FINAL TECHNICAL REPORT

Numerical Testing of a New Procedure to Determine Unbiased Seismicity Rates from Earthquake Catalogs

U.S. Geological Survey Award Number G17AP00059

by

James C. Pechmann, Relu Burlacu, and Walter J. Arabasz
University of Utah
Department of Geology and Geophysics

Gabriel R. Toro
Lettis Consultants International

Period of Award (with no-cost extensions):
June 1, 2017 to May 31, 2021

Principal Investigator
James C. Pechmann
University of Utah
Department of Geology and Geophysics
115 South 1460 East Room 383 FASB
Salt Lake City, UT 84112-0102
Tel: 801-581-3858
Fax: 801-585-5585
E-mail: james.pechmann@utah.edu

Report Date: March 29, 2022

This material is based upon work supported by the U.S. Geological Survey under Grant Number G17AP00059. The views and conclusions contained in this document are those of the authors and should not be interpreted as representing the opinions or policies of the U.S. Geological Survey. Mention of trade names or commercial products does not constitute their endorsement by the U.S. Geological Survey.

ABSTRACT

As part of the Working Group on Utah Earthquake Probabilities (WGUEP) project, Arabasz, Pechmann, and Burlacu (2016; hereafter APB2016) developed a moment magnitude (**M**) earthquake catalog for the Utah region and used it to compute seismicity rates for both this region and for the smaller Wasatch Front region. In the process, APB2016 developed a novel procedure for correcting observed seismicity rates for the overestimation caused by errors in magnitude determination in combination with the exponential frequency-magnitude distribution of earthquakes. The APB2016 methodology is a modification of the EPRI/DOE/NRC (2012) methodology, which involves the assignment of a uniform moment magnitude to each earthquake in a catalog, the assessment of magnitude uncertainties, and the application of bias corrections based on these uncertainties to determine unbiased recurrence parameters. APB2016 modified this methodology to use empirical relations from general orthogonal regression (GOR) instead of least squares regression (LSR) for estimating **M** from other earthquake size measurements. GOR is preferable to LSR for determining linear magnitude conversion relations because LSR tends to underestimate the slope of a linear relation when the uncertainty in the independent variable (x) is not negligible compared to the uncertainty in the dependent variable (y).

The APB2016 procedure for determining unbiased seismicity rates has four key elements: (A) the use of two-step (nested) magnitude conversion relations determined by GOR to estimate **M** and its uncertainty for some earthquakes, (B) the use of an inverse-variance-weighted average of multiple **M** estimates, all from GOR conversion relations, to obtain a best estimate of **M** and its uncertainty, (C) the application of the Tinti and Mulargia (1985) method to correct earthquake counts for magnitude error bias when the **M** estimates are from GOR conversion relations, and (D) the application of the Tinti and Mulargia (1985) method to correct earthquake counts for magnitude error bias when the **M** estimates are inverse-variance-weighted averages from GOR conversion relations. We carried out numerical simulations to test these four key elements of the APB2016 methodology. The results validate elements B, C, and D. For element A, our simulations show that the uncertainty in the **M** values computed from two-step magnitude conversion relations is overestimated by ~17% for the representative case that we analyzed. We are still investigating the cause of this discrepancy. We also conducted some numerical simulations of element C using LSR conversion relations and the related EPRI/DOE/NRC (2012) methodology. The results show that the latter methodology is unsatisfactory for the case that we analyzed. More work is needed to determine the circumstances under which the EPRI/DOE/NRC (2012) methodology may produce unreliable results.

INTRODUCTION

Analysis of earthquake catalog data to determine rates of seismic activity as a function of location, and less commonly time, is an essential part of any seismic hazard study. Such analyses are the only way to obtain information on rates of moderate-size but potentially damaging earthquakes of moment magnitude (**M**) 5.5 to 6.75 ± 0.25 that are below the usual size threshold for surface faulting and therefore not recorded in the near-surface geology. In areas where geologic studies of active faults are minimal or nonexistent, the historical earthquake record may be the primary information available for estimating seismic hazard.

Since the publication of the original papers on the topic by Veneziano and Van Dyck (1985) and Tinti and Mulargia (1985), it has been known that errors in magnitude determinations, in combination with the exponential frequency-magnitude distribution of earthquakes, can cause seismicity rates derived from earthquake catalogs to be overestimated. Figure 1 illustrates the reasons for this overestimation. Assume that the probability distribution for an observed magnitude is symmetical around the true magnitude $\mathbf{M}_{true}$, as illustrated schematically by the two curves shaded in yellow. Consider the number of earthquakes with observed magnitudes $\mathbf{M}_{obs}$ between M 4.0 and 4.5 (black box labeled “bin”). Because of the exponential frequency-magnitude distribution (red curve), there are more earthquakes with $\mathbf{M}_{true} \leq 4.0$ but $\mathbf{M}_{obs} \geq 4.0$ than there are with $\mathbf{M}_{true} \geq 4.0$ but $\mathbf{M}_{obs} \leq 4.0$. Similarly, at the other end of the bin there are more earthquakes with $\mathbf{M}_{true} \leq 4.5$ but $\mathbf{M}_{obs} \geq 4.5$ than there are with $\mathbf{M}_{true} \geq 4.5$ but $\mathbf{M}_{obs} \leq 4.5$. Because there are more earthquakes with magnitudes near the lower end of the bin than near the upper end, the number of earthquakes with $\mathbf{M}_{obs}$ between 4.0 and 4.5 is greater than the number of earthquakes with $\mathbf{M}_{true}$ between 4.0 and 4.5.

Tinti and Mulargia (1985) showed that the seismicity rate bias resulting from magnitude errors is independent of magnitude for a linear Gutenberg-Richter frequency-magnitude relation given by

$$\log_{10} N = a - b\mathbf{M} \quad (1)$$

where N is the cumulative number of earthquakes greater than or equal to moment magnitude **M** and a and b are constants. They also showed that observed seismicity rates could be corrected for this bias by multiplying by the factor

$$N^* = \exp\{-(b \ln(10))^2 \sigma^2/2\} \quad (2)$$

where σ is the standard deviation of the observed magnitudes (Figure 2). This factor can be used to correct observed earthquake counts on an earthquake-by-earthquake basis. As illustrated in Figure 2, it is also possible to correct for the seismicity rate bias by subtracting an adjustment from each observed magnitude to obtain **M*** (EPRI, 1988):

$$M^* = \mathbf{M}_{obs} - (b \ln(10))\, \sigma^2/2 \quad (3)$$

As an example of the size of the size of the N* corrections, σ varies from 0.1 to 0.5 for the moment magnitudes in the central and eastern United States catalog used for the U.S. Geological Survey (USGS) National Seismic Hazard Maps (Mueller, 2019). For the b-value of 1.0 that the USGS adopted for this catalog, equation (2) gives N* corrections of 0.97 to 0.52.

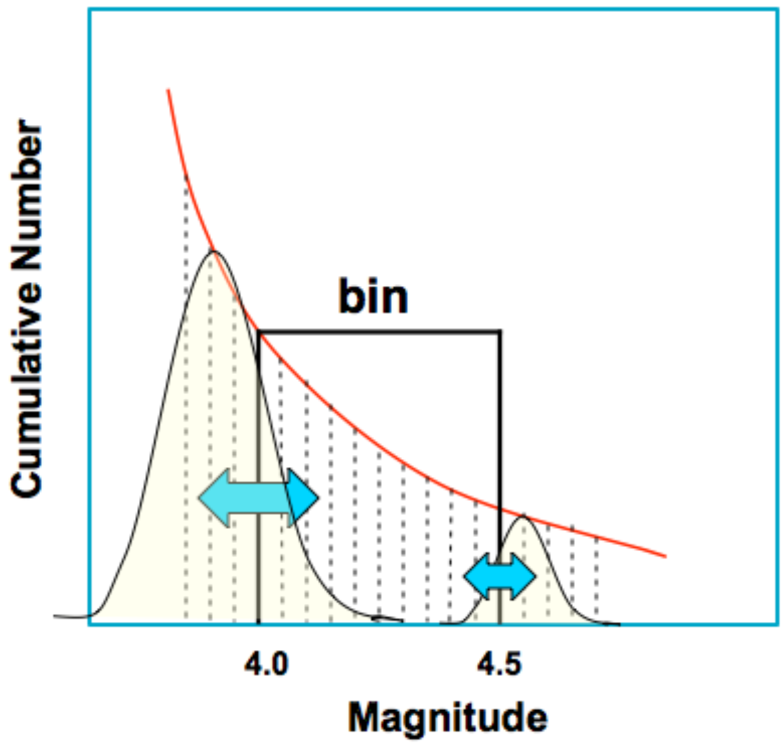

Figure 1. Schematic illustration of how magnitude errors affect observed earthquake counts within a magnitude bin (from Arabasz, 2012). The red curve shows the frequency-magnitude distribution, i.e., the cumulative number of earthquakes of magnitude greater than or equal to the magnitude on the horizontal axis. The bell-shaped curves, shaded in yellow, show probability distributions of observed magnitudes for earthquakes with true magnitudes of 3.9 and 4.55. As illustrated by the blue-green arrows, some earthquakes with true magnitudes inside the magnitude bin 4.0-4.5 (black box) have observed magnitudes outside of the bin, and vice-versa. See the text for further explanation.

