CHAPTER 3
UNREPLICATED LINEAR FUNCTIONAL RELATIONSHIP MODEL

This chapter proposes an initial attempt to solve the three issues in the existing ISMs as mentioned in Section 2.4, namely (i) perfect reference, (ii) bivariate case, and (iii) global versus localized measure. The discussion provides a clear direction for the study to develop a new unified similarity measure for one-to-one image comparison. The main purpose in this chapter is to develop a new similarity measure based on the unreplicated linear functional relationship (ULFR) model. This new similarity measure tries to overcome the first constraint by introducing a new concept of imperfect reference image. Some background readings of linear functional model and reasons not to choose other regression models and correlation coefficients are discussed. The new similarity measure defined by coefficient of determination for the ULFR model is derived and its properties will be discussed.

3.1 Why Some Regression Models Are Not Suitable?

As we noted in Table 2.1, there are some regression models or correlation coefficients that has been used for one-to-one image comparisons. Examples of application are the simple linear regression (Nielsen et al., 1997), multiple linear regression (Chang & Tan, 2006), canonical correlation (Tan & Chang, 2006) and ordinal correlation (Avcibas, 2002; Cramariuc et al., 2000; Bhat & Nayar, 1998). On the other hand, logistic regression, multinomial logistic regression and non-linear regression are the examples in which these models have not been used for comparing two images. Even though we have learnt from the literature review that performance indicators based on correlation measures generally perform better than other Statistical-based measures, they are still not considered the three issues mentioned in previous chapter. This section will show the reasons why some regression models and its corresponding correlation or coefficient of determination are inadequate for comparing two images.

Some existing or potential regression models and its correlations are summarized below:

(i) Simple linear regression model is $y_i = \alpha + \beta x_i + \varepsilon_i$ and its squared correlation or coefficient of determination is defined by

$$R^2 = \frac{S_{xy}^2}{S_{xx}S_{yy}} = \frac{\left[\sum\left(x-\bar{x}\right)\left(y-\bar{y}\right)\right]^2}{\sum\left(x-\bar{x}\right)^2\sum\left(y-\bar{y}\right)^2} \tag{3.1}$$

(ii) Multiple linear regression model is $y_i = \alpha + \beta_1 x_{1i} + \beta_2 x_{2i} + \cdots \beta_p x_{pi} + \varepsilon_i$ and its multiple correlation is defined by

$$R_M^2 = \frac{SS_R}{S_{GG}} \tag{3.2}$$

which is the generalisation of the Equation (3.1).

(iii) The canonical correlation for the model $V = b_1 Y_1 + \cdots + b_q Y_q$ and $W = a_1 X_1 + \cdots + a_p X_p$ is defined by

$$\lambda_i^2 = \left|\Sigma_{YY}^{-1}\Sigma_{YX}\Sigma_{XX}^{-1}\Sigma_{XY}\right|; \quad i = 1,2,\ldots,k \tag{3.3}$$

(iv) Spearman's rank Correlation:

$$Z4 = \sum_{\mu=1}^{b} \max_{k=1,2,\cdots,K}\left\{SRC_\mu^k\right\} \tag{3.4}$$

where SRC_μ^k is the Spearman's rank correlation of block number μ and of the *k*th spectral band

(v) Logistic regression is $\operatorname{logit}(p) = \alpha + \beta_1 x_{1i} + \beta_2 x_{2i} + \cdots \beta_p x_{pi}$ and multinomial logistic regression is equivalent to fitting n-1 binary logistic regression.

There are three methods of deriving the coefficient of determination (Maddala, 1983; Menard, 2001):

$$R_1^2 = \frac{\left(\ln L_0 - \ln L(\beta)\right)}{\left(\ln L_0 - \ln L(\beta)\right) + \left(\ln L(\beta) - \ln L_S\right)}$$

$$= \frac{\left(\ln L_0 - \ln L(\beta)\right)}{\left(\ln L_0 - \ln L_S\right)} \tag{3.5}$$

$$R_2^2 = 1 - \exp\left[-\frac{2}{n}\left\{\ln\left(L(\beta)\right) - \ln\left(L_0\right)\right\}\right]$$

$$= 1 - \left(\frac{L_0}{L(\beta)}\right)^{\frac{2}{n}} \tag{3.6}$$

$$R_3^2 = \frac{R_2^2}{\left(\max\text{imum possible } R_2^2\right)}$$

$$= \frac{1 - \left(\frac{L_0}{L(\beta)}\right)^{\frac{2}{n}}}{1 - \left(L_0\right)^{\frac{2}{n}}} \tag{3.7}$$

Table 3.1 compares different regression models and correlations for different model properties. Let *X* and *Y* represent the reference image and distorted image (the image being measured). Note that none of these correlation methods assume that both *X* and *Y* subject to errors and hence, the need for non-perfect reference was not met. To consider a multiple image features, both *X* and *Y* must be multivariate. However, it was shown that only the canonical correlation met this requirement. Furthermore, the squared Pearson correlation, multiple correlation and canonical correlation can be easily modified to measure the image quality both globally and locally. However, the logistic regression, due to its binary value of *Y*, it can only be used for global measure or localized measure, but not both. Lastly, the Spearman’s rank correlation is a localized measure.

