Comment on ‘Magnitude conversion problem using general orthogonal regression’ by H. R. Wason, Ranjit Das and M. L. Sharma, (Geophys. J. Int., 190, 1091–1096)

Paolo Gasperini1 and Barbara Lolli2

1Dipartimento di Fisica, Università di Bologna, Italy. E-mail: paolo.gasperini@unibo.it
2Istituto Nazionale di Geofisica e Vulcanologia, Sezione di Bologna, Italy

SUMMARY

The argument proposed by Wason et al. that the conversion of magnitudes from a scale (e.g., $M_s$ or mb) to another (e.g., $M_c$), using the coefficients computed by the general orthogonal regression method (Fuller et al.) is biased if the observed values of the predictor (independent) variable are used in the equation as well as the methodology they suggest to estimate the supposedly true values of the predictor variable are wrong for a number of theoretical and empirical reasons. Hence, we advise against the use of such methodology for magnitude conversions.

Key words: Earthquake source observations; Computational seismology; Statistical seismology.

INTRODUCTION

Wason et al. (2012) argued that the conversion between different magnitude scales, using the coefficients computed by the general orthogonal regression (GOR) method (Fuller et al., 1987), is biased if the observed values of the predictor (independent) variable are used in the equation. They also propose a procedure to compute the supposedly true values of the predictor variable and show that this reduces the standard deviation between observed and predicted magnitudes using an independent data set, not used for computing the regression. Unfortunately, the argument as well as the correction and the testing procedures are wrong owing to incorrect assumptions made by the authors.

First of all, Wason et al. (2012) erroneously assume that the GOR minimizes the orthogonal (minimum) distances between the regressed straight line and the observed data points, whereas the GOR minimizes instead the (Euclidean) distances between the regressed straight line and the observed data points, along paths, which are not necessarily perpendicular to the line itself. According to Fuller (1987, pp. 37) this corresponds to minimize the squared statistical distance that, if the uncertainties of the two variables are uncorrelated ($\sigma_{uu} = 0$), is given by (see eq. 1.3.19 of Fuller 1987)

$$(Y_{\text{obs}} - \hat{\beta}_0 - \hat{\beta}_1 X_{\text{obs}})^2 / \sigma_{ee} + \beta_1^2 / \sigma_{uu},$$

where $\hat{\beta}_1$ and $\hat{\beta}_0$ are the regression coefficient and intercept respectively. $Y_{\text{obs}}, X_{\text{obs}}, \sigma_{ee}$ and $\sigma_{uu}$ are the observed values and variances of response and predictor variables, respectively. As in GOR the variances are assumed constant over the entire data set and only their ratio

$$\eta = \frac{\sigma_{ee}}{\sigma_{uu}}$$

is known, the squared statistical distance (1) becomes

$$(\text{squared statistical distance}) = \frac{(Y_{\text{obs}} - \hat{\beta}_0 - \hat{\beta}_1 X_{\text{obs}})^2}{(1 + \beta_1^2 / \eta) \sigma_{ee}}.$$

We can also write (1) as

$$(\text{squared statistical distance}) = \frac{(Y_{\text{obs}} - \hat{\beta}_0 - \hat{\beta}_1 X_{\text{obs}})^2}{(\eta + \beta_1^2) \sigma_{uu}}.$$

If we assume constant $\sigma_{uu} \neq 0$ in eq. (3), the minimization of the sum of squared statistical distances (3) is equivalent to minimize the sum of squared distances

$$(\text{squared distance}) = \frac{(Y_{\text{obs}} - \hat{\beta}_0 - \hat{\beta}_1 X_{\text{obs}})^2}{1 + \beta_1^2 / \eta}.$$

On the other hand, it is well known, from analytic geometry, that the squared minimum (orthogonal) distance of a point ($X, Y$) from a straight line is

$$d^2 = \frac{(Y - \hat{\beta}_0 - \hat{\beta}_1 X)^2}{1 + \beta_1^2}.$$

Hence, eq. (5) represent the squared minimum distance of the data point ($X_{\text{obs}}, Y_{\text{obs}}$) from the straight line, only if $\eta = 1$ ($\sigma_{uu} = \sigma_{ee}$, which corresponds to the ordinary orthogonal regression).

