Chapter 303

Deming Regression

Introduction

Deming regression is a technique for fitting a straight line to two-dimensional data where both variables, X and Y, are measured with error. This is different from simple linear regression where only the response variable, Y, is measured with error. Deming regression is often used for method comparison studies in clinical chemistry to look for systematic differences between two measurement methods.

Deming regression uses paired measurements, (x_i, y_i), measured with errors, ε_i and δ_i, where

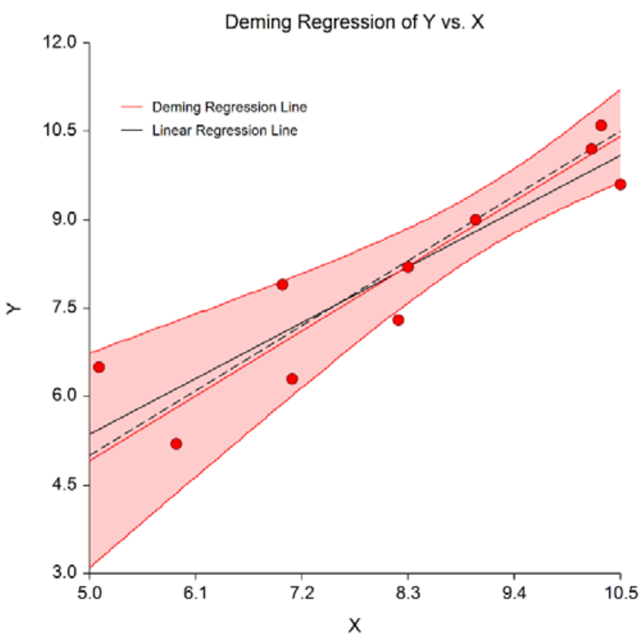
$$x_i = X_i + \varepsilon_i$$

$$y_i = Y_i + \delta_i,$$

to estimate the intercept, β_0, and the slope, β_1, in the equation

$$\hat{Y}_i = \beta_0 + \beta_1 \hat{X}_i.$$

$\hat{X}_i$ and $\hat{Y}_i$ are estimates of the "true" (or expected) values of X_i and Y_i, respectively.



Deming regression assumes that the measurement error ratio, $\lambda = V(\varepsilon_i)/V(\delta_i)$, is constant. The procedure requires the user to either input a known error ratio or to provide multiple measurements from each subject so that the error ratio can be estimated from the data. When $\lambda = 1$, Deming regression gives the same result as orthogonal regression.

Both simple (unweighted) and weighted Deming regression methods are available in this procedure. Regression coefficients and predicted values are calculated using the formulas given in Linnet K. (1990). The standard errors of the regression coefficients and predicted values are calculated using the jackknife leave-one-out method. A pair of tests for the overall hypothesis that $Y = X$ is also computed.

Experimental Design

Typical designs suitable for Deming regression involve *N* paired measurements, (x_i, y_i), $i = 1, \dots, N$, similar to the common input for simple linear regression. The measurement error ratio, λ, is required in Deming regression so the measurement error for each variable must be either entered as a known value or computed from multiple measurements within each subject.

Known Measurement Error

When the measurement error is known, only one response is required for each variable. Typical data of this type are shown in the table below.

Typical Data if Measurement Error is Known

Subject	X	Y
1	7	7.9
2	8.3	8.2
3	10.5	9.6
4	9	9
5	5.1	6.5
6	8.2	7.3
7	10.2	10.2
8	10.3	10.6
9	7.1	6.3
10	5.9	5.2

Unknown Measurement Error

When the measurement errors for each variable are unknown, they must be calculated using multiple measurements (i.e. 2 or more) within each subject. In this case, the paired measurements used in Deming regression, (x_i, y_i), represent the means of the individual replicates. That is

$$x_i = \bar{x}_i = \frac{\sum_{j=1}^{k_{x,i}} x_{i,j}}{k_{x,i}} \text{ and } y_i = \bar{y}_i = \frac{\sum_{j=1}^{k_{y,i}} y_{i,j}}{k_{y,i}}.$$

The *X* and *Y* variables do not necessarily have to have the same number of replicates (i.e. $k_{x,i}$ does not have to be equal to $k_{y,i}$). Typical data of this type are shown in the table below.

Typical Data if Measurement Error is Unknown

Subject	X1	X2	Y1	Y2
1	34	35	31	30
2	72	75	50	46
3	83	85	52	56
4	102	104	60	60
5	122	125	84	84
6	138	136	95	90
7	152	152	101	99
8	176	173	115	116
9	186	182	132	133
10	215	212	146	145

It is possible to perform Deming regression if the measurement error is known for one variable and unknown for the other. In this case, the data would be a hybrid of the two types described above.

Technical Details

The methods and results in this chapter are based on the formulas given in Linnet K. (1990).

Assumptions

Deming regression requires the following assumptions:

1. The measurement errors, ε_i and δ_i, are independent and Normally distributed with expected values of zero and variances $V(\varepsilon_i)$ and $V(\delta_i)$, respectively, which are constant or at least proportional.
2. The measurement error variance ratio $\lambda = V(\varepsilon_i)/V(\delta_i)$ is constant.
3. The subjects are independent of one another and were selected at random from a larger population.

Simple Deming Regression with Constant Errors

Define X_i and Y_i, $i = 1, \ldots, N$, as the "true" (or expected) values for two variables sampled with error to give the observed values x_i and y_i, respectively. We then have

$$x_i = X_i + \varepsilon_i$$

$$y_i = Y_i + \delta_i,$$

where ε_i and δ_i are error terms that are Normally distributed with means equal to zero and variances $V(\varepsilon)$ and $V(\delta)$, respectively.

Known Measurement Error

When the measurement errors for each variable are known, the variances $V(\varepsilon)$ and $V(\delta)$ are calculated from the known values of standard deviation as

$$V(\varepsilon) = SD(\varepsilon)^2 \text{ and } V(\delta) = SD(\delta)^2$$

and from coefficient of variation (COV) values as

$$V(\varepsilon) = (COV(\varepsilon) \times \bar{x})^2 \text{ and } V(\delta) = (COV(\delta) \times \bar{y})^2$$

Unknown Measurement Error

When the measurement errors for each variable are unknown and must be estimated using multiple measurements within each subject, x_i and y_i represent the means of the individual replicates. That is

$$x_i = \bar{x}_i = \frac{\sum_{j=1}^{k_x} x_{i,j}}{k_x}$$

$$y_i = \bar{y}_i = \frac{\sum_{j=1}^{k_y} y_{i,j}}{k_y}.$$

The variances $V(\varepsilon)$ and $V(\delta)$ are calculated as

$$V(\varepsilon) = \frac{\sum_{i=1}^{N}\sum_{j=1}^{k_{x,i}}\left(x_{i,j} - \bar{x}_i\right)^2}{\sum_{i=1}^{N}\left(k_{x,i} - 1\right)}$$

$$V(\delta) = \frac{\sum_{i=1}^{N}\sum_{j=1}^{k_{y,i}}\left(y_{i,j} - \bar{y}_i\right)^2}{\sum_{i=1}^{N}\left(k_{y,i} - 1\right)}$$

Coefficient Estimates

Let λ be the constant ratio of the two error variances such that

$$\lambda = \frac{V(\varepsilon)}{V(\delta)}.$$

Further define $\hat{X}_i$ and $\hat{Y}_i$ as the estimates of the "true" values X_i and Y_i, respectively. Using a series of N paired measurements, (x_i, y_i), the linear relationship is estimated by the equation

$$\hat{Y}_i = \beta_0 + \beta_1 \hat{X}_i,$$

where β_0 is the intercept, and β_0 is the slope. The least-squares approach of Deming regression minimizes the sum of squares

$$SS = \sum_{i=1}^{N} \left[\left(x_i - \hat{X}_i\right)^2 + \lambda\left(y_i - \hat{Y}_i\right)^2 \right].$$

The slope estimate, b_1, is computed as

$$b_1 = \frac{(\lambda q - u) + \sqrt{(u - \lambda q)^2 + 4\lambda p^2}}{2\lambda p}$$

with

$$u = \sum_{i=1}^{N} (x_i - \bar{x})^2,$$

$$q = \sum_{i=1}^{N} (y_i - \bar{y})^2,$$

$$p = \sum_{i=1}^{N} (x_i - \bar{x})(y_i - \bar{y}).$$

The intercept estimate, b_0, is computed as

$$b_0 = \bar{y} - b_1 \bar{x}.$$

Confidence Intervals

$100(1-\alpha)\%$ confidence intervals for the slope and intercept are

$$b_1 \pm t_{1-\alpha/2, N-2} \widehat{SE}(b_1)$$

and

$$b_0 \pm t_{1-\alpha/2, N-2} \widehat{SE}(b_0),$$

respectively. The standard errors are calculated using the jackknife method as described later. Note that some programs (e.g. MedCalc, SAS Macro by Allison Deal) use $N-1$ instead of $N-2$ degrees of freedom. CLSI EP09-A3, Appendix H, indicates that $N-2$ degrees of freedom should be used.