Table 3.1: Comparing different correlation based metrics.

	Squared Pearson correlation	Multiple correlation	COD for Logistic regression	Canonical correlation	Spearman's rank Correlation
Y subject to error	Yes	Yes	Yes	No	No
X subject to error	No	No	No	No	No
Multivariate for *X*	No	Yes	Yes	Yes	No
Multivariate for *Y*	No	No	No	Yes	No
Global or localized measure	Both	Both	Either	Both	Localized

This suggests the need to look for a more suitable relationship model for comparing two images. The first attempt is to adapt the error terms into both dependent variable and independent variable of a model; meant the reference image and the distorted image are both subject to errors. As an initial stage, the unreplicated linear functional relationship model and the development of its coefficient of determination is discussed in Section 3.3. A complete similarity measure for comparing two images that consider all constraint areas will be developed in the next Chapter.

3.2 Linear Functional Relationship Models

Over the centuries, linear regression model has become the central of study for many applications to investigate relationship between a response variable and a set of explanatory variables. In practice, however, such as engineering, economics, psychology, chemistry, biology and others often dealt with a situation where this relationship is obscured by random fluctuations associated with both variables (Sprent, 1969). Fuller (1987) made the same comment where the assumption that the explanatory variable can be measured exactly may not be realistic in many situations. Such experiences had lead to the development of a new type of linear relationship when both variables are subject to error or so called functional relationship although other terms have referred are 'law-like relationships', regression with errors in *x*', 'errors-in-models' and 'measurement error models'.

3.2.1. Basic Definition of Unreplicated Linear Functional Relationship Model

Suppose *X* and *Y* are two linearly related unobservable variables

$$Y_i = \alpha_F + \beta_F X_i \tag{3.8}$$

and the two corresponding random variables *x* and *y* are observed with errors δ and ε respectively

$$\left.\begin{aligned} x_i &= X_i + \delta_i \\ y_i &= Y_i + \varepsilon_i \end{aligned}\right\} \quad i = 1, 2, \cdots, n. \tag{3.9}$$

It is common to assume that the errors δ_i and ε_i are mutually independent and normally distributed random variables with $\delta_i \sim N\left(0, \sigma^2\right)$ and $\varepsilon_i \sim N\left(0, \tau^2\right)$. This implies that

(i) both errors have mean 0, that is $E\left(\delta_i\right) = E\left(\varepsilon_i\right) = 0$, $i = 1, 2, \ldots, n$

(ii) both errors have constant but different variance, that is $Var\left(\delta_i\right) = \sigma^2, \quad Var\left(\varepsilon_i\right) = \tau^2$, $i = 1, 2, \ldots, n$

(iii) the errors are uncorrelated, that is

$$Cov\left(\delta_i, \delta_j\right) = 0 = Cov\left(\varepsilon_i, \varepsilon_j\right), \ \forall i \neq j; \quad i, j = 1, 2, \ldots, n$$

$$Cov\left(\delta_i, \varepsilon_i\right) = 0, \ \forall i, j = 1, 2, \ldots, n$$

Hussin (1997) termed the model (3.8) and (3.9) as the unreplicated linear functional relationship (ULFR) model when there is only a single *x* and *y* observation for each level of *i*. The model is termed as replicated linear functional relationship (RLFR) model for multiple *x* and *y* observations at each level of *i*. Besides, Kendall (1951, 1952) formally made a distinction between functional and structural relationship between the two variables. In functional relationship, the unobservable X is a constant or a mathematical variable without specific distributional properties, whereas the unobservable X is usually assumed to be normally distributed in structural relationship. However, in practice, applications using data do not differentiate the functional or structural model (Sprent, 1990).