Most modern analyses of earthquake recurrence rates are done using measured or estimated moment magnitudes. Because measured **M** values are usually available for only a small fraction of the earthquakes in a catalog, primarily the larger and more recent ones, these analyses commonly rely on empirical relations to convert other size measurements to **M**. The other size measurements are typically other instrumental magnitudes or else intensity data such as maximum Modified Mercalli Intensity (MMI), felt area, and/or the area shaken at or greater than various levels of MMI. The empirical relations to convert these other size measurements to **M**, often called magnitude conversion relations, are generally linear relations obtained by least squares regression (LSR) or general orthogonal regression (GOR). In some cases, other types of magnitudes are simply judged or assumed to be equivalent to **M** without any formal regression analysis.

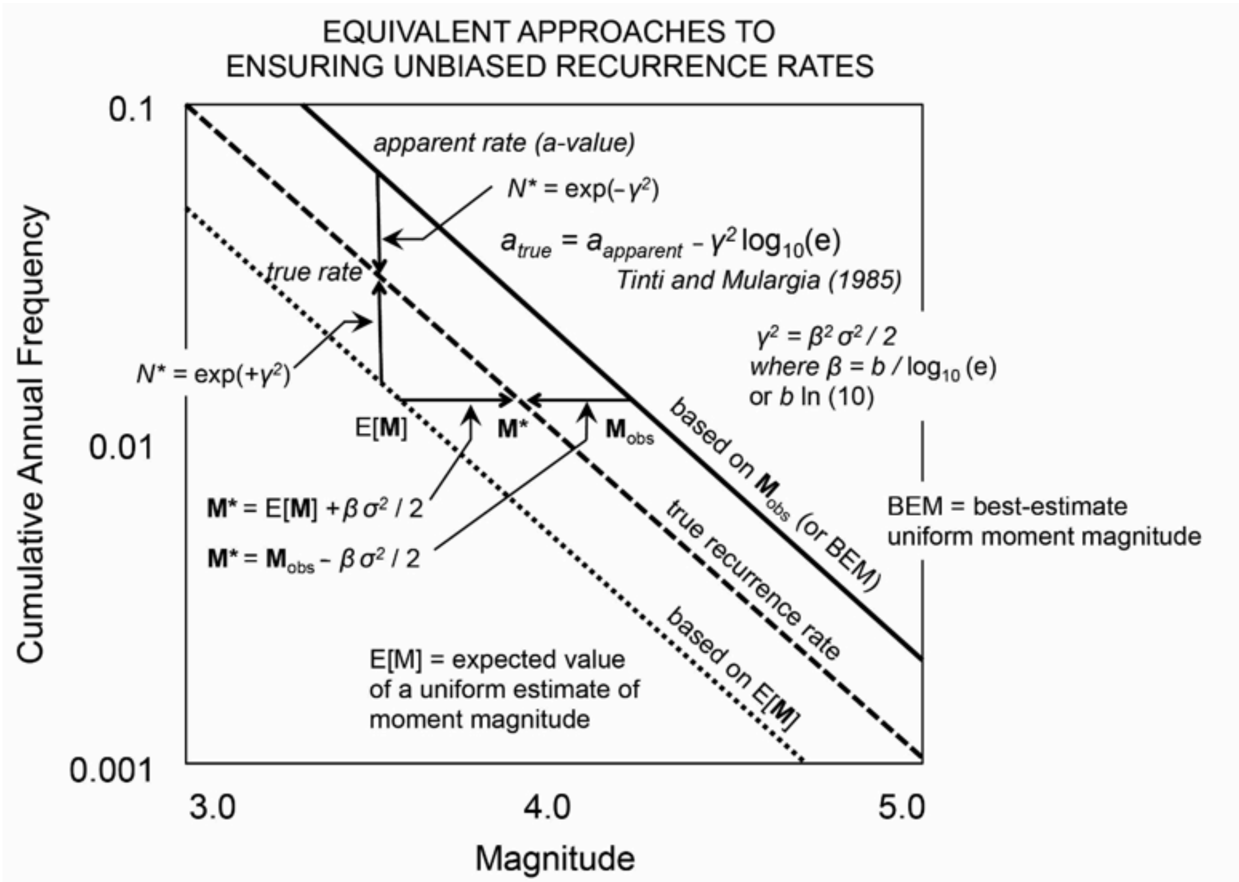

Figure 2. Schematic frequency-magnitude diagram from Arabasz, Pechmann, and Burlacu (2016) showing the true recurrence relation (heavy dashed line), the recurrence relation based on $\mathbf{M}_{obs}$ (solid line), and the recurrence relation based on E[**M**] (dotted line). The diagram shows how "true" (unbiased) recurrence rates can be determined by applying either an "*N** " correction to the rates or by correcting the magnitudes to "*M** ". *N** as defined on this figure is the equivalent count assigned to an individual earthquake.

Over the last two decades, a number of papers have been published showing that GOR is superior to LSR for the purpose of determining magnitude conversion relations (Castellaro et al., 2006; Castellaro and Bormann, 2007; Gasperini et al., 2015). In LSR, a linear relation between a dependent variable and an independent variable, in this case **M** and some other size measurement M_X, respectively, is found by minimizing the sum of the squared differences between the observed and predicted **M** values (vertical distances on an **M** versus M_X plot). As explained by Draper and Smith (1981), Castellaro et al. (2006), and others, a key assumption of LSR is that the uncertainties in the independent variable are negligible relative to the uncertainties in the dependent variable. This assumption is rarely valid for magnitude conversion problems. If this assumption is invalid, then LSR underestimates the true slope of the linear relation (Draper and Smith, 1981, pp. 122-125; Castellaro et al., 2006; Gasperini et al., 2015). To apply GOR, the data are first scaled so that the uncertainties in the dependent and independent variables are approximately equal. Then, a linear relation between the two variables is found by minimizing the sum of the squares of the orthogonal distances between the data points and the line. Based on theoretical considerations and numerical simulations, GOR methods work well in cases where the uncertainties in the independent variable are comparable to or even greater than the uncertainties in the dependent variable, as is usually the case for relations to predict **M** from some other earthquake size measurement.

Unfortunately, when the **M** estimates in an earthquake catalog are estimated from other size measurements via conversion relations, there appears to be considerable confusion in the seismological literature about the proper way to correct the earthquake rates for the bias caused by magnitude errors (Musson, 2012). Musson (2012) points out that in cases where magnitude conversions were used to obtain **M**, some authors have applied corrections for magnitude error bias that reduce the observed seismicity rates whereas other authors have applied corrections that increase the observed seismicity rates. Musson (2012) carried out numerical simulations which showed that an **M** computed from some other size measurement, M_X, using an equation determined from GOR can be treated just like an observed **M** for the purpose of correcting seismicity rates for magnitude error bias. For both $\mathbf{M}_{obs}$ and $\mathbf{M}(M_X)$ from GOR, the **M** errors tend to bias the seismicity rates high and the Tinti and Mulargia correction (equations (2) and (3)) works well. Toro (2013) reached the same conclusion based on his own numerical simulations. In contrast, other simulations by Musson (2012) showed that if $\mathbf{M}(M_X)$ is computed using an equation determined from LSR, the resulting seismicity rates are biased low even in the presence of magnitude errors. Remarkably, he found that adding the magnitude correction in equation (3) to each magnitude instead of subtracting it served to remove most of the bias in the seismicity rates. However, Musson (2012) does not draw any general, practical conclusions from his simulations or provide any clear recommendations on how to properly correct for magnitude error bias in catalogs with **M** values from magnitude conversions.

The EPRI/DOE/NRC (2012) Methodology

The "Central and Eastern United States Seismic Source Characterization for Nuclear Facilities" (CEUS-SSC) report of EPRI/DOE/NRC (2012) presents one of the most comprehensive and well-tested methodologies for constructing a moment magnitude earthquake catalog and then using it to compute bias-corrected earthquake rates. The methodology that they use is an extension of one originally developed by EPRI (1988). The first step in their method is to compute "the expected value of the true magnitude, E[**M**], for every earthquake in the catalog." For most earthquakes in their catalog, they computed E[**M**] from another size measurement using an empirical relation between **M** and the other size measurement. For other earthquakes, their E[**M**] was an inverse-variance weighted average of multiple E[**M**] values computed from such empirical relations. One crucial assumption underlying the E[**M**] calculations that is not stated in the EPRI/DOE/NRC (2012) report is that all of the magnitude conversion relations are from LSR (Robert Youngs, AMEC Foster Wheeler, verbal communication, Sept. 5, 2013). For an observed moment magnitude determined directly from seismic waveform data, $\mathbf{M}_{obs}$, EPRI/DOE/NRC (2012) defines

$$\mathrm{E}[\mathbf{M}] = \mathbf{M}_{\mathrm{obs}} - (b\ln(10))\,\sigma^2[\mathbf{M}|\mathbf{M}_{\mathrm{obs}}] \qquad (4)$$

where $\sigma^2[\mathbf{M}|\mathbf{M}_{obs}]$ is the variance in the observed **M** values. This conversion to E[**M**] is illustrated in Figure 2. It is clear from equation (4) that E[**M**] is not equivalent to the best estimate of **M**. Rather, E[**M**] is an artifical quantity that facilitates the correction of magnitude error bias when the magnitudes in the catalog are a mixture of observed magnitudes and magnitudes computed from LSR conversion relations.