True Value Estimates

Estimated true values are computed using the estimated regression coefficients and λ as

$$\hat{X}_i = x_i + \frac{\lambda b_1 d_i}{\left(1 + \lambda {b_1}^2\right)}$$

$$\hat{Y}_i = y_i - \frac{d_i}{\left(1 + \lambda {b_1}^2\right)}$$

with

$$d_i = y_i - (b_0 + b_1 x_i).$$

where d_i represents the raw y residuals.

Hypothesis Tests

The null hypothesis of identity (i.e., that $Y = X$), is tested by two individual tests. The first tests that the slope is equal to one:

$$H_0\text{: } \beta_1 = 1 \text{ vs. } H_1\text{: } \beta_1 \neq 1,$$

with t-statistic

$$t = \frac{(b_1 - 1)}{\widehat{SE}(b_1)}.$$

The second is a test of location difference, independent of the test of the slope:

$$H_0\text{: } \bar{Y} - \bar{X} = 0 \text{ vs. } H_1\text{: } \bar{Y} - \bar{X} \neq 0,$$

with t-statistic

$$t = \frac{(\bar{y} - \bar{x})}{\widehat{SE}(\bar{y} - \bar{x})}.$$

The standard errors are calculated using the jackknife method as described later. In linear regression, the standard error of regression parameter estimates is estimated using parametric methods, and $N - 2$ degrees of freedom is used. In our research and validation against other software that compute Deming regression, we found that some use $N - 2$ degrees of freedom (e.g., R, "MCR" package) while others use $N - 1$ degrees of freedom (e.g., MedCalc, SAS Macro by Allison Deal). CLSI (Clinical and Laboratory Standards Institute) EP09-A3, Appendix H indicates that $N - 2$ degrees of freedom should be used for the jackknife SE estimates, so $N - 2$ is used as the default in **NCSS**. The choice is given to allow the user to make the final decision as to what degrees of freedom to use and to allow the user to validate calculations in **NCSS** against other software. As the sample size increases, the effect of this choice diminishes.

Weighted Deming Regression with Proportional Errors

With X_i, Y_i, $\hat{X}_i$, $\hat{Y}_i$, x_i, y_i, ε_i, and δ_i all defined as in simple Deming regression, let λ be the ratio of the two error variances that are assumed proportional to the squares of the average of the true values, such that $\lambda = V(\varepsilon_i)/V(\delta_i)$ is constant. Using a series of N paired measurements, (x_i, y_i), the linear relationship is still estimated by the equation

$$\hat{Y}_i = \beta_0 + \beta_1 \hat{X}_i,$$

but the sum of squares minimized becomes

$$SS_w = \sum_{i=1}^{N} \left[w_i \left(x_i - \hat{X}_i\right)^2 + \lambda w_i \left(y_i - \hat{Y}_i\right)^2 \right].$$

The weights, w_i, are

$$w_i = \frac{1}{\left[\frac{X_i + \lambda Y_i}{1 + \lambda}\right]^2}$$

and estimated as

$$\hat{w}_i = \frac{1}{\left[\frac{x_i + \lambda y_i}{1 + \lambda}\right]^2}$$

Note that these weights are different from those outlined in Linnet K. (1990). Linnet K. (1993) suggests that these weights are probably more precise than the simple average.

Coefficient Estimates

The slope estimate, b_1, is computed as

$$b_1 = \frac{(\lambda q_w - u_w) + \sqrt{(u_w - \lambda q_w)^2 + 4\lambda {p_w}^2}}{2\lambda p_w}$$

with

$$u_w = \sum_{i=1}^{N} \widehat{w}_i (x_i - \bar{x}_w)^2,$$

$$q_w = \sum_{i=1}^{N} \widehat{w}_i (y_i - \bar{y}_w)^2,$$

$$p_w = \sum_{i=1}^{N} \widehat{w}_i (x_i - \bar{x}_w)(y_i - \bar{y}_w),$$

$$\bar{x}_w = \sum_{i=1}^{N} \widehat{w}_i x_i / \sum_{i=1}^{N} \widehat{w}_i,$$

$$\bar{y}_w = \sum_{i=1}^{N} \widehat{w}_i y_i / \sum_{i=1}^{N} \widehat{w}_i.$$

The intercept estimate, b_0, is computed as

$$b_0 = \bar{y}_w - b_1 \bar{x}_w.$$

Confidence Intervals

$100(1-\alpha)\%$ confidence intervals for the slope and intercept are computed the same as in the simple case as

$$b_1 \pm t_{1-\alpha/2, N-2} \widehat{SE}(b_1)$$

and

$$b_0 \pm t_{1-\alpha/2, N-2} \widehat{SE}(b_0),$$

respectively. Note that some programs (e.g. MedCalc, SAS Macro by Allison Deal) use $N-1$ instead of $N-2$ degrees of freedom. The standard errors are calculated using the jackknife method as described later.

True Value Estimates

Estimated true values are computed using the estimated regression coefficients and λ as

$$\hat{X}_i = x_i + \frac{\lambda b_1 d_i}{\left(1 + \lambda {b_1}^2\right)}$$

$$\hat{Y}_i = y_i - \frac{d_i}{\left(1 + \lambda {b_1}^2\right)}$$

with

$$d_i = y_i - (b_0 + b_1 x_i).$$

Where d_i represents the raw y residuals.

Iterative Re-weighting to Yield Better Estimates

To obtain better estimates of the true values X_i and Y_i, and therefore better estimates of the weights and coefficients, an iterative procedure is employed. First, initial estimates of the model coefficients are calculated without using weights, or equivalently with all weights equal to one and estimated true values are computed. The process is then iterated by substituting $\hat{X}_i$ for x_i and $\hat{Y}_i$ for y_i and recalculating the weights and regression coefficients. If λ is being estimated from the data, then new estimates for λ are generated with each iteration. The process terminates when the difference between coefficient estimates in subsequent iterations falls below a threshold value.

Hypothesis Tests

The null hypothesis of identity (i.e. that $Y = X$), is tested by two individual tests. The first tests that the slope is equal to one:

$$H_0: \beta_1 = 1 \text{ vs. } H_1: \beta_1 \neq 1,$$

with t-statistic

$$t = \frac{(b_1 - 1)}{\widehat{SE}(b_1)},$$

where b_1 and $\widehat{SE}(b_1)$ are estimated using the weighted method.

The second is a test of location difference, independent of the test of the slope:

$$H_0: \bar{Y} - \bar{X} = 0 \text{ vs. } H_1: \bar{Y} - \bar{X} \neq 0,$$

with t-statistic

$$t = \frac{(\bar{y}_w - \bar{x}_w)}{\widehat{SE}(\bar{y}_w - \bar{x}_w)}.$$

The standard errors are calculated using the jackknife method as described later. In linear regression, the standard error of regression parameter estimates is estimated using parametric methods, and $N - 2$ degrees of freedom is used. In our research and validation against other software that compute Deming regression, we found that some use $N - 2$ degrees of freedom (e.g. R, "MCR" package) while others use $N - 1$ degrees of freedom (e.g. MedCalc, SAS Macro by Allison Deal). CLSI (Clinical and Laboratory Standards Institute) EP09-A3, Appendix H indicates that $N - 2$ degrees of freedom should be used for the jackknife SE estimates, so $N - 2$ is used as the default in **NCSS**. The choice is given to allow the user to make the final decision as to what degrees of freedom to use and to allow the user to validate calculations in **NCSS** against other software. As the sample size increases, the effect of this choice diminishes.

Jackknife Standard Error Estimation

The standard errors (SE) for all estimates (coefficients, predicted values, etc.) in the Deming regression procedure are calculated using the jackknife method. The jackknife is a non-parametric technique and has been shown to perform adequately in the case of Deming regression (see Linnet K. (1990)).