The log likelihood function is given by

$$L\left(\alpha_F, \beta_F, X_1, \ldots, X_n, \sigma_\delta^2, \sigma_\varepsilon^2\right)$$
$$= -n\log\left(2\pi\right) - \frac{1}{2}n\left(\log\sigma_\delta^2 + \log\sigma_\varepsilon^2\right) - \sum\frac{\left(x_i - X_i\right)^2}{2\sigma_\delta^2} - \sum\frac{\left(y_i - \alpha_F - \beta_F X_i\right)^2}{2\sigma_\varepsilon^2}$$

When the ratio of the error variances is known, that is $\frac{\sigma_\varepsilon^2}{\sigma_\delta^2} = \lambda$, then the maximum likelihood estimators of parameters α_F, β_F, σ_δ^2 and X_i are (Hussin, 1997)

$$\hat{\alpha}_F = \overline{y} - \hat{\beta}_F\overline{x} \tag{3.10}$$

$$\hat{\beta}_F = \frac{\left(S_{yy} - \lambda S_{xx}\right) + \left\{\left(S_{yy} - \lambda S_{xx}\right)^2 + 4\lambda S_{xy}\right\}^{1/2}}{2S_{xy}} \tag{3.11}$$

$$\hat{\sigma}_\delta^2 = \frac{1}{n-2}\left[\sum\left(x_i - \hat{X}_i\right)^2 + \frac{1}{\lambda}\sum\left(y_i - \hat{\alpha}_F - \hat{\beta}_F\hat{X}_i\right)^2\right] \tag{3.12}$$

and $$\hat{X}_i = \frac{\lambda x_i + \hat{\beta}_F\left(y_i - \hat{\alpha}\right)}{\lambda + \hat{\beta}_F^2} \tag{3.13}$$

where $\overline{y} = \frac{\sum y_i}{n}$, $\overline{x} = \frac{\sum x_i}{n}$, $S_{yy} = \sum\left(y_i - \overline{y}\right)^2$, $S_{xx} = \sum\left(x_i - \overline{x}\right)^2$ and $S_{xy} = \sum\left(x_i - \overline{x}\right)\left(y_i - \overline{y}\right)$.

3.2.2 Historical Remarks In Brief

Sprent (1969, 1990) had documented a comprehensive overview of the history of functional and structural relationships in 1990. The study covered the major developments of the functional and structural relationships for more than a century, back to year 1877 to 1987. With a little extra effort, some recent papers were included in this study and reviewed the topic from a different perspective.

(i) Study on the error variances

Adcock (1877, 1878) is the first person who had investigated the problem of fitting a linear relationship when both dependent variable and independent variable are subject to error (Sprent, 1990). Adcock (1877, 1878) obtained the least squares estimation for the slope, β_F, for a special case of the model (3.8) and (3.9) where the two error variances are equal; i.e. $\sigma^2 = \tau^2$. Kummel (1879) extended the result to the case where the ratio of the error variances, $\lambda = \frac{\sigma^2}{\tau^2}$ is known. In 1940, Wald (1940) dealt with group data for parameters estimation where the n observations are divided into q groups of equal size and the errors variance are finite. It was observed that the distribution of the errors is not affected by the grouping.

Furthermore, Lindley (1947) obtained conditions on the error distribution that would preserve linearity of regression for the observable random variables. This result was linked indirectly to the work by Geary (1949), where assumption concerning the error variances can be waived using product-cumulants for parameters estimation. However, it is not applicable to normal distribution.

Madansky (1959) and Moran (1971) reviewed the models using a variety of error structures for both single relationships in the two and higher variate case. Many anomalies associated with errors assumptions and confidence intervals problems were resolved in 1960s and 1970s (Sprent, 1990).

In an early paper published in 1966, Sprent (1966) used generalized least squares for point estimate when there is information on the correlated departures, i.e.

$\sigma_{\delta\varepsilon}$, $\sigma_{\delta\delta}$ and $\sigma_{\varepsilon\varepsilon}$. In this case, it is no longer true that $Y_i = \alpha_F + \beta_F X_i$ reduces to $Y_i = \beta_F X_i$. At about the same time, Fisk (1966) and Malinvaud (1966) had separately considered some special patterns of correlation of departures in economic situations. One example of these departures patterns is when the covariance matrix has the following form:

$$\Sigma = \sigma^2 \begin{bmatrix} 1 & \rho & \rho^2 & \cdots & \cdots & \cdots & \rho^{n-1} \\ \rho & 1 & \rho & \rho^2 & \cdots & \cdots & \rho^{n-2} \\ \vdots & \vdots & & & & & \vdots \\ \vdots & \vdots & & & \rho & 1 & \rho \\ \rho^{n-1} & \cdots & \cdots & \cdots & \rho^2 & \rho & 1 \end{bmatrix}$$

Lastly, Chan and Mak (1984) considered heteroscedastic errors in multivariate linear functional relationship, in which the error variances and covariances are not necessary homogeneous. They applied a modified least squares method to estimate the structural parameters of the model. This method does not require the distributional assumptions on the errors and it implies that the derived estimating equations are still applicable when the errors are not necessarily normal.