After computing an E[**M**] value for each earthquake in the catalog, an equivalent count N* for each earthquake is computed by dividing by equation (2) with σ equal to the standard deviation of the observed or estimated **M** value for the earthquake. Finally, as described on p. 3-17 of EPRI/DOE/NRC (2012), the rate of earthquakes within a given magnitude interval is obtained by summing the values of N* for earthquakes with E[**M**] values in this magnitude interval and then dividing by the length of the time period for the count. It is important to note that the EPRI/DOE/NRC (2012) methodology was carefully tested using numerical simulations tailored to their actual data sets. However, as this methodology appears to rely on systematic errors in the **M** values calculated from LSR magnitude conversion relations, it is unclear if this methodology can be successfully applied, without restrictions, to any data set.

The APB2016 Methodology

Arabasz, Pechmann, and Burlacu (2016; hereafter APB2016) developed a moment magnitude earthquake catalog for the Utah region and used it to determine unbiased background seismicity rates for this region and for the Wasatch Front region, which is the most heavily populated part of Utah. This work was done under the auspices of the Working Group on Utah Earthquake Probabilities (WGUEP), which used the results for their earthquake probability calculations. In order to accomplish this work, APB2016 essentially developed a new procedure for determining unbiased seismicity rates from earthquake catalogs. This new procedure is patterned after the procedures used by EPRI/DOE/NRC (2012). However, it uses magnitude conversion relations that are derived from GOR instead of LSR. Although none of the steps in the APB2016 procedure are entirely original, it is the only study that we are aware of in which this combination of steps, along with the corresponding rigorous treatment of uncertainties, has been employed to analyze background seismicity rates.

We consider the new APB2016 procedure for determining unbiased seismicity rates to be a significant improvement over the EPRI/DOE/NRC (2012) procedure for the following reasons discussed by APB2016. First, the APB2016 procedure makes use of magnitude conversion relations determined by GOR instead of LSR, which provide more accurate **M** estimates for the reasons discussed above. Second, the APB2016 procedure does not require the calculation and use of the "expected value of the true magnitude," E[**M**], which is required for the EPRI/DOE/NRC (2012) methodology. Reporting of E[**M**] is important for studies that use the EPRI/DOE/NRC (2012) methodology, but can be confusing to users of the catalog because E[**M**] is systematically lower than the best-estimate magnitude. Third, the APB2016 procedure is simpler and more straightforward to use than the EPRI/DOE/NRC (2012) procedure, in part because the observed moment magnitudes can be used directly without modification. There is also no ambiguity about which magnitude should be used to define the boundaries of the magnitude bins for the earthquake counts.

In order for the new APB2016 methodology to gain acceptance within the seismic hazard analysis community and be used by others, the accuracy of this method must be verified by numerical simulations. In this study, we use numerical simulations to test the key elements of the APB2016 procedure for determining unbiased seismicity rates from earthquake catalogs. We also carry out some more limited numerical simulations of the related LSR-based EPRI/DOE/NRC methodology in order to improve our understanding of the conditions necessary for this methodology to produce reliable results.

NUMERICAL SIMULATIONS

In this section we introduce, describe, and present the results of our numerical testing of four key elements of APB2016's analysis of background seismicity rates for the Wasatch Front and Utah regions. The four elements are: (A) the use of two-step (nested) magnitude conversion relations determined by GOR to estimate **M** and its uncertainty for some earthquakes, (B) the use of an inverse-variance-weighted average of multiple **M** estimates, all from GOR conversion relations, to obtain a best estimate of **M** and its uncertainty, (C) the application of the N* method of Tinti and Mulargia (1985) to correct earthquake counts for magnitude error bias when the **M** estimates are from GOR conversion relations, and (D) the application of the N* method of Tinti and Mulargia (1985) to correct earthquake counts for magnitude error bias when the **M** estimates are inverse-weighted averages from GOR conversion relations. The tests of these elements are roughly based on some of their applications to the Utah region analysis in order to ensure that the tests are realistic. We also test a variation of element (C) that has been used in studies in which the **M** estimates are from LSR conversion relations (e.g., EPRI/DOE/NRC, 2012). These studies used the inverse of the N* corrections of Tinti and Mulargia to correct earthquake counts for magnitude error bias.

Synthetic Catalogs for the Magnitude Conversion Relations

All four of our tests require synthetic data sets for determining magnitude conversion relations. We constructed these synthetic data sets to be similar to the one that APB2016 used to develop a conversion relation from University of Utah Seismograph Stations (UUSS) local magnitude (M_L) to **M**. This data set consists of 65 local earthquakes for which both magnitudes are available: five events with $3.1 < \mathbf{M} < 3.5$ and 60 with $3.5 < \mathbf{M} < 6.0$ (Figure 3).

In order to generate synthetic versions of this M_L-**M** data set for testing purposes, we modeled its frequency-**M** distribution as follows. For earthquakes of $3.0 \leq \mathbf{M} < 3.5$, we assumed a uniform probability density function. For earthquakes in the **M** range $3.5 \leq \mathbf{M} < 6.5$, we used a standard truncated exponential distribution (Kramer, 1996, p. 123):

$$N(\mathbf{M}) = \mathrm{N}(\mathbf{M}_{\min}) \, \frac{10^{-b(\mathbf{M} - \mathbf{M}_{\min})} - 10^{-b(\mathbf{M}_{\max} - \mathbf{M}_{\min})}}{1 - 10^{-b(\mathbf{M}_{\max} - \mathbf{M}_{\min})}} \qquad (5)$$

where N(**M**) is the number of earthquakes of magnitude **M** or larger, $\mathbf{M}_{\min}$ is the minimum magnitude of 3.5, and $\mathbf{M}_{\max}$ is the upper bound magnitude which we set to 6.5. To determine the maximum likelihood value of b in (1), we iteratively solved the following equation from McGuire (2004, p. 42):

$$b = 1/\ln(10)\left[\overline{\mathbf{M}} - \mathbf{M}_{\min} + \frac{(\mathbf{M}_{\max} - \mathbf{M}_{\min})10^{-b(\mathbf{M}_{\max} - \mathbf{M}_{\min})}}{1 - 10^{-b(\mathbf{M}_{\max} - \mathbf{M}_{\min})}}\right] \tag{6}$$

where $\overline{\mathbf{M}}$ is the mean magnitude of the data set. As the starting value for b, we used the maximum likelihood solution for the non-truncated version of (1) which is:

$$\mathrm{b} = \frac{1}{\ln(10)(\overline{\mathbf{M}} - \mathbf{M}_{\min})} \tag{7}$$

(McGuire, 2004, p. 42). The resulting frequency-**M** model, with N($\mathbf{M}_{\min}$) = 60 and b = 0.63, is a good fit to the **M** distribution of the APB2016 M_L-**M** data set (Figure 4)

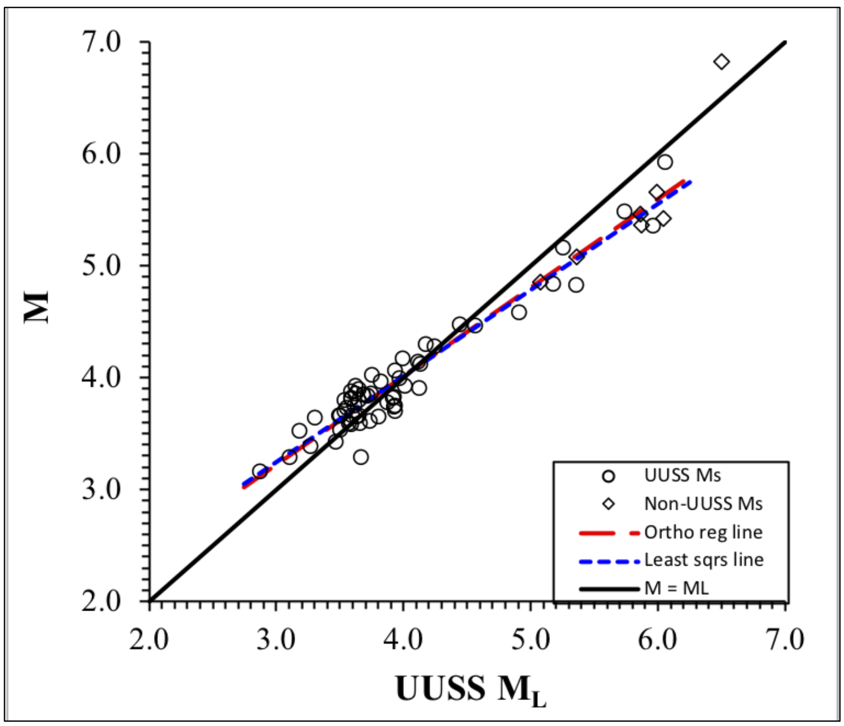

Figure 3. Data used by APB2016 to develop a linear orthogonal regression relation for estimating **M** from UUSS M_L (red dashed line). The least squares regression line (blue dashes) and the **M** = M_L line (solid black) are shown for comparison. Circles and diamonds distinguish UUSS and non-UUSS **M**s, respectively. The regressions excluded the largest earthquake.