Suppose you have a parameter called θ (θ could be the intercept, slope, a predicted value, the difference in location parameters, etc.) and wish to compute the standard error of $\hat{\theta}$, the estimate of θ. The jackknife estimate for the standard error of $\hat{\theta}$, i.e. $\widehat{SE}(\hat{\theta})$, is computed as follows:

1. Compute $\hat{\theta}$, the estimate of θ using all of the available data.
2. Compute $\hat{\theta}_{-i}$, the estimate of θ based on the data subset that does not contain the pair (x_i, y_i), $i = 1, \dots, N$. Therefore, the end result is that we'll have N estimates of θ, each with a different data pair left out.
3. Compute the i^{th} pseudovariate, $\hat{\theta}_i^*$, $i = 1, \dots, N$, as
$$\hat{\theta}_i^* = N\hat{\theta} - (N - 1)\hat{\theta}_{-i}.$$
4. Compute $\hat{\theta}_{Jackknife}$, the jackknifed estimator as
$$\hat{\theta}_{Jackknife} = \sum_{i=1}^{N} \frac{\hat{\theta}_i^*}{N}.$$

5. Finally, the jackknifed estimate for the variance of $\hat{\theta}$ is calculated as

$$V_J(\hat{\theta}) = \sum_{i=1}^{N} \frac{\left(\hat{\theta}_i^* - \hat{\theta}_{Jackknife}\right)^2}{(N-1)},$$

and the jackknifed estimate for the standard error of $\hat{\theta}$ is calculated as

$$\widehat{SE}(\hat{\theta}) = \sqrt{\frac{V_J(\hat{\theta})}{N}}.$$

Residuals

The residuals can be used to check the Deming regression assumptions. There are 4 different types of residuals that are computed: *X* Residuals, *Y* Residuals, Raw *Y* Residuals, and Optimized Residuals. They are calculated as

$$X\ Residual = d_i^X = x_i - \hat{X}_i = x_i - \left(\hat{Y}_i - b_0\right)/b_1,$$

$$Y\ Residual = d_i^Y = y_i - \hat{Y}_i = y_i - \left(b_0 + b_1\hat{X}_i\right),$$

$$Raw\ Y\ Residual = d_i = y_i - \hat{Y}_i|x_i = y_i - (b_0 + b_1 x_i),$$

$$Optimized\ Residual = d_i^{Opt} = sign(d_i)\sqrt{\widehat{w}_i\left(d_i^X\right)^2 + \widehat{w}_i\lambda\left(d_i^Y\right)^2}.$$

The optimized residuals correspond to the distances from the points (x_i, y_i) to the estimated line at an angle determined by λ, which distances are minimized by Deming regression. The weights, $\widehat{w}_i$, are all equal to one in the case of simple Deming regression.

Predicted Values for Y at a Given X

The predicted value of *Y* for a given value of *X* is

$$\hat{Y} = b_0 + b_1 X,$$

with $100(1-\alpha)\%$ confidence interval

$$\hat{Y} \pm t_{1-\alpha/2,N-2}\widehat{SE}(\hat{Y}).$$

The standard error is calculated using the jackknife method as described above.

Example 1 – Simple Deming Regression with Known Measurement Error

This section presents an example of how to run a Deming regression analysis of the data in the *DemingReg1* dataset. In this example, we will run a Deming regression where the measurement errors are assumed known for both *Y* and *X* with Variance(Y) = 0.008 and Variance(X) = 0.032 (Error ratio: $\lambda = 4.0$). Predicted values of *Y* are wanted at *X* values between 6 and 10. In this example we'll add a simple linear regression line to the plot to see the difference between the two methods.

Setup

To run this example, complete the following steps:

1 Open the DemingReg1 example dataset

- From the File menu of the NCSS Data window, select **Open Example Data**.
- Select **DemingReg1** and click **OK**.

2 Specify the Deming Regression procedure options

- Find and open the **Deming Regression** procedure using the menus or the Procedure Navigator.
- The settings for this example are listed below and are stored in the **Example 1** settings template. To load this template, click **Open Example Template** in the Help Center or File menu.

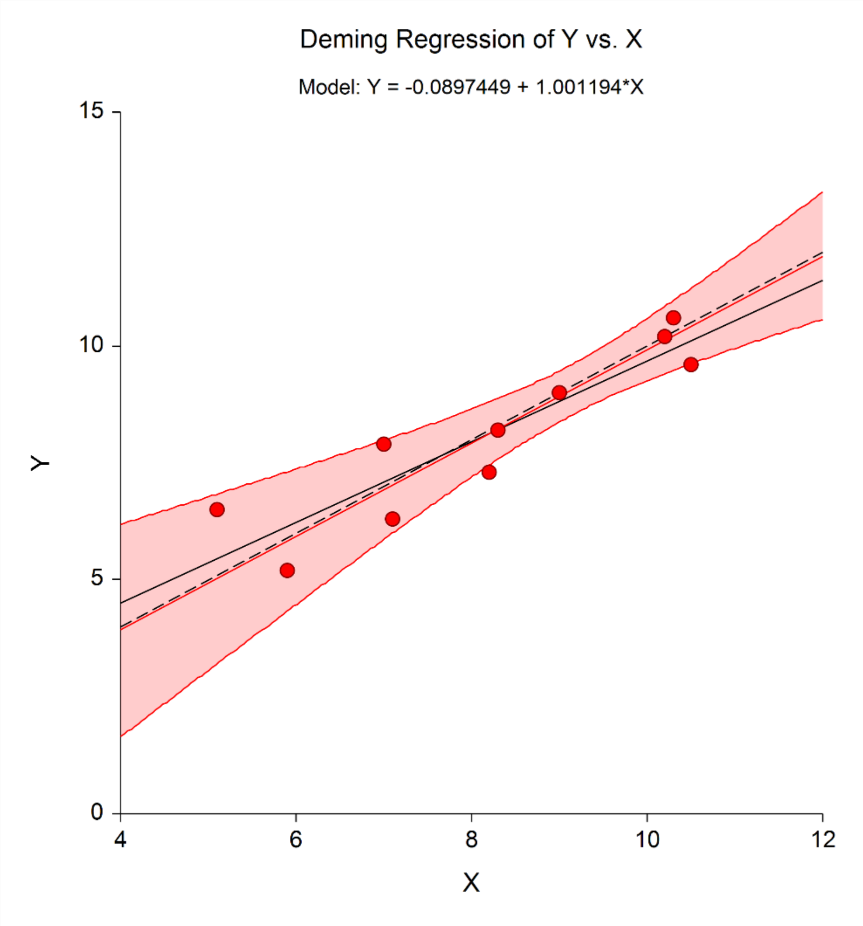
Option	Value
Variables Tab	
Y Measurement Error	**Known SD, Variance, or COV**
Input Type	**Variance**
=	**0.008**
Y Variable	**Y**
X Measurement Error	**Known SD, Variance, or COV**
Input Type	**Variance**
=	**0.032**
X Variable	**X**
Reports Tab	
Data List	**Checked**
Residuals	**Checked**
Predicted Y's at Data X's	**Checked**
Predicted Y's at Specific X's	**Checked**
X's	**6 7 8 9 10**
Plots Tab	
Deming Regression Scatter Plot Format (*Click the Button*)	
More Lines Tab	
Regression Line	**Checked**

3 Run the procedure

- Click the **Run** button to perform the calculations and generate the output.

Deming Regression Scatter Plot

Deming Regression Scatter Plot



The plot shows the data, the Deming regression line, and associated confidence interval bounds in red. The black dashed line represents the 45° $Y = X$ line. The Deming regression line has a slope very nearly equal to one (1.001194 to be exact). The black line is the simple linear regression line. Notice that it is not as steep as the Deming regression line indicating that simple linear regression underestimates the slope when not taking into account the measurement error.

Run Summary Report

Run Summary

Item	Value	Item	Value
Y Measurement Error	Known	Rows Processed	10
Y Variable	Y	Rows with Missing Values	0
X Measurement Error	Known	Rows Used in Estimation	10
X Variable	X		
		Error Variance Ratio (λ)	4.00000
Deming Regression Type	Simple	Y Error Variance	0.00800
SE Estimation	Jackknife	X Error Variance	0.03200
Jackknife DF	N-2	Y Coefficient of Variation	0.01107
Intercept	-0.08974	X Coefficient of Variation	0.02192
Slope	1.00119		

This report gives a summary of the input and various descriptive measures about the Deming regression.

Descriptive Statistics Report

Descriptive Statistics

Variable(s)	N	Mean	Error SD	Error Variance	Error COV
Y: Y	10	8.08	0.08944	0.00800	0.01107
X: X	10	8.16	0.17889	0.03200	0.02192

Error Variance Ratio (λ) = 4.00000

This report gives descriptive statistics about the variables used in the regression.