(ii) Types of variable *X*

An important research direction for functional relationship model is the consideration of various types of variable *X*. In 1901, Pearson (1901) extended Adcock's solution in Equations (3.8) and (3.9) to the multiple (principle component) relationship with X_{ki}, $k = 1, 2, \ldots, p$ with $\sigma^2 = \tau^2$. On the other hand, Reiersol (1945) introduced an instrumental variable, that is, variable *Z* correlated with the true *X* and *Y*, but independent of the errors. It was further studied by Barnett (1969) in 1969 in a structural relationship problem and indicated that the maximum likelihood estimator works without further assumptions on the error variance.

The Berkson model is a special functional relationship model introduced by Berkson (1950) in 1950. In this model, the controlled variable x is fixed a *priori* at a value chosen by the experimenter. On the other hand, Moran (1971) considered a special case where $\alpha = 0$. He showed that the shift of origin to a data determined point is not equivalent to fitting a line through a predetermined origin that is not the data mean.

Sprent (1969) discussed the multidimensional functional relationship with a single linear functional relationship given by $Y_i = \alpha + \beta_1 X_{1i} + \beta_2 X_{2i} + \cdots + \beta_p X_{pi}$. There is at least one or more independent linear relationships or replication, each represents a space of $p-1$ dimensions. Ramsay & Silverman (1997) termed this as functional linear models for scalar response. Furthermore, Chan & Mak (1983) considered the estimation of multivariate linear functional relationships

$$\begin{bmatrix} \boldsymbol{x}_i \\ \boldsymbol{y}_i \end{bmatrix} = \begin{bmatrix} \mathbf{0} \\ \boldsymbol{a} \end{bmatrix} + \begin{bmatrix} \boldsymbol{I} \\ \boldsymbol{B} \end{bmatrix} \boldsymbol{X}_i + \begin{bmatrix} \boldsymbol{\delta}_i \\ \boldsymbol{\varepsilon}_i \end{bmatrix} \tag{3.14}$$

where $\boldsymbol{x}_i = \begin{bmatrix} x_{1i} \\ \vdots \\ x_{pi} \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix} + \begin{bmatrix} 1 & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} X_{1i} \\ \vdots \\ X_{pi} \end{bmatrix} + \begin{bmatrix} \delta_{1i} \\ \vdots \\ \delta_{pi} \end{bmatrix}$ and

$$\boldsymbol{y}_i = \begin{bmatrix} y_{1i} \\ \vdots \\ y_{qi} \end{bmatrix} = \begin{bmatrix} \beta_0 \\ \vdots \\ \beta_0 \end{bmatrix} + \begin{bmatrix} \beta_{11} & \beta_{12} & \cdots & \beta_{1p} \\ \beta_{21} & \beta_{22} & \cdots & \beta_{2p} \\ \vdots & \vdots & & \vdots \\ \beta_{q1} & \beta_{q2} & \cdots & \beta_{qp} \end{bmatrix} \begin{bmatrix} X_{1i} \\ \vdots \\ X_{pi} \end{bmatrix} + \begin{bmatrix} \varepsilon_{1i} \\ \vdots \\ \varepsilon_{qi} \end{bmatrix}.$$

(iii) Types of relationship

Dent (1935) proposed the geometric mean functional relationship estimator of slope

$$\beta_F = \operatorname{sgn}\left(m_{xy}\right)\left(\frac{m_{yy}}{m_{xx}}\right)^{1/2} \tag{3.15}$$

where m_{yy}, m_{xx} and m_{xy} are the sample variances and covariance, respectively. The estimator in Equation (3.15) is used when assessment of the error variances or their ratio is not possible. However, this estimator is generally not consistent.

There was also interest in considering several simultaneous linear relationships between p variates subject to error by Gleser & Watson (1973) considered several linear relationships (simultaneous relationships). Sprent (1969) regarded this as multivariate regression. These simultaneous relationships frequently occurred in economics and physical sciences where it is often involving large measurement errors.

Instead of linear functional relationship, a non-linear relationships and transformation of data may occur in some problems (Bhat & Nayar, 1998). One example of the non-linear relationships study was carried out by Huxley (1924). Huxley (1924) transformed the data using logarithm and studying the linear relationship between the transformed log x and log y in biological studies. In this context, Huxley referred it as the "simple allometry relationship".

An excellent work on functional data analysis had also been done by Ramsay & Silverman (1997). They had considered a wide range of functional linear models which including functional canonical correlation analysis, relationship for both response y and covariate x are functions as $\hat{y}_i(t) = \alpha(t) + \int_{\Upsilon_X} x_i(t)\beta(s,t)\,ds$, functional responses with multivariate covariates, functional linear models for scalar responses and functional linear models for functional responses.

