Be and ^{26}Al 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.

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

819 Phillips, F. M., Hinz, M., Marrero, S. M., Nishiizumi, K., & Stone, J. O.
820 (2014). CRONUS-Earth cosmogenic-nuclide calibration sites in the Sierra
821 Nevada, California. Submitted for Publication in this issue of *Quaternary*
822 *Geochronology*.

823 Phillips, F. M., Zreda, M. G., Plummer, M. A., Elmore, D., & Clark, D.
824 (2009). Glacial geology and chronology of Bishop Creek and vicinity, east-
825 ern Sierra Nevada, California. *Geological Society of America Bulletin*, *121*,
826 1013–1033.

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

829 Putnam, A. E., Schaefer, J. M., Barrell, D. J. A., Vandergoes, M., Denton,
830 G. H., Kaplan, M. R., Finkel, R. C., Schwartz, R., Goehring, B. M., &
831 Kelley, S. E. (2010). In situ cosmogenic ^{10}Be production-rate calibration
832 from the Southern Alps, New Zealand. *Quaternary Geochronology*, *5*,
833 392–409.

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

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

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

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

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

847 Stuiver, M., & Reimer, P. J. (1993). Extended database and revised CALIB
848 radiocarbon dating program. *Radiocarbon*, *35*, 215–230.

849 Stuiver, M., Reimer, P. J., & Reimer, R. (2005). CALIB radiocarbon cali-
850 bration. Http://radiocarbon.pa.qub.ac.uk/calib.

851 Swanson, T. W., & Caffee, M. L. (2001). Determination of ^{36}Cl production
852 rates derived from the well-dated deglaciation surfaces of Whidbey and
853 Fidalgo Islands, Washington. *Quaternary Research*, *56*, 366–382.

854 Uppala, S. M., Kållberg, P., Simmons, A., Andrae, U., Bechtold, V., Fiorino,
855 M., Gibson, J., Haseler, J., Hernandez, A., Kelly, G. et al. (2005). The
856 ERA-40 re-analysis. *Quarterly Journal of the Royal Meteorological Society*,
857 *131*, 2961–3012.