To generate a synthetic **M** catalog with a similar distribution, we first generate five **M**s randomly distributed between 3.0 and 3.5. (Here and throughout, **M**s denotes the plural of **M** and should not be confused with M_s, the abbreviation for surface-wave magnitude.) Then, we apply the inversion method described by Zhuang and Touati (2015) to select 60 random **M**s from the truncated exponential distribution that we fit to the APB2016 data set for **M** ≥ 3.5. This inversion method requires the cumulative probability function, which for the truncated exponential distribution of equation (1) is (Kramer, 1996, p. 124):

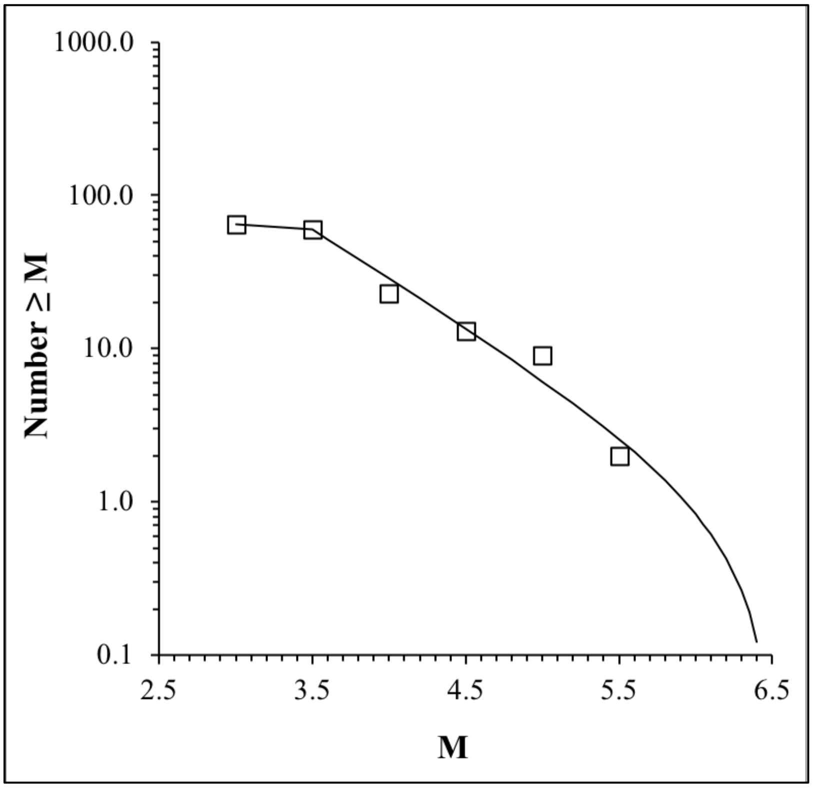

Figure 4. Cumulative frequency-**M** distribution for the APB2016 data set shown in Figure 3 (squares) and our frequency-**M** model for these data (solid line).

$$F_{\mathbf{M}'}(\mathbf{M}) = \mathrm{P}[\mathbf{M}' < \mathbf{M} | \mathbf{M}_{\min} \leq \mathbf{M} \leq \mathbf{M}_{\max}] = \frac{1 - 10^{-b(\mathbf{M} - \mathbf{M}_{\min})}}{1 - 10^{-b(\mathbf{M}_{\max} - \mathbf{M}_{\min})}} \qquad (8)$$

Substituting U for $F_{\boldsymbol{M}'}(\boldsymbol{M})$, where U is a random variable uniformly distributed between 0 and 1, and solving for $\boldsymbol{M}$ gives an equation for $\boldsymbol{M}$ as a function of U:

$$\mathbf{M} = \mathbf{M}_{\min} - \frac{1}{b}\log\left[1 - U\left(1 - 10^{-b(\mathbf{M}_{\max} - \mathbf{M}_{\min})}\right)\right] \qquad (9)$$

From the proof of the inversion method in Zhuang and Touati (2015), a set of random values of **M** computed with equation (9) will have the truncated exponential distribution given by equation (5). We used equation (9) with $b = 0.63$, $\mathbf{M}_{\min} = 3.5$, and $\mathbf{M}_{\max} = 6.5$ to generate 60 values of **M** for the synthetic catalogs, in addition to the five **M** values between 3.0 and 3.5. We consider these synthetic catalogs to be realistic simulations of catalogs in which the **M** values are determined from regional moment tensor inversions, which can typically be done only for earthquakes of **M** ~3.5 and larger (Herrmann et al., 2011; Whidden and Pankow, 2012). These catalogs also minimize regression artifacts which can occur when a realistic synthetic earthquake catalog is simply truncated at some minimum magnitude. Figure 5 shows a plot of **M** versus M_L for one of our synthetic catalogs, along with the GOR and LSR relations.

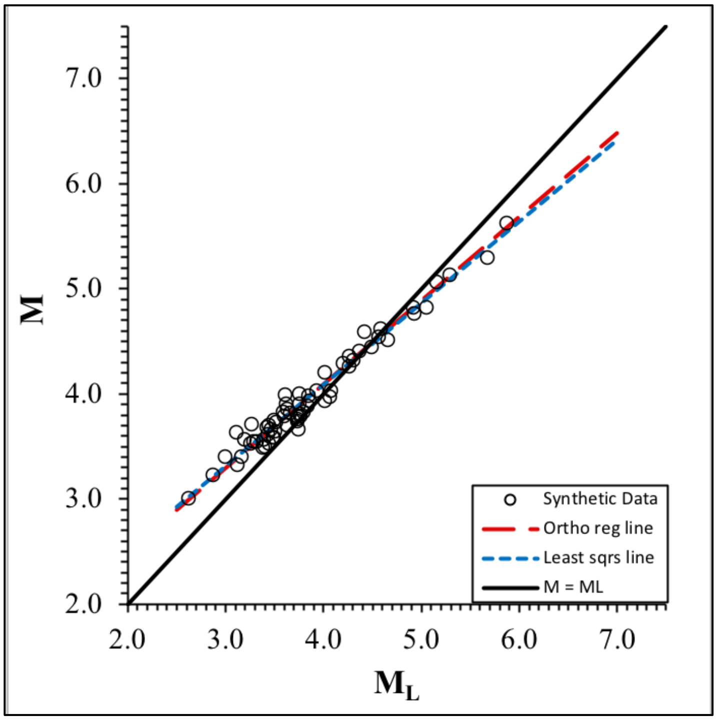

Figure 5. Plots of **M** versus M_L for one of the synthetic catalogs that we generated for testing magnitude conversion relations. The red and blue dashed lines show the orthogonal and least squares regression lines, respectively. The solid black line is the **M** = M_L line.

Test A. Two-step GOR magnitude conversion relations

For many earthquakes in the Utah region catalog, especially the older ones, the only size measurements available are those that APB2016 termed "secondary" size measures, X_2. For these secondary size measurements, a direct empirical relation with **M** cannot be developed due to the lack of earthquakes with both **M** and X_2 values. Consequently, APB2016 used a two-step procedure to estimate **M** and its uncertainty for earthquakes with secondary size measurements only. They used GOR to obtain a linear relation between X_2 and a primary size measurement, X_1, for which they had a GOR conversion relation to **M**. The first step of the two-step procedure was to calculate X_1 from X_2 using the first GOR linear relation $X_1(X_2) = b_1 X_2 + b_2$. The second step was to calculate **M** from $X_1(X_2)$ using the second GOR linear relation, $\mathbf{M} = a_1 X_1(X_2) + a_2$. Assuming that X_1 and X_2 are independent random variables and applying a basic theorem for the linear combination of variances (e.g., Chapman and Schaufele, 1970, p. 131), APB2016 derived the following equation for the variance in **M** estimates derived from X_2 using the two-step procedure:

$$\sigma^2[\mathbf{M}|X_2] = \sigma^2[\mathbf{M}|X_1] + a_1^2 (\sigma^2[X_1|X_2] - \sigma^2 [X_1]) \tag{10}$$

Here, $\sigma^2 [X_1]$ is the variance of values of X_1 used to develop the regression between X_1 and **M** and $\sigma^2[X_1|X_2]$ is the square of the standard error in the predicted value of X_1 from the regression relation between X_1 and X_2. To calculate this standard error, and the standard error in the predicted values from all of their linear regressions (both GOR and LSR), APB2016 used the following equation for the standard error of estimate of y on x, $S_{y,x}$:

$$S_{y,x} = \sqrt{\frac{\sum_i^N (y_i - y_i')^2}{N-2}} \quad (11)$$

where y_i and y_i' are the observed and predicted values of y, respectively, for the ith data point and N is the total number of data points. Finally, $\sigma^2[\mathbf{M}|X_1]$ in equation (10) is the variance in true **M**, given X_1, as specified by equation (3.3.1-8) in EPRI/DOE/NRC (2012):

$$\sigma^2[\mathbf{M}|X_1] = \sigma^2[\mathbf{M}_{\text{obs}}|X_1] - \sigma^2[\mathbf{M}|\mathbf{M}_{\text{obs}}] \quad (12)$$

Here, $\sigma^2[\mathbf{M}_{\text{obs}}|X_1]$ is the square of the standard error in the predicted value of $\mathbf{M}_{\text{obs}}$ from the $\mathbf{M}_{\text{obs}}$-X_1 regression, given by (11), and $\sigma^2[\mathbf{M}|\mathbf{M}_{\text{obs}}]$ is the variance in the observed values of **M**, $\mathbf{M}_{\text{obs}}$, used in the **M**- X_1 regression.

To check the accuracy of the variance estimates for the two-step estimates of M given by equation (10), we did the following.