Several other types of functional relationship had also been studied by Abdul Ghapor Hussin since late 1990s. Hussin (1998) considered the unreplicated complex linear functional relationship model to analyses the wind direction data in 1998. The complex form of the model in (3.8) and (3.9) can be written as $(\cos Y + i\sin Y) = \alpha_C + \beta_C(\cos X + i\sin X)$ where $(\cos x_i + i\sin x_i) = (\cos X_i + i\sin X_i) + \delta_i$

and $\left(\cos y_i + i\sin y_i\right) = \left(\cos Y_i + i\sin Y_i\right) + \varepsilon_i$. After that, Hussin (2001) and Hussin & Chik (2003) considered the unreplicated linear circular functional relationship model where Equation (3.8) replaced by $Y_i = \alpha_F + \beta_F X_i \left(\bmod 2\pi\right)$. The former paper discussed parameter estimation using the maximum likelihood method, while the latter paper focused on the estimation of error concentration for circular functional model. In 2005, Hussin et al. (2005) and Hussin (2005) considered the pseudo-replicates and replicated linear circular functional relationship model, respectively.

Lastly, Ferraty & Vieu (2006) had discussed the functional nonparametric regression model together with some practical illustrations such as speech recognition and electricity consumption problems.

(iv) Parameter estimations

In view of the inconsistency problem of ML estimation, a number of studies were searching for improved methods on parameter estimation. For instance, Lindley & El Sayyad (1968) introduced Bayesian method for parameter estimation and Morton (1981) considered unbiased estimating equations using pivotals. Hussin (2005) applied the Fisher information matrix of parameters to estimate the variance of the estimated parameters. Other related works are generalized MLE (Chan & Mak, 1984), modified estimating equation (Buonaccorsi, 1996), the use of grouping in the bivariate case (Wald, 1940) and Karni & Weissman (1974) obtained a consistent estimator of β_F based on serial (rank) correlation. Hussin (2004) compared various estimators of slope parameters for ULFR model using simulation approach. Six estimation techniques were discussed, which are the two-group method of Wald (1940), the three-group method, the weighted regression, Housner-Brennan's method (Housner & Brennan, 1948), Durbin's ranking method (Durbin, 1954), and maximum likelihood method ($\left(\lambda = 1\right)$. The study concluded that maximum likelihood method and the weighted regression are favorable compared to other methods.

3.3 Coefficient of Determination for ULFR Model

It is the main objective of this study to propose a new image similarity measure that is able to overcome the three constraint areas. In view of the inadequacy of other regression models or correlations discussed in Section 3.1, this section begins with the ULFR model as defined in Equations (3.8) and (3.9) which fulfills the need to consider the reference image and distorted image where both images are subject to errors. The remaining task is to develop a correlation-based similarity measure via coefficient of determination (COD), which is appropriate for the purpose of 1-to-1 image comparison in Section 3.3.1.

3.3.1 Derivation of the Coefficient of Determination

In an ordinary simple linear regression (SL) analysis, we look at the COD as a measure of the variability in y explained by the regression model. The construction of the COD remains a good practice when fitting the ULFR model. We show the results of COD for ULFR in this section.

The Equations (3.8) and (3.9) can be rewritten as

$$y_i = \alpha_F + \beta_F X_i + \varepsilon_i \quad \text{for } i = 1, 2, \ldots, n \tag{3.16}$$

If we substitute X_i by $\left(x_i - \delta_i\right)$ in Equation (3.16), then we have the expression

$$\begin{aligned} y_i &= \alpha_F + \beta_F x_i + (\varepsilon_i - \beta_F \delta_i) \\ &= \alpha_F + \beta_F x_i + V_i \end{aligned}$$

where the errors of the model $V_i = (\varepsilon_i - \beta_F \delta_i) = y_i - (\alpha_F + \beta_F x_i)$, for $i = 1, 2, \ldots, n$ is a normally distributed random variable with zero mean and variance $\sigma_\varepsilon^2 + \beta_F^2 \sigma_\delta^2$. If $\hat{\alpha}_F$ and $\hat{\beta}_F$ are the estimates of α_F and β_F respectively, then

$$\hat{V}_i = y_i - \hat{y}_i = y_i - (\hat{\alpha}_F + \hat{\beta}_F x_i), \text{ for } i = 1, 2, \ldots, n \tag{3.17}$$

will be the residual of the model. From Anderson (1984) and Kendall & Stuart (1979) that the sum of squared distances of the observed points from the fitted line or the residual sum of squares (SS_E) is given as:

$$\text{SS}_E = \frac{\sum \left\{ y_i - \left(\hat{\alpha} + \hat{\beta}_F x_i \right) \right\}^2}{\left(\lambda + \hat{\beta}_F^2 \right)}$$

$$= \frac{\text{S}_{yy} - 2\hat{\beta}_F \text{S}_{xy} + \hat{\beta}_F^2 \text{S}_{xx}}{\left(\lambda + \hat{\beta}_F^2 \right)} \tag{3.18}$$