Regression Coefficient Estimation Report

Regression Coefficient Estimation

Parameter	Regression Coefficient b(i)	Jackknife Standard Error SE[b(i)]	DF	Lower 95% Conf. Limit of β(i)	Upper 95% Conf. Limit of β(i)
Intercept	-0.08974	1.72199	8	-4.06065	3.88117
Slope (X)	1.00119	0.18718	8	0.56956	1.43283

The T-value used to calculate the confidence limits was 2.30600, with N-2 = 8 degrees of freedom.

Estimated Model
Y = -0.0897448990070444 + 1.00119422781949 * X

This section reports the Deming regression coefficients, along with their standard errors, degrees of freedom, and their confidence limits. The standard errors are calculated using the jackknife method.

The Estimated Model section displays an equation with coefficients that have all available decimals displayed (i.e. full-precision).

Hypothesis Test of Y = X Report

Hypothesis Test of Y = X

Test	Parameter Estimate	Jackknife Standard Error	DF	T-Statistic	Prob Level	Reject H0 at α = 0.025?
H0: Slope = 1 vs. H1: Slope ≠ 1	1.00119	0.18718	8	0.00638	0.99507	No
H0: Mean Y - Mean X = 0 vs. H1: Mean Y - Mean X ≠ 0	-0.08000	0.24846	8	-0.32198	0.75572	No

The section reports the pair of hypothesis tests for the overall hypothesis that $Y = X$. As expected based on the Deming regression plot, there is no evidence to suggest that the line does not follow the 45° identity line. Each hypothesis is tested at the Bonferroni adjusted value of $\alpha = 0.025$ because two tests are conducted simultaneously. The standard errors are calculated using the jackknife method.

Data List Report

Data List

Row	X	Y	Estimate of True X (Xhat)	Estimate of True Y (Yhat)	Predicted Y (Yhat|X)
1	7	7.9	7.78455	7.70410	6.91861
2	8.3	8.2	8.28388	8.20403	8.22017
3	10.5	9.6	9.84224	9.76424	10.42279
4	9	9	9.06315	8.98423	8.92100
5	5.1	6.5	6.28607	6.20384	5.01635
6	8.2	7.3	7.54443	7.46370	8.12005
7	10.2	10.2	10.26201	10.18452	10.12244
8	10.3	10.6	10.60174	10.52466	10.22256
9	7.1	6.3	6.52542	6.44347	7.01873
10	5.9	5.2	5.40651	5.32322	5.81730

This section reports the estimated and predicted values for each of the input data points. This report may be quite lengthy if you have a large dataset. You can save these values back to the database using the Storage tab on the procedure window.

Residuals Report

Residuals

Row	X	Y	X Residual (X-Xhat)	Y Residual (Y-Yhat)	Raw Y Residual (Y-Yhat|X)	Optimized Residual
1	7	7.9	-0.78455	0.19590	0.98139	0.87694
2	8.3	8.2	0.01612	-0.00403	-0.02017	-0.01802
3	10.5	9.6	0.65776	-0.16424	-0.82279	-0.73523
4	9	9	-0.06315	0.01577	0.07900	0.07059
5	5.1	6.5	-1.18607	0.29616	1.48365	1.32575
6	8.2	7.3	0.65557	-0.16370	-0.82005	-0.73277
7	10.2	10.2	-0.06201	0.01548	0.07756	0.06931
8	10.3	10.6	-0.30174	0.07534	0.37744	0.33727
9	7.1	6.3	0.57458	-0.14347	-0.71873	-0.64224
10	5.9	5.2	0.49349	-0.12322	-0.61730	-0.55160

This section reports the residuals for each of the input data points. There are 4 different types of residuals that are reported. This report may be quite lengthy if you have a large dataset. You can save these values back to the database using the Storage tab on the procedure window.

Predicted Y's at Data X's Report

Predicted Y's at Data X's

Row	X	Predicted Y (Yhat|X)	Jackknife Standard Error SE[Yhat|X]	Lower 95% Conf. Limit of Y|X	Upper 95% Conf. Limit of Y|X
1	7	6.91861	0.46020	5.85739	7.97984
2	8.3	8.22017	0.28011	7.57424	8.86610
3	10.5	10.42279	0.35015	9.61535	11.23024
4	9	8.92100	0.23605	8.37667	9.46533
5	5.1	5.01635	0.78719	3.20109	6.83160
6	8.2	8.12005	0.29071	7.44967	8.79043
7	10.2	10.12244	0.31083	9.40566	10.83921
8	10.3	10.22256	0.32338	9.47683	10.96828
9	7.1	7.01873	0.44421	5.99437	8.04309
10	5.9	5.81730	0.64583	4.32802	7.30658

The T-value used to calculate the confidence limits was 2.30600, with N-2 = 8 degrees of freedom.

This section reports the predicted values of *Y* for each of the *X* data values, along with their standard errors and confidence limits. The standard errors are calculated using the jackknife method. This report may be quite lengthy if you have a large dataset. You can save these values back to the database using the Storage tab on the procedure window.

Predicted Y's at Specific X's Report

Predicted Y's at Specific X's

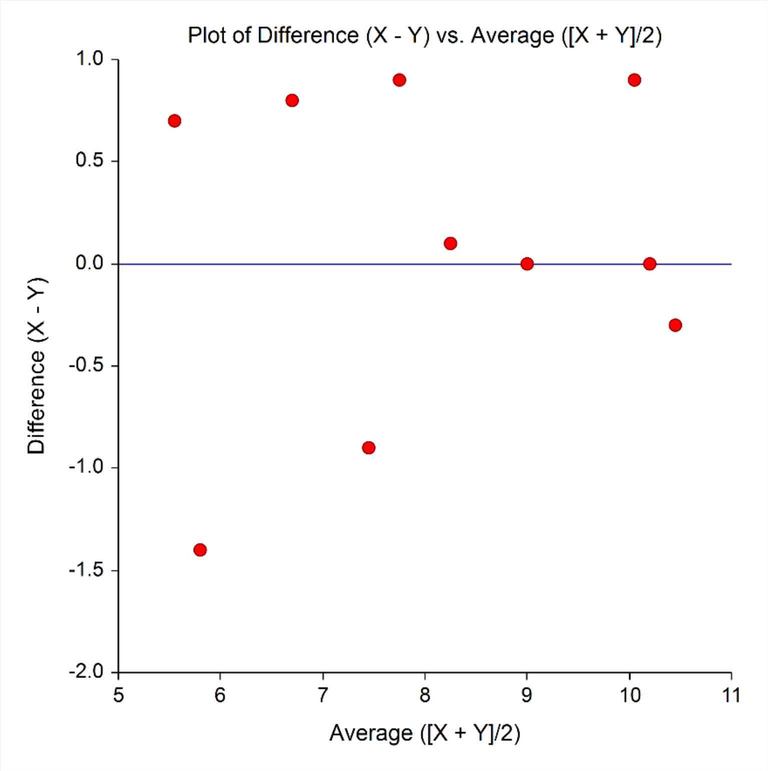
X	Predicted Y (Yhat|X)	Jackknife Standard Error SE[Yhat|X]	Lower 95% Conf. Limit of Y|X	Upper 95% Conf. Limit of Y|X
6	5.91742	0.62843	4.46826	7.36658
7	6.91861	0.46020	5.85739	7.97984
8	7.91981	0.31419	7.19528	8.64434
9	8.92100	0.23605	8.37667	9.46533
10	9.92220	0.28774	9.25868	10.58572

The T-value used to calculate the confidence limits was 2.30600, with N-2 = 8 degrees of freedom.

This section reports the predicted values of *Y* for each of the *X* values specified on the Reports tab of the procedure window, along with their standard errors and confidence limits. The standard errors are calculated using the jackknife method. This report may be quite lengthy if you have a large dataset.