A1. We created a synthetic data set of 65 true **M** values distributed between **M** 3.0 and 6.5, following the procedure described in the previous section.
A2. For each earthquake in the synthetic data set, we calculated a true M_L (local magnitude) and a true m_b (body wave magnitude) from the true **M** value using assumed linear relations similar to GOR conversion relations CR-1 and CR-9 of APB2016: $\mathbf{M}_{\text{pred}} = 0.8\ M_L + 0.9$ and $M_L = 1.1\ m_b – 0.7$, respectively.
A3. We added random errors to the true values of **M**, M_L, and m_b in order to create a synthetic data set of "observed" magnitudes. The random errors were drawn from normal distributions with zero mean and with standard deviations σ comparable to those of the observed magnitudes that APB2016 used to develop their **M**-M_L and M_L-m_b relations: $\sigma[\mathbf{M}|\mathbf{M}_{\text{obs}}] = \sigma[\mathbf{M}] = 0.05$, $\sigma[M_L] = 0.10$, and $\sigma[m_b] = 0.15$.
A4. We used GOR to determine an empirical linear relation for observed M_L as a function of observed m_b and then calculated $\sigma[M_L|m_b]$ from equation (11).
A5. We used GOR to determine an empirical linear relation for $\mathbf{M}_{\text{obs}}$ as a function of observed M_L and then calculated $\sigma[\mathbf{M}_{\text{obs}}|M_L]$ and $\sigma[\mathbf{M}|M_L]$ from equations (11) and (12), respectively.
A6. We calculated and saved $\sigma[\mathbf{M}|m_b]$, the standard deviation of true **M** given observed m_b, predicted from equation (10).
A7. For each of the 65 earthquakes in the synthetic data set, we computed $M_L(m_b)$ using the observed m_b values and the GOR relation from step A4 and then used the result in the GOR relation from step A5 to compute $\mathbf{M}(M_L(m_b))$.
A8. For each earthquake in the synthetic data set, we computed the difference between the true **M** and the **M** estimated from the two-step procedure in step A7, $\mathbf{M}|m_b = \mathbf{M}(M_L(m_b))$. Then, we computed and saved the mean and the sample standard deviation of these differences. The latter is an estimate of the "observed" value of $\sigma[\mathbf{M}|m_b]$, the standard deviation of true **M** given m_b.
A9. We repeated steps A1-A8 for 500 simulations.

Table 1 presents the mean values of the key Test A outputs for the 500 simulations. The mean difference between true **M** and the **M** from the two-step GOR magnitude conversion relation, **M**($M_L(m_b)$), is essentially zero. However, the mean standard deviation of true **M** given m_b, $\sigma[\mathbf{M}|m_b]$, that we predict from equation (10) is 17.2% larger the mean observed value computed in step A8. In four additional sets of 500 simulations that we ran, the overprediction of $\sigma[\mathbf{M}|m_b]$ ranged from 17.3% to 18.1%. The amount of the overprediction depends on the standard deviations of the magnitudes used in the simulations. For example, if we reduce $\sigma[m_b]$ from 0.15 to 0.10 and leave $\sigma[\mathbf{M}]$ and $\sigma[M_L]$ unchanged, the predicted and observed values of $\sigma[\mathbf{M}|m_b]$ are both reduced but the difference between them increases to ~35%.

We are still working to find the cause of the discrepancy between the predicted and observed values of $\sigma[\mathbf{M}|m_b]$. As a check on the first term of equation (10), which we used to compute the predicted $\sigma[\mathbf{M}|m_b] = \sigma[\mathbf{M}|M_L(m_b)]$, we compared the mean value of $\sigma[\mathbf{M}|M_L]$ predicted from equation (12) in step A5 with the mean observed value. To estimate the mean observed value, we computed the sample standard deviation of the differences between the true **M** and **M**(M_L). The results in the right two columns of Table 1 indicate that the mean predicted and observed values of $\sigma[\mathbf{M}|M_L]$ agree to within 0.1%. Given this result, and a separate successful check of the second term in equation (10), the next step will be to re-examine APB2016's derivation of equation (10).

Table 1. Test A: Mean Results from 500 Simulations

Two-Step GOR Magnitude Conversion: $\mathbf{M}|m_b = \mathbf{M}(M_L(m_b))$		One-Step GOR Magnitude Conversion: $\mathbf{M}|M_L = \mathbf{M}(M_L)$	
$\mathbf{M} - \mathbf{M}(M_L(m_b))$	0.0002	$\mathbf{M} - \mathbf{M}(M_L)$	0.0002
Observed* $\sigma[\mathbf{M}|m_b]$	0.1310	Observed* $\sigma[\mathbf{M}|M_L]$	0.0789
Predicted† $\sigma[\mathbf{M}|m_b]$	0.1535	Predicted† $\sigma[\mathbf{M}|M_L]$	0.0788
$\frac{\text{Predicted } \sigma[\mathbf{M}|m_b]}{\text{Observed } \sigma[\mathbf{M}|m_b]}$	1.172	$\frac{\text{Predicted } \sigma[\mathbf{M}|M_L]}{\text{Observed } \sigma[\mathbf{M}|M_L]}$	0.999

*Based on simulated data sets

†From equations (10) through (12), with $X_1 = M_L$ and $X_2 = m_b$

Test B. Inverse-variance-weighted averages of GOR M estimates

For earthquakes with no observed moment magnitude **M**, both EPRI/DOE/NRC(2012) and APB2016 used inverse-variance weighted averages of **M** estimates computed from other size measures to obtain best estimates of **M** and its uncertainty. However, EPRI/DOE/NRC (2012) used **M** estimates from conversion relations determined with LSR whereas APB2016 used **M** estimates from conversion relations determined with GOR. APB2016 used the following equations to compute a best estimate of **M**, $\mathbf{M}_{best}$, and a corresponding variance, $\sigma^2[\mathbf{M}|\mathbf{X}]$, from a vector **X** of R estimated **M** values:

$$\mathbf{M}_{best} = \sum\nolimits_{i}^{R} \frac{\sigma^2[\mathbf{M}|\boldsymbol{X}]}{\sigma^2[\mathbf{M}|X_i]} \cdot \mathbf{M}|X_i \tag{13}$$

$$\sigma^2[\mathbf{M}|\mathbf{X}] = \frac{1}{\sum_i \frac{1}{\sigma^2\,[\mathbf{M}|X_i]}} \tag{14}$$

where $\mathbf{M}|X_i$ is an estimate of **M** given X_i, and $\sigma^2[\mathbf{M}|X_i]$ is the variance from this estimate. The corresponding equations in EPRI/DOE/NRC (2012), (3.3.1-9) and (3.3.1-10), respectively, are the same except for the replacement of **M** by E[**M**] and the presence of an additional term in (3.3.1-9), given by $+(R-1)\beta\,\sigma^2[\mathbf{M}|\mathbf{X}]$, where $\beta = \mathrm{b}\ln(10)$. APB2016 showed that equations (13) and (14) are equivalent to equations (3.3.1-9) and (3.3.1-10) in EPRI/DOE/NRC (2012), given the definition of E[**M**] and the assumption that an $\mathbf{M}|X_i$ from orthogonal regression can be treated the same an an observed value of **M** for the purposes of magnitude bias correction.

To check the accuracy of combined variance estimates from equation (14), we carried out a numerical test for the case of $\mathbf{M}_{best}$ computed from two GOR $\mathbf{M}|X_i$ estimates.

B1. We created a synthetic data set of 65 true **M** values distributed between **M** 3.0 and 6.5, following the procedure in the section on "Synthetic Catalogs for the Magnitude Conversion Relations."
B2. For each earthquake in the synthetic data set, we calculated a true M_L and a true M_C (coda magnitude) from the true **M** using assumed linear relations similar to GOR conversion relations CR-1 and CR-3 of APB2016: $\mathbf{M}_{pred} = 0.8\ M_L + 0.9$ and $\mathbf{M}_{pred} = 0.9\ M_C + 0.2$, respectively.
B3. We added random errors to the true values of **M**, M_L, and M_C in order to create a synthetic data set of observed magnitudes. The random errors were drawn from normal distributions with zero mean and with standard deviations comparable to those of the observed magnitudes that APB2016 used to develop their **M**-M_L and **M**-M_C relations: $\sigma[\mathbf{M}] = 0.05$, $\sigma[M_L] = 0.10$, and $\sigma[M_C] = 0.10$.
B4. We used GOR to determine an empirical linear relation for $\mathbf{M}_{obs}$ as a function of observed M_L and then calculated $\sigma[\mathbf{M}_{obs}|M_L]$ and $\sigma[\mathbf{M}|M_L]$ from equations (11) and (12), respectively.
B5. We used GOR to determine an empirical linear relation for $\mathbf{M}_{obs}$ as a function of observed M_C and then calculated $\sigma[\mathbf{M}_{obs}|M_C]$ and $\sigma[\mathbf{M}|M_C]$ from equations (11) and (12), respectively.
B6. We calculated and saved $\sigma[\mathbf{M}|\mathbf{X}] = \sigma[\mathbf{M}_{best}]$ using equation (14) with $\sigma[\mathbf{M}|X_1] = \sigma[\mathbf{M}|M_L]$ from step B4 and $\sigma[\mathbf{M}|X_2] = \sigma[\mathbf{M}|M_C]$ from step B5.
B7. From the observed magnitudes for each of the 65 earthquakes in the synthetic data set, we computed $\mathbf{M}|X_1 = \mathbf{M}(M_L)$ using the GOR relation from step B4, $\mathbf{M}|X_2 = \mathbf{M}(M_C)$ using the GOR relation from step B5, and $\mathbf{M}_{best}$ from equation (13).
B8. For each earthquake in the synthetic data set, we computed the difference between the true **M** and $\mathbf{M}_{best}$ from step B7. Then, we computed and saved the mean and the sample standard deviation of these differences. The latter is an estimate of the "observed" value of $\sigma[\mathbf{M}_{best}]$.
B9. We repeated steps B1-B8 for 500 simulations.