We shall consider here that the ratio of the error variances is equal to one (λ = 1). For those cases when $\lambda \neq 1$, we can always reduce this to the case of λ = 1 by dividing the observed values of y by $\lambda^{1/2}$ (see e.g. Kendall & Stuart, 1979). Hence, we have

$$\text{SS}_E = \frac{\text{S}_{yy} - 2\hat{\beta}_F \text{S}_{xy} + \hat{\beta}_F^2 \text{S}_{xx}}{1 + \hat{\beta}_F^2} \tag{3.19}$$

In the same way as ordinary linear regression, we can now define the COD of the ULFR $\left(R_F^2 \right)$ as the proportion of variation explained by the variable *x*, that is

$$R_F^2 = \frac{\text{SS}_R}{\text{S}_{yy}} \tag{3.20}$$

where SS_R is the regression sum of squares which can be derived as

$$\text{SS}_R = \text{S}_{yy} - \text{SS}_E$$

$$= \text{S}_{yy} - \frac{\text{S}_{yy} - 2\hat{\beta}_F \text{S}_{xy} + \hat{\beta}_F^2 \text{S}_{xx}}{1 + \hat{\beta}_F^2}$$

$$= \frac{\hat{\beta}_F^2 \mathrm{S}_{yy} + 2\hat{\beta}_F \mathrm{S}_{xy} - \hat{\beta}_F^2 \mathrm{S}_{xx}}{1 + \hat{\beta}_F^2}$$

$$= \frac{\hat{\beta}_F^2 \left(\mathrm{S}_{yy} - \mathrm{S}_{xx}\right) + 2\hat{\beta}_F \mathrm{S}_{xy}}{1 + \hat{\beta}_F^2} \tag{3.21}$$

We can summarise our proposed COD with the following results.

Result 1. Let the ratio of the error variances be known and equals to one, $\frac{\sigma_\varepsilon^2}{\sigma_\delta^2} = \lambda = 1$, then the COD for the ULFR model is

$$R_F^2 = \frac{\mathrm{SS}_R}{\mathrm{S}_{yy}} = \frac{\hat{\beta}_F \mathrm{S}_{xy}}{\mathrm{S}_{yy}} \tag{3.22}$$

Proof. To show that $\frac{\mathrm{SS}_R}{\mathrm{S}_{yy}} = \frac{\hat{\beta}_F \mathrm{S}_{xy}}{\mathrm{S}_{yy}}$, we need to prove $\mathrm{SS}_R = \hat{\beta}_F \mathrm{S}_{xy}$.

When $\lambda = 1$, $\hat{\beta}$ in Equation (3.11) becomes

$$\hat{\beta}_F = \frac{\left(\mathrm{S}_{yy} - \mathrm{S}_{xx}\right) + \left[\left(\mathrm{S}_{yy} - \mathrm{S}_{xx}\right)^2 + 4\mathrm{S}_{xy}^2\right]^{1/2}}{2\mathrm{S}_{xy}}$$

and $\hat{\beta}_F^2 = \frac{\left(\mathrm{S}_{yy} - \mathrm{S}_{xx}\right)\hat{\beta}_F + \mathrm{S}_{xy}}{\mathrm{S}_{xy}}$

From Equation (3.21), we have

$$\mathrm{SS}_R = \frac{\hat{\beta}_F^2 \left(\mathrm{S}_{yy} - \mathrm{S}_{xx}\right) + 2\hat{\beta}_F \mathrm{S}_{xy}}{1 + \hat{\beta}_F^2} = \hat{\beta}_F \mathrm{S}_{xy}$$

$$\Leftrightarrow \hat{\beta}_F^2 \left(\mathrm{S}_{yy} - \mathrm{S}_{xx}\right) = \hat{\beta}_F^3 \mathrm{S}_{xy} - \hat{\beta}_F \mathrm{S}_{xy}$$

$$\Leftrightarrow \hat{\beta}_F \left(\mathrm{S}_{yy} - \mathrm{S}_{xx}\right) = \left(\hat{\beta}_F^2 - 1\right)\mathrm{S}_{xy}$$

From the R.H.S.