Difference vs. Average Plot

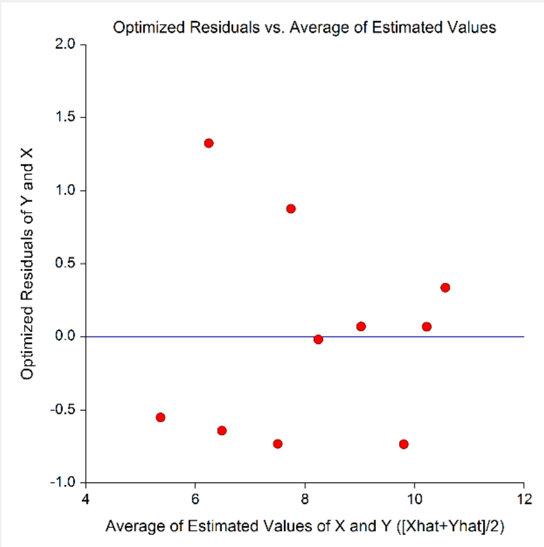
Difference vs. Average Plot

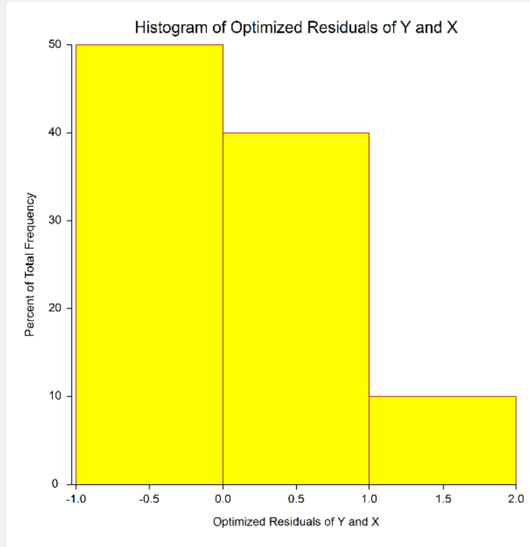


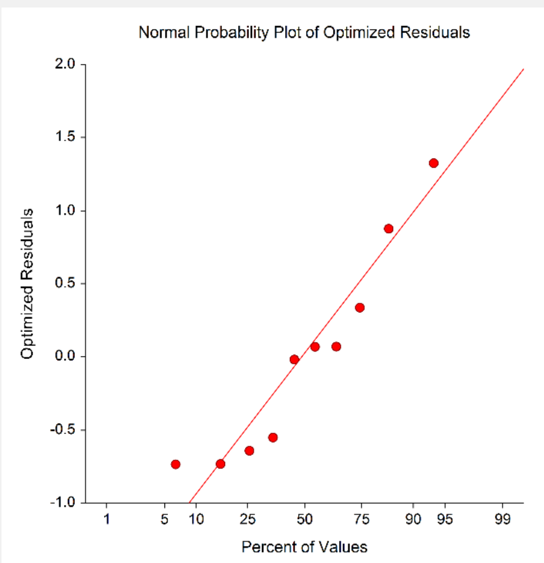
This plot shows the differences between *X* and *Y* for each data pair versus their averages. Use this plot to help determine whether simple Deming regression is sufficient or whether you should use the weighted approach. When the scatter is constant, like it is here, then simple Deming regression is appropriate. If there is a cone-shaped pattern across the horizontal axis, then you should consider using weighted Deming regression instead.

Residual Diagnostic Plots

Residual Diagnostic Plots







The residual diagnostic plots should be used to check the Deming regression assumptions of residual normality. The points in the residual scatter plot should be evenly distributed across the horizontal axis, with no apparent patterns or definite shapes. The points in the normal probability plot should follow closely to the straight line. The histogram is really only useful if you have a sufficient number of residuals, which is clearly not the case here.

You have the option of generating many different kinds of

Example 2 – Simple Deming Regression with Unknown Measurement Error

This section presents an example of how to run a simple Deming regression analysis when the measurement errors for both *Y* and *X* are unknown and must be estimated from duplicate data values. This example uses the *DemingReg2* dataset.

Setup

To run this example, complete the following steps:

1 Open the DemingReg2 example dataset

- From the File menu of the NCSS Data window, select **Open Example Data**.
- Select **DemingReg2** and click **OK**.

2 Specify the Deming Regression procedure options

- Find and open the **Deming Regression** procedure using the menus or the Procedure Navigator.
- The settings for this example are listed below and are stored in the **Example 2** settings template. To load this template, click **Open Example Template** in the Help Center or File menu.

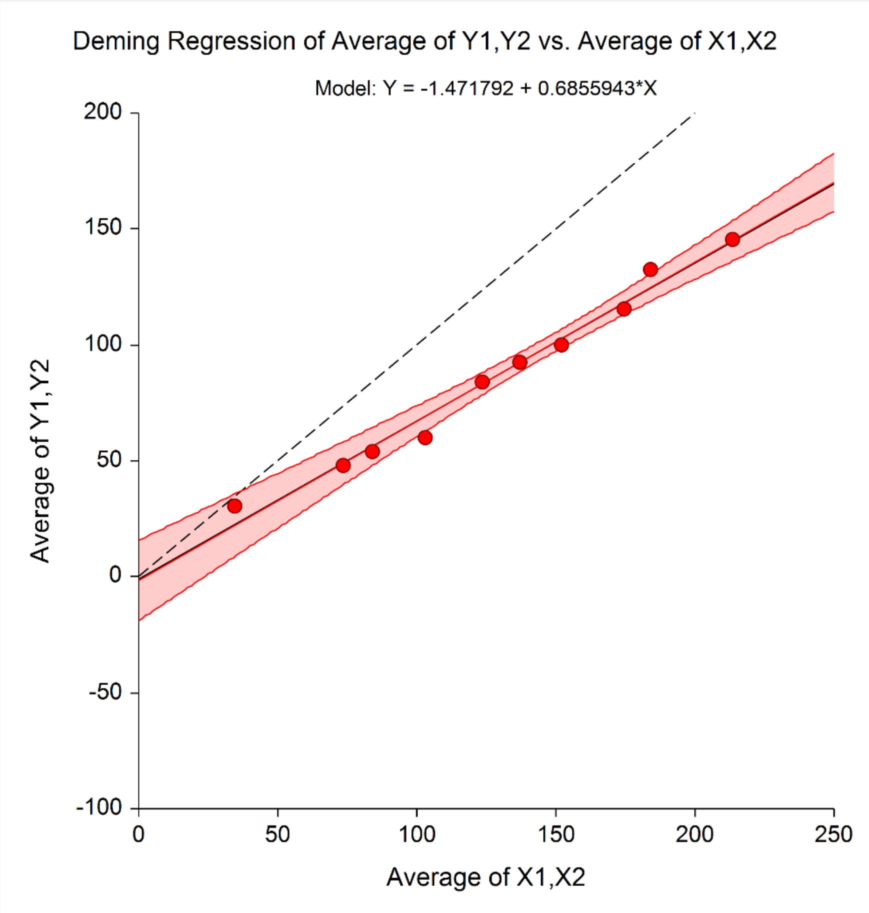
Option	Value
Variables Tab	
Y Measurement Error	**Unknown (Estimate from Multiple Measurements)**
Y Variables	**Y1-Y2**
X Measurement Error	**Unknown (Estimate from Multiple Measurements)**
X Variables	**X1-X2**
Plots Tab	
Deming Regression Scatter Plot	**Checked**
All Other Plots	**Unchecked**
Deming Regression Scatter Plot Format (*Click the Button*)	
More Lines Tab	
Regression Line	**Checked**

3 Run the procedure

- Click the **Run** button to perform the calculations and generate the output.

Output

Deming Regression Scatter Plot



Run Summary

Item	Value	Item	Value
Y Measurement Error	Unknown	Rows Processed	10
Y Variables	Y1,Y2	Rows with Missing Values	0
X Measurement Error	Unknown	Rows Used in Estimation	10
X Variables	X1,X2		
		Error Variance Ratio (λ)	1.00000
Deming Regression Type	Simple	Y Error Variance	3.25000
SE Estimation	Jackknife	X Error Variance	3.25000
Jackknife DF	N-2	Y Coefficient of Variation	0.02090
Intercept	-1.47179	X Coefficient of Variation	0.01409
Slope	0.68559		

Descriptive Statistics

Variable(s)	N	Mean	Error SD	Error Variance	Error COV
Y: Y1,Y2	10	86.25	1.80278	3.25000	0.02090
X: X1,X2	10	127.95	1.80278	3.25000	0.01409

Error Variance Ratio (λ) = 1.00000

Regression Coefficient Estimation

Parameter	Regression Coefficient b(i)	Jackknife Standard Error SE[b(i)]	DF	Lower 95% Conf. Limit of β(i)	Upper 95% Conf. Limit of β(i)
Intercept	-1.47179	7.49434	8	-18.75378	15.81019
Slope (X1,X2)	0.68559	0.04981	8	0.57074	0.80045

The T-value used to calculate the confidence limits was 2.30600, with N-2 = 8 degrees of freedom.