Table 2 presents the mean values of the results of Test B for the 500 simulations. The mean difference between true **M** and $\mathbf{M}_{best}$ is nearly zero. There is good agreement between the means of the predicted and observed values of $\sigma[\mathbf{M}_{best}]$, which differ by less than 1%. The right column of Table 2 compares the mean value of $\sigma[\mathbf{M}|M_C]$ predicted from equation (12) in step B5 with the mean observed value. The mean observed value is the sample standard deviation of the difference between **M** and $\mathbf{M}[M_C]$. The means of the predicted and observed $\sigma[\mathbf{M}|M_C]$ also differ by less than 1%.

Table 2. Test B: Mean Results from 500 Simulations

Inverse-Variance-Weighted Average of Two GOR Magnitude Conversions: $\mathbf{M}_{best} = \mathbf{M}(M_L, M_C)$		One-Step GOR Magnitude Conversion: $\mathbf{M}|M_C = \mathbf{M}(M_C)$	
$\mathbf{M} - \mathbf{M}_{best}$	-0.0005	$\mathbf{M} - \mathbf{M}(M_C)$	-0.0005
Observed* $\sigma[\mathbf{M}_{best}]$	0.0591	Observed* $\sigma[\mathbf{M}|M_C]$	0.0886
Predicted† $\sigma[\mathbf{M}_{best}]$	0.0587	Predicted† $\sigma[\mathbf{M}|M_C]$	0.0888
$\frac{\text{Predicted } \sigma[\mathbf{M}_{best}]}{\text{Observed } \sigma[\mathbf{M}_{best}]}$	0.993	$\frac{\text{Predicted } \sigma[\mathbf{M}|M_C]}{\text{Observed } \sigma[\mathbf{M}|M_C]}$	1.002

*Based on simulated data sets and equation (13), with $X_1 = M_L$ and $X_2 = M_C$

†From equation (14)

Test C. Magnitude bias corrections for a catalog with M from a single GOR conversion relation

As mentioned previously, APB2016 corrected their earthquake rate data for the Wasatch Front and Utah regions by using the N^* approach of Tinti and Mulargia (1985). They applied this approach by carrying out the following procedure: (i) compute $N^* = \exp\{-(b \ln(10))^2\sigma^2/2\} \leq 1.0$ for each earthquake in the catalog, (ii) sum the N^* values for the earthquakes in various moment magnitude bins over the time intervals of completeness, and then (iii) divide by the lengths of the completeness intervals to obtain bias-corrected earthquake rates for each **M** bin. Their N^* corrections rely on the premise that **M** values from GOR magnitude conversion relations can be treated as equivalent to observed **M** values for the purposes of magnitude bias corrections. This premise is supported by the results of numerical tests by Musson (2012) and Toro (2013). Note that the N^* corrections for use with observed **M**-values and **M**-values calculated from GOR magnitude conversion relations serve to reduce raw earthquake rates. In contrast, the rate corrections suitable for use with **M** values from LSR magnitude conversion relations, which we call N*' to distinguish it from N^*, serve to increase raw earthquake rates: $N^{*\prime} = \exp\{+(b \ln(10))^2\sigma^2/2\} \geq 1.0$. The validity of the N*' corrections for use with **M** values from LSR magnitude conversions is supported by numerical tests in EPRI/DOE/NRC (2012), at least for their data sets.

Toro (2013) offered an explanation for the difference between the N* corrections for **M** values from GOR and LSR magnitude conversion relations. His explanation is based on the fact that the slope of an LSR line fit to data with noise tends to underestimate the true slope, especially if the uncertainties in the independent variable M_X, e.g. M_L or m_b, are not negligible compared to the uncertainties in the dependent variable **M** (Draper and Smith, 1981, pp. 122-125; Castellaro et al., 2006; Gasperini et al., 2015). As a result, for magnitudes greater than the mean magnitude, least-squares regression lines tend to underpredict **M** (e.g., Figure 5). This underprediction of **M** would lead to underestimation of earthquake rates. Linear magnitude conversion relations derived using GOR do not have this problem. It is unclear if Toro's (2013) explanation holds for the numerical tests in EPRI/DOE/NRC (2012) because, for these tests, the lower magnitude bound appears to be the same for both the LSR fits and the recurrence modeling.

We carried out the following test to check the accuracy of the Tinti and Mulargia (1985) N* method that APB2016 used to correct earthquake counts for magnitude error bias when the **M** estimates are from GOR conversion relations. In the process, we also checked the accuracy of the N*' method used to correct earthquake counts when the **M** estimates are from LSR conversion relations.

C1. We simulated a catalog of 65 true **M** values distributed between **M** 3.0 and 6.5, following the method described in the previous section on "Synthetic Catalogs for the Magnitude Conversion Relations."
C2. We simulated a larger catalog of 10,000 true **M** values that follows a truncated exponential distribution (equation (5)) with a b-value of 1.0, $\mathbf{M}_{min}$ = 2.0, and $\mathbf{M}_{max}$ = 7.0.
C3. For each earthquake in both synthetic catalogs, we calculated a true M_L value from the true **M** value using an assumed linear relation similar to GOR conversion relation CR-1 of APB2016: $\mathbf{M}_{pred} = 0.8\ M_L + 0.9$.
C4. We added random errors to the true values of **M** and M_L in both synthetic catalogs in order to create synthetic observed magnitudes. The random errors added were drawn from normal distributions with zero mean and standard deviations similar to those of the observed magnitudes that APB2016 used to develop their **M**-M_L relations: $\sigma[\mathbf{M}] = 0.05$ and $\sigma[M_L] = 0.10$.
C5. We used GOR and LSR to determine empirical linear relations that predict moment magnitudes $\mathbf{M}_{obs}$ and $\mathbf{M}'_{obs}$, respectively, as a function of observed M_L, using only data from the 65 earthquakes in the smaller catalog.
C6. For the GOR linear relation, we calculated $\sigma[\mathbf{M}_{obs}|M_L]$ and $\sigma[\mathbf{M}|M_L]$ from equations (11) and (12), respectively. Similarly, for the LSR linear relation, we calculated $\sigma[\mathbf{M}'_{obs}|M_L]$ and $\sigma[\mathbf{M}'|M_L]$ from equations (11) and (12), respectively.
C7. From the observed M_L for each earthquake in the larger synthetic M_L catalog, we computed $\mathbf{M}(M_L)$ using the GOR relation from step C5 and $\mathbf{M}'(M_L)$ using the LSR relation from step C5.
C8. Assuming a provisional b-value of 1.0, we computed N* and N*′ from

$$N^* = \exp\{-(b\ln(10))^2\sigma^2[\mathbf{M}|M_L]\ /2\} \quad (15)$$

$$N^{*\prime} = \exp\{+(b\ln(10))^2\sigma^2[\mathbf{M}'|M_L]\ /2\} \quad (16)$$

C9. For the larger $\mathbf{M}(M_L)$ catalog, we counted the number of earthquakes of $\mathbf{M} \geq \mathbf{M}_{min}$, $N(\mathbf{M}_{min})$, for $\mathbf{M}_{min}$ = 2.5, 3.0, 3.5, and 4.0. Similarly, for the larger $\mathbf{M}'(M_L)$ catalog, we counted the number of earthquakes of $\mathbf{M}' \geq \mathbf{M}'_{min}$, $N(\mathbf{M}'_{min})$, for $\mathbf{M}'_{min}$ = 2.5, 3.0, 3.5, and 4.0. We did not use minimum magnitudes of less than 2.5 in order to avoid truncation effects resulting from the original minimum magnitude being **M** 2.0.

C10. We multiplied the N($\mathbf{M}_{min}$) values from step C9 by N* from equation (15) to obtain N*($\mathbf{M}_{min}$), the number of earthquakes of $\mathbf{M} \geq \mathbf{M}_{min}$ corrected for magnitude bias. We also multiplied the N($\mathbf{M}'_{min}$) values from step C9 by N*′ to obtain N*′($\mathbf{M}'_{min}$), the number of earthquakes of $\mathbf{M}' \geq \mathbf{M}'_{min}$ corrected for magnitude bias.
C11. We applied the methodology described in the section on "Synthetic Catalogs for the Magnitude Conversion Relations" to fit truncated exponential distributions to the **M**(M_L) catalog with $\mathbf{M}_{max}$ = 7.0, $\mathbf{M}_{min}$ = 2.5, 3.0, 3.5, and 4.0, and the N*($\mathbf{M}_{min}$) from step C10 (Figure 6). We applied the same methodology to fit truncated exponential distributions to the **M**′(M_L) catalog with $\mathbf{M}'_{max}$ = 7.0, $\mathbf{M}'_{min}$ = 2.5, 3.0, 3.5, and 4.0, and the N*′($\mathbf{M}'_{min}$) values from step C10.
C12. We repeated steps C1-C11 for 500 simulations.