$$\left(\hat{\beta}_F^2 - 1\right)\mathrm{S}_{xy} = \left[\frac{\left(\mathrm{S}_{yy} - \mathrm{S}_{xx}\right)\hat{\beta}_F + \mathrm{S}_{xy}}{\mathrm{S}_{xy}} - 1\right]\mathrm{S}_{xy}$$

$$= \left[\frac{\left(\mathrm{S}_{yy} - \mathrm{S}_{xx}\right)\hat{\beta}_F + \mathrm{S}_{xy} - \mathrm{S}_{xy}}{\mathrm{S}_{xy}}\right]\mathrm{S}_{xy}$$

$$= \hat{\beta}_F\left(\mathrm{S}_{yy} - \mathrm{S}_{xx}\right)$$

3.3.2 Relationship Between R_F^2 and R_S^2

Result 2. Let $\hat{\beta}_s$ and $\hat{\beta}_F$ be the slope estimator for the simple linear regression and the ULFR, respectively. The corresponding CODs are

$$R_S^2 = \frac{\hat{\beta}_S \mathrm{S}_{xy}}{\mathrm{S}_{yy}} \quad \text{and} \quad R_F^2 = \frac{\hat{\beta}_F \mathrm{S}_{xy}}{\mathrm{S}_{yy}} \quad (\text{when } \lambda = 1)$$

Then $0 \le R_S^2 \le R_F^2 \le 1$.

Proof. From the regression sum of squares, we obtain

$$0 \le \mathrm{SS}_R = \mathrm{S}_{yy} - \mathrm{SS}_E \le \mathrm{S}_{yy}$$

$$0 \le R_F^2 = \frac{\mathrm{SS}_R}{\mathrm{S}_{yy}} \le 1$$

we also know that simple linear regression has $0 \le R_S^2 \le 1$.

We now proof $R_S^2 \le R_F^2$ by contradiction. Assuming that

$$R_S^2 = \frac{\mathrm{S}_{xy}^2}{\mathrm{S}_{xx}\mathrm{S}_{yy}} > \frac{\left\{(\mathrm{S}_{yy} - \mathrm{S}_{xx}) + \sqrt{(\mathrm{S}_{yy} - \mathrm{S}_{xx})^2 + 4\mathrm{S}_{xy}^2}\right\}\mathrm{S}_{xy}}{2\mathrm{S}_{xy}\mathrm{S}_{yy}} = R_F^2$$

$$\frac{2\mathrm{S}_{xy}^2}{\mathrm{S}_{xx}} - \left(\mathrm{S}_{yy} - \mathrm{S}_{xx}\right) > \sqrt{\left(\mathrm{S}_{yy} - \mathrm{S}_{xx}\right)^2 + 4\mathrm{S}_{xy}^2}$$

$$\frac{4\mathrm{S}_{xy}^4}{\mathrm{S}_{xx}^2} + \left(\mathrm{S}_{yy} - \mathrm{S}_{xx}\right)^2 - \frac{4\mathrm{S}_{xy}^2\left(\mathrm{S}_{yy} - \mathrm{S}_{xx}\right)}{\mathrm{S}_{xx}} > \left(\mathrm{S}_{yy} - \mathrm{S}_{xx}\right)^2 + 4\mathrm{S}_{xy}^2$$

$$S_{xy}^4 - S_{xy}^2 S_{xx}\left(S_{yy} - S_{xx}\right) > S_{xy}^2 S_{xx}^2$$

$$S_{xy}^4 > S_{xy}^2 S_{xx}\left(S_{xx} + S_{yy} - S_{xx}\right)$$

$$S_{xy}^2 > S_{xx} S_{xy}$$

Also from $R_S^2 = \dfrac{S_{xy}^2}{S_{xx}S_{yy}} \le 1 \Rightarrow S_{xy}^2 \le S_{xx}S_{yy}$, this is a contradiction. Hence, we have $R_S^2 \le R_F^2$.