Estimated Model
Y = -1.47179171648271 + 0.685594308061608 * X

Hypothesis Test of Y = X

Test	Parameter Estimate	Jackknife Standard Error	DF	T-Statistic	Prob Level	Reject H0 at α = 0.025?
H0: Slope = 1 vs. H1: Slope ≠ 1	0.68559	0.04981	8	-6.31245	0.00023	Yes
H0: Mean Y - Mean X = 0 vs. H1: Mean Y - Mean X ≠ 0	-41.70000	5.80048	8	-7.18906	0.00009	Yes

In this case the slope if very different from one (reject H0 with p-value = 0.00023) and the location parameters are not equal (reject H0 with p-value = 0.00009), indicating a systematic difference between the two variables. Interestingly, the simple linear regression line is not too much different from the Deming regression line in this case.

Example 3 – Weighted Deming Regression

This section presents an example of how to run a weighted Deming regression analysis. This example uses the *DemingReg3* dataset. In the example, we will assume a constant error ratio of 1.

Setup

To run this example, complete the following steps:

1 Open the DemingReg3 example dataset

- From the File menu of the NCSS Data window, select **Open Example Data**.
- Select **DemingReg3** and click **OK**.

2 Specify the Deming Regression procedure options

- Find and open the **Deming Regression** procedure using the menus or the Procedure Navigator.
- The settings for this example are listed below and are stored in the **Example 3** settings template. To load this template, click **Open Example Template** in the Help Center or File menu.

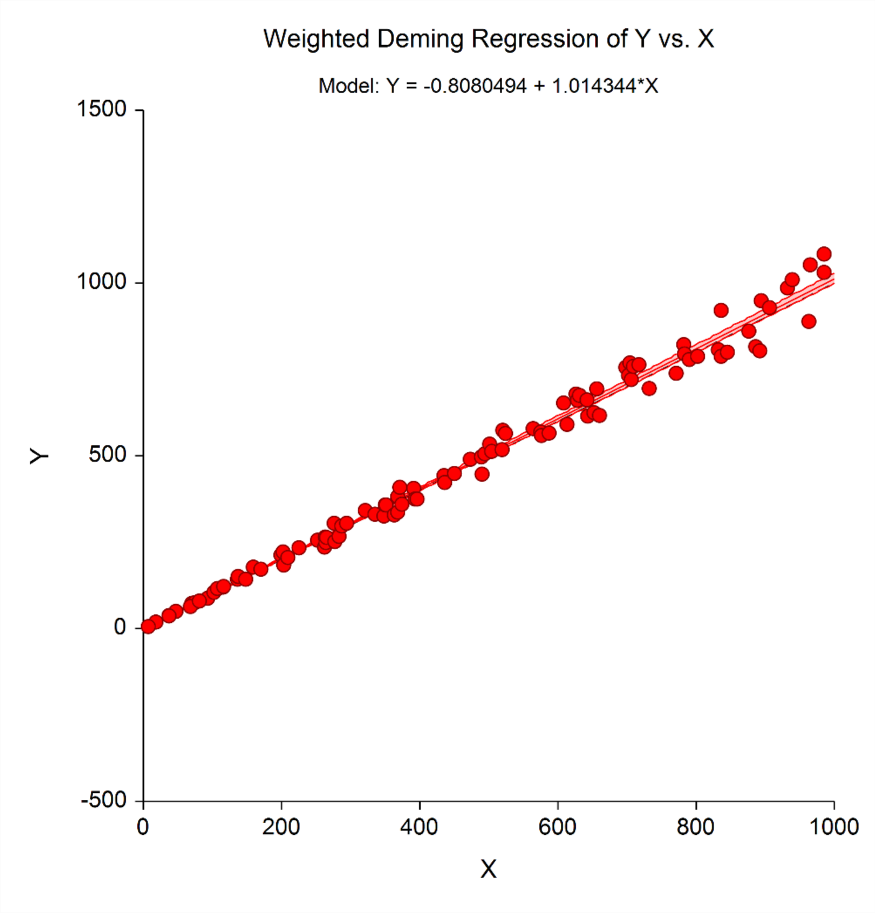
Option	Value
Variables Tab	
Y Variable	**Y**
X Variable	**X**
Deming Regression Type	**Weighted**

3 Run the procedure

- Click the **Run** button to perform the calculations and generate the output.

Output

Deming Regression Scatter Plot



Run Summary

Item	Value	Item	Value
Y Measurement Error	Known	Rows Processed	100
Y Variable	Y	Rows with Missing Values	0
X Measurement Error	Known	Rows Used in Estimation	100
X Variable	X		
		Error Variance Ratio (λ)	1.00000
Deming Regression Type	Weighted	Y Error Variance	1.00000
SE Estimation	Jackknife	X Error Variance	1.00000
Jackknife DF	N-2	Y Coefficient of Variation	0.00209
Iterations	3 of 100	X Coefficient of Variation	0.00212
Convergence	Normal		
Intercept	-0.80805		
Slope	1.01434		

Descriptive Statistics

Variable(s)	N	Mean	Error SD	Error Variance	Error COV
Y: Y	100	477.75	1.00000	1.00000	0.00209
X: X	100	471.76	1.00000	1.00000	0.00212

Error Variance Ratio (λ) = 1.00000

Regression Coefficient Estimation

Parameter	Regression Coefficient b(i)	Jackknife Standard Error SE[b(i)]	DF	Lower 95% Conf. Limit of β(i)	Upper 95% Conf. Limit of β(i)
Intercept	-0.80805	1.42057	98	-3.62713	2.01103
Slope (X)	1.01434	0.00847	98	0.99753	1.03116

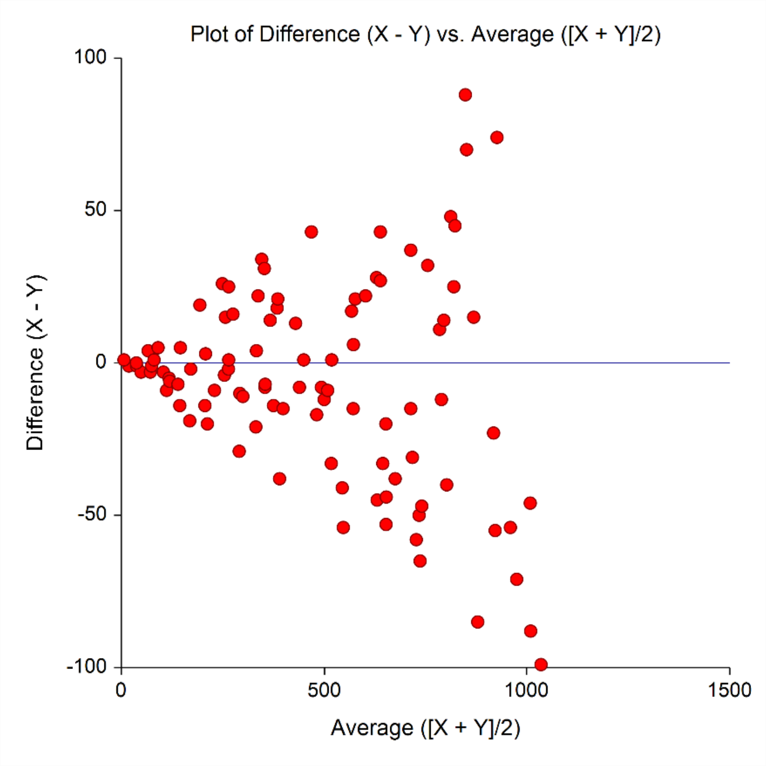
The T-value used to calculate the confidence limits was 1.98447, with N-2 = 98 degrees of freedom.