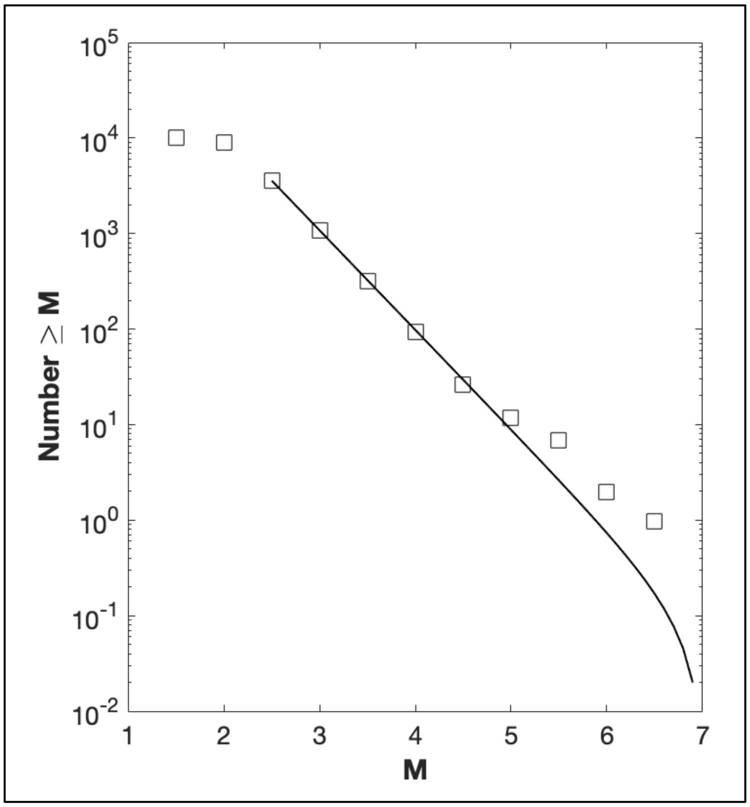

Figure 6. Recurrence plot for a synthetic **M** catalog of 10,000 earthquakes (squares), created by applying an **M**-M_L conversion relation to the M_Ls in this catalog and then applying the Tinti and Mulargia (1985) N* correction for magnitude error bias. The solid line shows the truncated exponential distribution fit to the data: $\mathbf{M}_{min}$ = 2.5, $\mathbf{M}_{max}$ = 7.0, N($\mathbf{M}_{min}$) = 3576, and b = 1.04.

Table 3 shows that the Tinti and Mulargia (1985) magnitude bias corrections work very well for our simulated catalogs of **M** estimates from GOR conversion relations of M_L to **M**. The Tinti and Mulargia (1985) corrections are fairly small for the set of simulations that we did, with an average correction factor of N* = 0.983. Nevertheless, these corrections bring the means of the observed event counts N*($\mathbf{M}_{min}$) to within 0.1% of the true values for the four $\mathbf{M}_{min}$ thresholds that we examined. (See the row in Table 3 with N*($\mathbf{M}_{min}$) / True N($\mathbf{M}_{min}$) in the first column.)

The mean b-values fit to the synthetic catalogs are all very close to the expected value of 1.0, with the differences all well within one mean standard deviation.

Table 3. Test C: Mean Recurrence Parameters from 500 Simulations

Parameter	$\mathbf{M}_{min}$, $\mathbf{M}'_{min}$			
	2.5	**3.0**	**3.5**	**4.0**
With True M				
N($\mathbf{M}_{min}$)	3165.5	999.9	316.4	100.0
b	1.000	1.000	1.000	1.000
With M from a General Orthogonal Regression Conversion Relation				
N($\mathbf{M}_{min}$)	3217.0	1016.0	321.7	101.8
N*($\mathbf{M}_{min}$); N* = 0.983	3163.3	999.0	316.3	100.1
N*($\mathbf{M}_{min}$) / True N($\mathbf{M}_{min}$)	0.999	0.999	1.000	1.001
b	1.000	1.001	1.005	1.018
σ[b]	0.025	0.036	0.057	0.108
With M from a Least Squares Regression Conversion Relation				
N($\mathbf{M}'_{min}$)	3421.6	1059.5	329.1	102.1
N*′ ($\mathbf{M}'_{min}$); N*′ = 1.017	3478.9	1077.3	334.6	103.8
N*′ ($\mathbf{M}'_{min}$) / True N($\mathbf{M}_{min}$)	1.099	1.077	1.058	1.038
b	1.017	1.018	1.023	1.036
σ[b]	0.025	0.036	0.057	0.109

In contrast, the N*′ magnitude bias corrections work poorly for our simulated catalogs of **M** estimates from LSR conversion relations of M_L to **M**. In fact these corrections increase, rather than reduce, the differences between the means of the observed event counts N($\mathbf{M}'_{min}$) and the true event counts. This increase occurs because, contrary to our expectations, the observed event counts N($\mathbf{M}'_{min}$) are higher than the true event counts for all four of the $\mathbf{M}'_{min}$ values in Table 3. The N*′ corrections from equation (16) increase these observed event counts further. After the

N*′ corrections are applied, the mean event counts range from 3.8% to 9.9% above the mean true counts, with the amount of overestimation decreasing as $\mathbf{M}'_{min}$ increases from 2.5 to 4.0. (See the row in Table 3 with N*′ ($\mathbf{M}'_{min}$) / True N($\mathbf{M}_{min}$) in the first column.) The mean b-values fit to the synthetic catalogs range from 1.017 to 1.036, but they are all well within one standard deviation of the expected value of 1.0. More work is needed in order to understand the situations in which the EPRI/DOE/NRC (2012) N*′ magnitude bias corrections do and do not produce accurate results for **M**s estimated from LSR relations.

Test D. Magnitude bias corrections for a catalog with inverse-variance weighted average Ms from GOR conversion relations

Most of the **M** values in the APB2016 Utah region catalog are inverse-variance weighted averages of two or more predicted values of **M** computed from GOR magnitude conversion relations. Our last test investigates the accuracy of the Tinti and Mulargia (1985) N* magnitude bias corrections for a catalog of this sort, thus serving as a realistic overall check on the earthquake recurrence analysis procedure employed by APB2016. We use as our test case a synthetic catalog composed of inverse-variance weighted averages of **M** computed from (i) a GOR **M**-M_L relation and (ii) a GOR **M**-M_C relation. The following are the steps that we followed to conduct this test.

D1. We simulated a catalog of 65 true **M** values and a separate, larger catalog of 10,000 true **M** values following the procedure in steps C1 and C2 of Test C.
D2. For each earthquake in both synthetic catalogs, we calculated a true M_L and a true M_C from the true **M** using assumed linear relations similar to GOR conversion relations CR-1 and CR-3 of APB2016: $\mathbf{M}_{pred} = 0.8\ M_L + 0.9$ and $\mathbf{M}_{pred} = 0.9\ M_C + 0.2$.
D3. We added random errors to the true values of **M**, M_L, and M_C in both synthetic catalogs in order to create synthetic "observed" magnitudes. The random errors added were drawn from normal distributions with zero mean and standard deviations that are roughly twice those of the observed magnitudes that APB2016 used to develop their **M**-M_L and **M**-M_C relations: $\sigma[\mathbf{M}] = 0.10$, $\sigma[M_L] = 0.20$, and $\sigma[M_C] = 0.20$. We used these larger, but still realistic, standard deviations for the observed magnitudes in order to increase the size of the magnitude bias corrections and thereby provide a better test of them.
D4. We used GOR to determine an empirical linear relation for $\mathbf{M}_{obs}$ as a function of observed M_L, using only data from the 65 earthquakes in the smaller catalog. We then calculated $\sigma[\mathbf{M}_{obs}|M_L]$ and $\sigma[\mathbf{M}|M_L]$ from equations (11) and (12), respectively.
D5. We used GOR to determine an empirical linear relation for $\mathbf{M}_{obs}$ as a function of observed M_C, using only data from the 65 earthquakes in the smaller catalog. We then calculated $\sigma[\mathbf{M}_{obs}|M_C]$ and $\sigma[\mathbf{M}|M_C]$ from equations (11) and (12), respectively.
D6. We calculated and saved $\sigma[\mathbf{M}|\mathbf{X}] = \sigma[\mathbf{M}_{best}]$ using equation (14) with $\sigma[\mathbf{M}|X_1] = \sigma[\mathbf{M}|M_L]$ from step D4 and $\sigma[\mathbf{M}|X_2] = \sigma[\mathbf{M}|M_C]$ from step D5.
D7. From the observed magnitudes for each of the 10,000 earthquakes in the larger synthetic catalog, we computed $\mathbf{M}|X_1 = \mathbf{M}(M_L)$ using the GOR relation from step D4, $\mathbf{M}|X_2 = \mathbf{M}(M_C)$ using the GOR relation from step D5, and $\mathbf{M}(M_L, M_C) = \mathbf{M}_{best}$ from equation (13).
D8. Assuming a provisional b-value of 1.0, we computed N* from

$$N^* = \exp\{-(b \ln(10))^2 \sigma^2[\mathbf{M}|\mathbf{X}] / 2\} \quad (17)$$

D9. For the larger $\mathbf{M}_{best}$ catalog, we counted the number of earthquakes of $\mathbf{M}_{best} \geq \mathbf{M}_{min}$, N($\mathbf{M}_{min}$), for $\mathbf{M}_{min}$ = 2.5, 3.0, 3.5, and 4.0.
D10. We multiplied the N($\mathbf{M}_{min}$) values from step D9 by N* from equation (17) to obtain N*($\mathbf{M}_{min}$), the number of earthquakes of $\mathbf{M}_{best} \geq \mathbf{M}_{min}$ corrected for magnitude bias.
D11. We applied the methodology described in the section on "Synthetic Catalogs for the Magnitude Conversion Relations" to fit truncated exponential distributions to the $\mathbf{M}_{best}$ catalog with M_{max} = 7.0, $\mathbf{M}_{min}$ = 2.5, 3.0, 3.5, and 4.0, and the N*($\mathbf{M}_{min}$) values from step D10.
D12. We repeated steps D1-D11 for 500 simulations.