If instead, for example, $\eta \to \infty$ ($\sigma_{uu}$ is negligible with respect to $\sigma_{ee}$), as in case of direct ordinary least squares (OLS) regression,
eq. (5) becomes
\[
\lim_{\eta \to \infty} \frac{(Y_{\text{obs}} - \beta_0 - \beta_1 X_{\text{obs}})^2}{(1 + \beta_1^2/\eta)} = (Y_{\text{obs}} - \beta_0 - \beta_1 X_{\text{obs}})^2, \tag{7}
\]
that is the squared distance along a vertical line from \(Y_{\text{obs}}\) to the \(Y\) value predicted by the straight line equation at \(X_{\text{obs}}\), which corresponds exactly to the statistics minimized by the direct OLS regression.

If instead \(\eta \to 0 (\sigma_\alpha^2\) is negligible with respect to \(\sigma_\omega^2\), as in case of the inverse OLS regression, we can assume constant \(\sigma_\omega^2\neq0\) in eq. (4). In this case, minimizing the sum of squared statistical distances (4) is equivalent to minimize the sum of squared distances
\[
(\text{squared distance}) = \frac{(Y_{\text{obs}} - \beta_0 - \beta_1 X_{\text{obs}})^2}{\beta_1^2} = \left(\frac{X_{\text{obs}} + \beta_0 - Y_{\text{obs}}}{\beta_1}\right)^2, \tag{8}
\]
that is the distance along a horizontal line from \(X_{\text{obs}}\) to the \(X\) value predicted by the inverse straight line equation at \(Y_{\text{obs}}\), which corresponds exactly to the statistics minimized by the inverse OLS regression.

Moreover, when formulating the procedure to correct the observed abscissa to the supposedly unbiased orthogonally projected value, Wason et al. (2012) do not demonstrate (as they should do) that the empirical linear regression between observed and projected abscissas is a good estimator of the true value in general, and in particular for independent data not used for the regression. By the way, they neither explain how they compute such regressions (using OLS or GOR).

Finally, to proof the presumed superiority of their approach, Wason et al. (2012) erroneously assume, as goodness-of-fit statistics, the standard deviation between observed and predicted values whereas, as shown above, the correct statistics, to compare different orthogonal regression models, is the sum of squared statistical distances or any other normalization of such (e.g. its square root divided by \(N-2\)).

Actually, the best method to minimize the standard deviation between observed and predicted values is using the OLS coefficients but this does not grant that the predicted values are the best in the statistical sense, owing to the presence of errors in the predictor variable. We can note, in particular, that the supposedly corrected regression lines computed by Wason et al. (2012), plotted in their fig. 2, are generally rotated towards the axis with larger variance, similarly to those that can be computed using OLS. Hence, they obviously have smaller standard deviations between observed and predicted values.

Even the use of a different set for testing the goodness-of-fit with respect to that used for computing the regression coefficients does not represent a proof on independent data because, if the two data sets are large enough and come from the same data distribution, the regressed coefficient would reasonably be very similar among the two.

An empirical proof of the inconsistency of the bias suggested by Wason et al. (2012) was given by Castellaro et al. (2006) that, by numerical simulations, demonstrated that it is possible to obtain unbiased estimates of the slope of the frequency–magnitude distribution using \(X_{\text{obs}}\) in conversion equations computed through the GOR.

In conclusion, we have to totally reject the argument proposed by Wason et al. (2012) and particularly the procedure they suggest to correct the presumed bias. We can then confidently, use in magnitude conversions, the observed values, which are certainly affected by errors, but represent the best estimate we have at our disposition of the (unknown) true value of the predictor variable.

This implies that a more reliable conversion from \(M_s\) and \(m_b\) to \(M_w\) can be made, for example, using the regression coefficients computed in a previous paper of the same authors (Das et al. 2011), whereas similar objections have to be raised against another paper of them (Das et al. 2012), using the same assumptions and methodologies criticized in the present paper.

REFERENCES