Estimated Model
Y = -0.808049379959723 + 1.01434367011481 * X

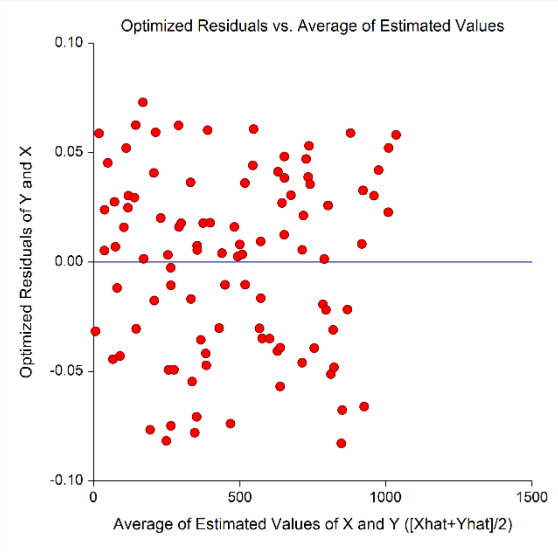
Hypothesis Test of Y = X

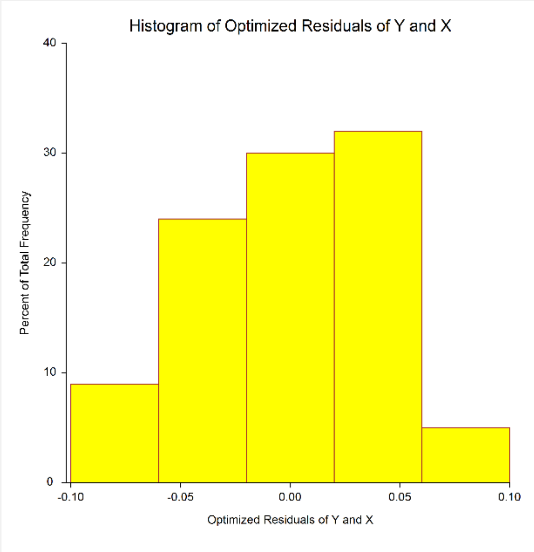
Test	Parameter Estimate	Jackknife Standard Error	DF	T-Statistic	Prob Level	Reject H0 at α = 0.025?
H0: Slope = 1 vs. H1: Slope ≠ 1	1.01434	0.00847	98	1.69317	0.09360	No
H0: Mean Y - Mean X = 0 vs. H1: Mean Y - Mean X ≠ 0	-0.52396	1.68469	98	-0.31101	0.75645	No

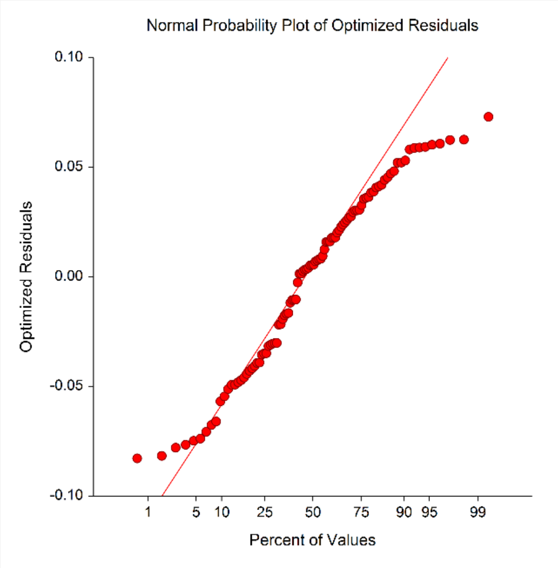
Difference vs. Average Plot



Residual Diagnostic Plots







The cone-shaped pattern in the Difference vs. Average plot indicates the need for weighted Deming regression. The proportional increase in standard deviation is also apparent on the Deming regression scatter plot. The results indicate that the regression line does not differ significantly from unity.

Example 4 – Simple Deming Regression with a Grouping Variable

This section presents an example of how to run a simple Deming regression analysis with a grouping variable. This example uses the *DemingReg4* dataset. In the example, we will assume a constant error ratio of 1.

Setup

To run this example, complete the following steps:

1 Open the DemingReg4 example dataset

- From the File menu of the NCSS Data window, select **Open Example Data**.
- Select **DemingReg4** and click **OK**.

2 Specify the Deming Regression procedure options

- Find and open the **Deming Regression** procedure using the menus or the Procedure Navigator.
- The settings for this example are listed below and are stored in the **Example 4** settings template. To load this template, click **Open Example Template** in the Help Center or File menu.

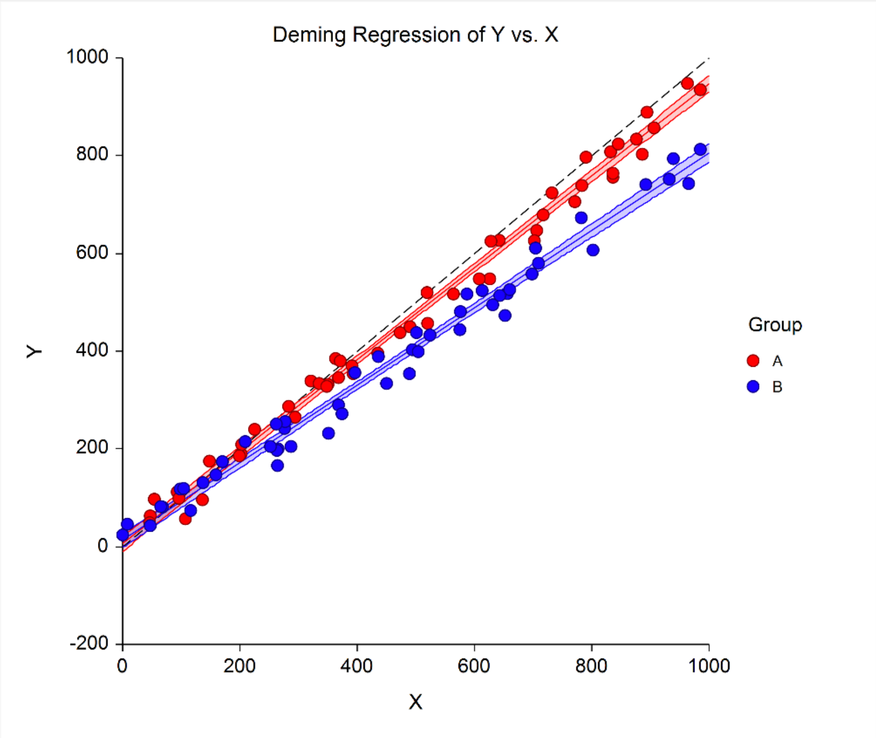
Option	Value
Variables Tab	
Y Variable	**Y**
X Variable	**X**
Grouping Variable	**Group**
Reports Tab	
Regression Coefficient Estimation	**Checked**
All Other Reports	**Unchecked**
Plots Tab	
Deming Regression Scatter Plot	**Checked**
Show Combined Plot (If Grouping Variable Present)	**Checked**
All Other Plots	**Unchecked**

3 Run the procedure

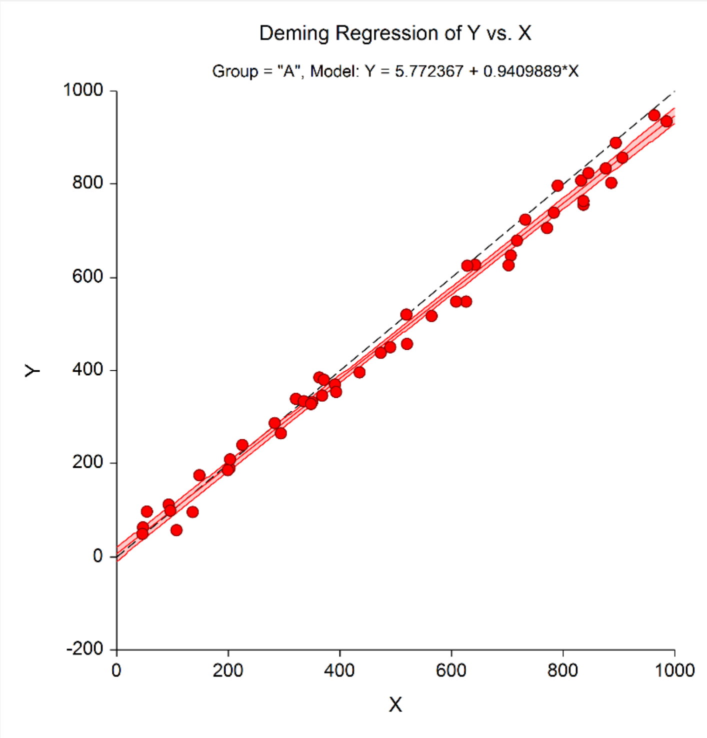
- Click the **Run** button to perform the calculations and generate the output.