Table 4 compares the recurrence parameters obtained from the simulations of the catalog data processing with the recurrence parameters for the original synthetic catalog. The results validate the use of the Tinti and Mulargia (1985) magnitude bias corrections for a catalog with inverse-variance-weighted mean **M**s from GOR conversion relations. After these corrections, which average a factor of N* = 0.964 for our test, the means of the observed event counts N*($\mathbf{M}_{min}$) are within 1% of the true values for the four $\mathbf{M}_{min}$ thresholds that we examined. (See the row in Table 4 with N*($\mathbf{M}_{min}$) / True N($\mathbf{M}_{min}$) in the first column.) The mean b-values fit to the synthetic catalogs are all within 1.1% of the expected value of 1.0, with the differences all well within one mean standard deviation.

Table 4. Test D: Mean Recurrence Parameters from 500 Simulations

	$\mathbf{M}_{min}$			
Parameter	**2.5**	**3.0**	**3.5**	**4.0**
With True M				
N($\mathbf{M}_{min}$)	3161.1	999.8	315.2	99.4
b	1.000	1.000	1.000	1.000
With Inverse-Variance-Weighted Average M from Two GOR Conversion Relations				
N($\mathbf{M}_{min}$)	3306.4	1041.6	327.4	103.4
N*($\mathbf{M}_{min}$); N* = 0.964	3187.2	1003.9	315.5	99.6
N*($\mathbf{M}_{min}$) / True N($\mathbf{M}_{min}$)	1.008	1.004	1.001	1.002
b	1.001	1.004	1.005	1.011
σ[b]	0.037	0.046	0.065	0.115

CONCLUSIONS

We used synthetic earthquake catalogs to conduct numerical tests of four components of the APB2016 procedure for determining unbiased seismicity rates: (A) estimation of **M** and its uncertainty from a two-step GOR magnitude conversion relation, (B) estimation of **M** and its uncertainty from an inverse-variance-weighted average of **M**s from two GOR magnitude conversion relations, (C) application of the Tinti and Mulargia (1985) magnitude bias corrections to a catalog with **M**s obtained from a GOR conversion relation, and (D) application of the Tinti and Mulargia (1985) magnitude bias corrections to a catalog with inverse-variance-weighted average **M**s from two GOR conversion relations. Our simulations indicate that components (B), (C), and (D) of the APB2016 methodology work well, as expected. For the two-step GOR magnitude conversion relation that we tested in (A), the predicted uncertainty in the **M** values overestimates the true uncertainty by ~17%. The cause of this overestimation is still undetermined. We also carried out some limited tests of the EPRI/DOE/NRC (2012) magnitude bias corrections for a catalog with **M**s obtained from an LSR conversion relation. The results for the case that we analyzed showed that these corrections increased rather than decreased the bias in the earthquake counts.

REFERENCES

Arabasz, W.J. (2012). Issues in considering alternative earthquake catalogs for the NSHMP—Utah catalog example, slide presentation at the National Seismic Hazard Mapping Program Intermountain West Workshop, June 12, 2012 (unpublished).

Arabasz, W.J., J.C. Pechmann, and R. Burlacu (2016). A uniform moment magnitude earthquake catalog and background seismicity rates for the Wasatch Front and surrounding Utah region, Appendix E in Working Group on Utah Earthquake Probabilities, Earthquake probabilities for the Wasatch Front region in Utah, Idaho, and Wyoming, *Utah Geol. Surv., Misc. Publ. 16-3*, 126 pp. and 10 electronic supplements.

Castellaro, S., F. Mulargia, and Y.Y. Kagan (2006). Regression problems for magnitudes, *Geophys. J. Int.* **165,** 913-930.

Castellaro, S., and P. Bormann (2007). Performance of different regression procedures on the magnitude conversion problem, *Bull. Seism. Soc. Am*. **97,** 1167-1175.

Chapman, D.G., and R.A. Schaufele (1970). *Elementary Probability Models and Statistical Inference*, Xerox College Publishing Company, Waltham, Massachusetts, 358 pp.

Draper, N.R., and H. Smith (1981). *Applied Regression Analysis* (Second Edition), John Wiley and Sons, New York, New York, 709 pp.

Electric Power Research Institute (EPRI) (1988). *Seismic Hazard Methodology for the Central and Eastern United States*, 10 volumes, EPRI-NP-4726, Palo Alto, California.

Electric Power Research Institute (EPRI), U.S. Department of Energy (DOE), and U.S. Nuclear Regulatory Commission (NRC) (2012). *Technical Report: Central and Eastern United States Seismic Source Characterization for Nuclear Facilities*, v. 1, EPRI, Palo Alto, California, http://www.ceus-ssc.com/index.htm (last accessed March 2022).

Gasperini, P., B. Lolli, and S. Castellaro (2015). Comparative analysis of regression methods used for seismic magnitude conversions, *Bull. Seism. Soc. Am.* **105,** 1787-1791, doi: 10.1785/012015001.

Herrmann, R.B., H. Benz, and C.J. Ammon (2011). Monitoring the earthquake source process in North America, *Bull. Seism. Soc. Am.* **101,** 2609-2625.

Kramer, S.L. (1996). *Geotechnical Earthquake Engineering*, Prentice Hall, Upper Saddle River, New Jersey, 653 pp.

McGuire, R.K. (2004). *Seismic Hazard and Risk Analysis*, Earthquake Engineering Research Institute, Oakland, California, 221 pp.

Mueller, C.S. (2019). Earthquake catalogs for the USGS National Seismic Hazard Maps, *Seism. Res. Lett.* **90,** 251-261.

Musson, R.M.W. (2012). The effect of magnitude uncertainty on earthquake activity rates, *Bull. Seism. Soc. Am*. **102,** 2771-2775, doi: 10.1785/0120110224.

Pechmann, J.C., and K.M. Whidden (2013). The relation of University of Utah local and coda magnitudes to moment magnitudes: The sequel, *Seism. Res. Lett*. **84,** 299.

Tinti, S., and F. Mulargia (1985). Effects of magnitude uncertainties on estimating the parameters in the Gutenberg-Richter frequency-magnitude law, *Bull. Seism. Soc. Am*. **75,** 1681-1697.

Toro, G. (2013). Uncertainty in magnitude: Numerical experiment with ML→Mw using orthogonal regression, slide presentation during U.S. Geol. Surv. National Seismic Hazard Mapping Project conference call, Sept. 5, 2013 (unpublished).

Veneziano, D., and J. Van Dyck (1985). Analysis of earthquake catalogs for incompleteness and recurrence rates, Appendix A-6 in *Seismic Hazard Methodology for Nuclear Facilities in the Eastern United States*, R.K. McGuire (Project Manager),Vol. 2, Appendix A: Electric Power Research Institute Project No. P101-29, EPRI/Seismicity Owners Group Draft 85-1 (April 30, 1985), A-220–297.

Weichert, D.H. (1980). Estimation of the earthquake recurrence parameters for unequal observation periods for different magnitudes, *Bull. Seism. Soc. Am*. **70,** 1337-1346.

Whidden, K.M., and K.L. Pankow (2012). A catalog of regional moment tensors in Utah from 1998 to 2011, *Seism. Res. Lett.* **83,** 775-783.

Wong, I., W. Lund, C. DuRoss, P. Thomas, W. Arabasz, A. Crone, M. Hylland, N. Luco, S. Olig, J. Pechmann, S. Personius, M. Petersen, D. Schwartz, R. Smith, and S. Bowman (2016). Earthquake probabilities for the Wasatch Front region in Utah, Idaho, and Wyoming, *Utah Geological Survey, Misc. Publ. 16-3*, 251 pp. (excluding appendices).

Zhuang, J., and S. Touati (2015). Stochastic simulation of earthquake catalogs, *Community Online Resource for Statistical Seismicity Analysis*, doi:10.5078/corssa-43806322, available at http://www.corssa.org (last accessed March 2022).

PROJECT DATA

There were no data used in this project. The data used by APB2016 for the Working Group on Utah Earthquake Probabilities study are publicly available at: https://quake.utah.edu/publications/reports/a-uniform-moment-magnitude-earthquake-catalog-and-background-seismicity-rates-for-the-wasatch-front-and-surrounding-utah-region-appendix-e-in-working-group-on-utah-earthquake-probabilities-wguep .

BIBLIOGRAPHY OF PUBLICATIONS RESULTING FROM WORK PERFORMED UNDER THE AWARD

We have not yet published any of the work performed under this award.