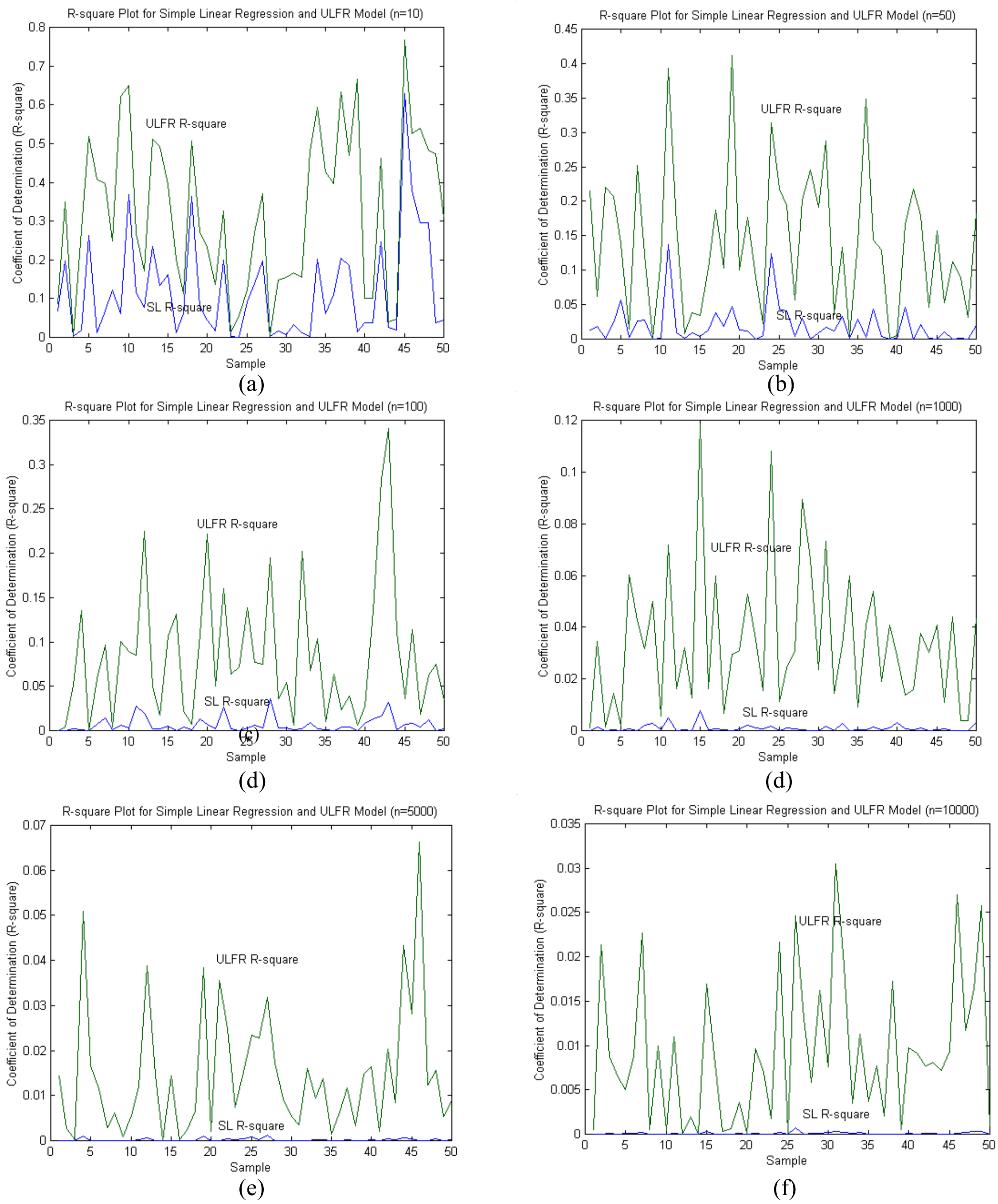

Figure 3.1: Simulation values for R_S^2 and R_F^2 with (a) n = 10, (b) n = 50, (c) n = 100, (d) n = 1000, (e) n = 5000, (f) n = 10000.

A simulation work is carried out to study the Result 2. The numerical example of this result is given in Figures 3.1(a) – 3.1(f). To explain the relationship between R_F^2 and R_S^2, the value of CODs are compared using different sample sizes, $n = 10, 50, 100, 1000, 5000, 10000$. For a given sample size, fifty samples were generated from a uniform distribution, each contains two random data sets (y_i, x_i). As an example, Figure 3.1(a) shows fifty values of COD obtained from simple linear regression and ULFR models with sample size $n = 10$. Figure 3.1(b) – Figure 3.1(f) show the same plot with different sample sizes. It is clearly shown that R_F^2 is always greater or equal to R_S^2. One interesting observation shown in Figure 3.1 is that both R_F^2 and R_S^2 decreased when the sample size increased.

Since R_F^2 is computed following the same method as R_S^2, some of the properties of R_S^2 may also remain for R_F^2. It is known that R_S^2 does not measure the appropriateness of the linear model (see e.g. Montgomery & Peck, 1992). This holds for R_F^2 too. As an example, when a nonlinear model has a large R_S^2 value, it is obvious that R_F^2 will also be large (see Result 2). More properties of R_F^2 are discussed in Chang et al. (2007) and will be further illustrated in Chapter 5 as a special case of multivariate version.

3.4 Discussion

This chapter explained the reasons why most of the conventional regression models are not suitable for solving the three issues stated in Chapter 2. The development of ULFR models as a potential performance indicator was also discussed. This ULFR model allows the two reference image and distorted image are both subject to errors, and hence provide a good solution to the need to consider an imperfect reference image. The derivation of COD for ULFR, labeled as R_F^2 was showed and it was used as an initial performance indicator in this study. The applications of R_F^2 will be discussed in the next chapters together with some selected ISMs. This ULFR model will be extended to a multidimensional version in the next chapter.