Output

Combined Deming Regression Scatter Plot



Deming Regression Scatter Plot for Group = "A"



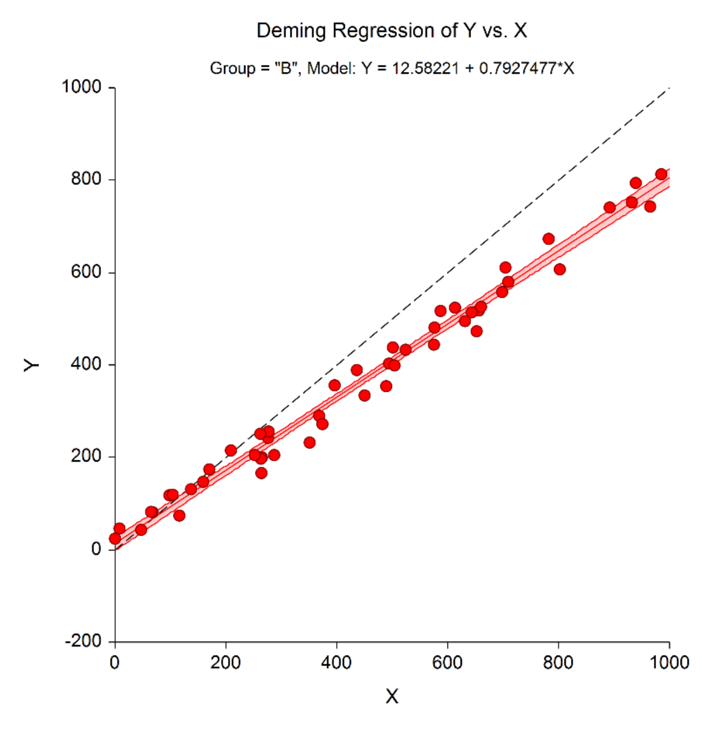
Regression Coefficient Estimation for Group = "A"

Parameter	Regression Coefficient b(i)	Jackknife Standard Error SE[b(i)]	DF	Lower 95% Conf. Limit of β(i)	Upper 95% Conf. Limit of β(i)
Intercept	5.77237	7.54389	48	-9.39564	20.94037
Slope (X)	0.94099	0.01413	48	0.91258	0.96940

The T-value used to calculate the confidence limits was 2.01063, with N-2 = 48 degrees of freedom.

Estimated Model
Y = 5.77236644152418 + 0.940988919637653 * X

Deming Regression Scatter Plot for Group = "B"



Regression Coefficient Estimation for Group = "B"

Parameter	Regression Coefficient b(i)	Jackknife Standard Error SE[b(i)]	DF	Lower 95% Conf. Limit of β(i)	Upper 95% Conf. Limit of β(i)
Intercept	12.58221	7.21137	48	-1.91722	27.08165
Slope (X)	0.79275	0.01431	48	0.76398	0.82151

The T-value used to calculate the confidence limits was 2.01063, with N-2 = 48 degrees of freedom.

Estimated Model
Y = 12.5822121029879 + 0.792747665759649 * X

Separate regressions are performed for each group. The combined plot displays both results on a single graph. Here we can see that group “B” deviates more from the unity line than group “A”. After the combined plot, the results are given separately for each group.

Example 5 – Validation of Simple Deming Regression using R ("mcr" package by Manuilova E. et al. (2014))

This section uses the data from Example 1, *DemingReg1*, to validate Simple Deming Regression using R, with the "mcr" package (version 1.2.1) created by Manuilova, E. et al. (2014). Again suppose that the measurement errors are known for both *Y* and *X* with Variance(Y) = 0.008 and Variance(X) = 0.032 (Error ratio: $\lambda = 4.0$). Furthermore, recall that the "mcr" package in R uses $N - 2$ degrees of freedom, the default in **NCSS**.

The R script to perform this Deming regression with the "mcr" package is:

```
library(mcr)
X=c(7,8.3,10.5,9,5.1,8.2,10.2,10.3,7.1,5.9)
Y=c(7.9,8.2,9.6,9,6.5,7.3,10.2,10.6,6.3,5.2)
model=mcreg(X,Y,error.ratio=4,alpha=0.05,method.reg="Deming",method.ci="jackknife")
printSummary(model)
```

This script generates the following results in R:

```
DEMING REGRESSION FIT:

                   EST          SE         LCI        UCI
Intercept   -0.0897449   1.7219874  -4.0606550   3.881165
Slope        1.0011942   0.1871771   0.5695632   1.432825
```

Setup

To run this example, complete the following steps:

1 Open the DemingReg1 example dataset

- From the File menu of the NCSS Data window, select **Open Example Data**.
- Select **DemingReg1** and click **OK**.

2 Specify the Deming Regression procedure options

- Find and open the **Deming Regression** procedure using the menus or the Procedure Navigator.
- The settings for this example are listed below and are stored in the **Example 5** settings template. To load this template, click **Open Example Template** in the Help Center or File menu.

Option	Value
Variables Tab	
Y Measurement Error	**Known SD, Variance, or COV**
Input Type	**Variance**
=	**0.008**
Y Variable	**Y**
X Measurement Error	**Known SD, Variance, or COV**
Input Type	**Variance**
=	**0.032**
X Variable	**X**
Deming Regression Type	**Simple**

Reports Tab

Regression Coefficient Estimation.......... **Checked**
All Other Reports .... **Unchecked**

Decimal Places

Betas (Coefficients) **7**
SD, SE, Variance, COV.......................... **7**

Plots Tab

All Plots... **Unchecked**

3 Run the procedure

- Click the **Run** button to perform the calculations and generate the output.

Output

Regression Coefficient Estimation

Parameter	Regression Coefficient b(i)	Jackknife Standard Error SE[b(i)]	DF	Lower 95% Conf. Limit of β(i)	Upper 95% Conf. Limit of β(i)
Intercept	-0.0897449	1.7219874	8	-4.0606550	3.8811652
Slope (X)	1.0011942	0.1871771	8	0.5695632	1.4328253

The T-value used to calculate the confidence limits was 2.30600, with N-2 = 8 degrees of freedom.

Estimated Model
Y = -0.0897448990070444 + 1.00119422781949 * X

All of the Deming regression coefficient output from **NCSS** matches the R output exactly.

Example 6 – Validation of Weighted Deming Regression using R ("mcr" package by Manuilova E. et al. (2014))

This section uses the data from Example 1, *DemingReg1*, to validate Weighted Deming Regression using R, with the "mcr" package (version 1.2.1) created by Manuilova, E. et al. (2014). Again suppose that the measurement errors are known for both *Y* and *X* with Variance(Y) = 0.008 and Variance(X) = 0.032 (Error ratio: $\lambda = 4.0$). Furthermore, recall that the "mcr" package in R uses *N* – 2 degrees of freedom, the default in **NCSS**.

The R script to perform this Deming regression with the "mcr" package is:

```
library(mcr)
X=c(7,8.3,10.5,9,5.1,8.2,10.2,10.3,7.1,5.9)
Y=c(7.9,8.2,9.6,9,6.5,7.3,10.2,10.6,6.3,5.2)
model=mcreg(X,Y,error.ratio=4,alpha=0.05,method.reg="WDeming",method.ci="jackknife")
printSummary(model)
```

This script generates the following results in R:

```
WEIGHTED DEMING REGRESSION FIT:

                  EST          SE        LCI        UCI
Intercept  -0.3283761   1.9743380 -4.8812076   4.224455
Slope       1.0312280   0.2202037  0.5234374   1.539019
```

Setup

To run this example, complete the following steps:

1 Open the DemingReg1 example dataset

- From the File menu of the NCSS Data window, select **Open Example Data**.
- Select **DemingReg1** and click **OK**.

2 Specify the Deming Regression procedure options

- Find and open the **Deming Regression** procedure using the menus or the Procedure Navigator.
- The settings for this example are listed below and are stored in the **Example 6** settings template. To load this template, click **Open Example Template** in the Help Center or File menu.

Option	Value
Variables Tab	
Y Measurement Error	**Known SD, Variance, or COV**
Input Type	**Variance**
=	**0.008**
Y Variable	**Y**
X Measurement Error	**Known SD, Variance, or COV**
Input Type	**Variance**
=	**0.032**
X Variable	**X**
Deming Regression Type	**Weighted**

Reports Tab
Regression Coefficient Estimation.......... **Checked**
All Other Reports .... **Unchecked**

Decimal Places
Betas (Coefficients) **7**
SD, SE, Variance, COV.......................... **7**

Plots Tab
All Plots.. **Unchecked**

3 Run the procedure

- Click the **Run** button to perform the calculations and generate the output.

Output

Regression Coefficient Estimation

Parameter	Regression Coefficient b(i)	Jackknife Standard Error SE[b(i)]	DF	Lower 95% Conf. Limit of β(i)	Upper 95% Conf. Limit of β(i)
Intercept	-0.3283761	1.9743380	8	-4.8812076	4.2244554
Slope (X)	1.0312280	0.2202037	8	0.5234374	1.5390186

The T-value used to calculate the confidence limits was 2.30600, with N-2 = 8 degrees of freedom.

Estimated Model
Y = -0.328376138786767 + 1.03122798996277 * X

All of the weighted Deming regression coefficient output from **NCSS** matches the R output exactly.