

## Using the Geographically Weighted Regression to Modify the Residential Flood Damage Function

L.F Chang,<sup>1</sup> and M.D. Su<sup>2</sup>

<sup>1</sup> Room 101, Water Management and Geographic Information Research Lab,  
Department of Bioenvironmental Systems Engineering, National Taiwan University,  
No.1, sec. 4, Rd. Roosevelt, Taipei City, Taiwan 10617, PH (886) 2-3366-3471;  
FAX (886) 2-2363-5854; email: [kiki@upland.ae.ntu.edu.tw](mailto:kiki@upland.ae.ntu.edu.tw)

<sup>2</sup> Room 101, Water Management and Geographic Information Research Lab,  
Department of Bioenvironmental Systems Engineering, National Taiwan University,  
No.1, sec. 4, Rd. Roosevelt, Taipei City, Taiwan 10617, PH (886) 2-3366-3451;  
FAX (886) 2-2363-5854; email: [sumd@ntu.edu.tw](mailto:sumd@ntu.edu.tw)

### ***Abstract.***

Flood damage functions are necessary to ensure comprehensive risk management. This study attempts to establish a residential flood damage function and explores residents living in the Keelung basin, where flood disasters occur frequently in Taiwan. Ordinary least squares (OLS) method is used to construct a flood damage function. Analytical results indicate that flood depth, is the significant variable, but the residual is non-stationary with spatial. The Geographically Weighted Regression (GWR) model is applied to modify the traditional regression model, which cannot capture spatial variations, and to solve the spatial non-stationary. Analytical results demonstrate significant spatial variation in the local parameter estimates for the variable flood depth and intercept. Therefore, a dummy variable, Zone, is added to the OLS model. The R-square value is found to increase from 0.15 to 0.24, and the residual is spatially stationary. In conclusion, the residential flood damage is determined by flood depth and zone, and the GWR model not only captures the spatial variations of the affecting factors, but also helps to discover the explanatory variable to modify the traditional regression model.

**Key words :** flood damage、flood depth、OLS、GWR、non- stationary

### ***Introduction***

Floods occur frequently in Taiwan because of its geographical location, climate and topography. Statistical data from the Ministry of the Interior show that Taiwan had an average of 4.47 typhoons and many storms annually from 1958 to 2004. These events have caused serious damage to agriculture, fishery, hydrologic engineering, houses, traffic facilities, electric power, telecommunications and economical activities. Risk management plays a very important role in mitigating the effects flood disasters, which cause damage to property and threaten lives. A complete flood management and mitigation system comprises a hydrological module for channel discharge calculation, an economic module for damage estimation, and a risk analysis process (Grigg, 1985). Although many studies have been performed in hydrology and hydraulics, few focus on flood damage in Taiwan (Chang, 2000). Since the level flood damage varies regionally, studies relating to other parts of the world cannot be applied directly to Taiwan. For those reasons, this study plans to establish the residential flood damage functions, and the result can be considered as the reference of regional risk management.

Flood damage function is traditionally estimated by an Empirical Depth-Damage Curve. The curve can be constructed in two ways (Kang *et al.*, 2005 ), from the investigation of damage after the disaster (TVA, 1969; FIA, 1970; Grigg and Helweg, 1975; Smith *et al.*, 1994 ; Su *et al.*, 2005), and from Synthesis (Chang, 2000; Chang and Su, 2001; Kang *et al.*, 2005). In the synthesis approach, data of the property items, possessive rate, and the height of the arrangement of the furniture are collected, and the possible damage of each item during flooding at each depth is investigated. These two methods are different in the way of establishing the curve, both assume that the flood depth is the only factor in the flood damage function.

Nevertheless, the flood depth may not be sufficient for consideration by a household flood damage function. McBean *et al.* (1988) pointed out that there were many factors besides flood depth could affect the flood damage, such as time of year of flooding, velocity of floodwaters, duration of flooding, sediment load and warning time, and therefore recommended adjusting the flood damage function should be adjusted. Yang L. *et al.* (2005) also noted that some meteorological, physiographic

and human factors such as rainfall, terrain and drainage could influence the actual flood damage. Hence, the relationships between various factors and flood damage are now widely examined. The most common factor being considered is the building type ( Grigg, 1974; FEMA, 1977; McBean *et al.*, 1988; Smith, 1994; Taiwan Water Resource Agency, 1997; Chang, 2000; Kang *et al.*, 2005 ). Other factors include area of main floor, family income ( McBean *et al.*, 1988), flood warning system ( Wind *et al.* , 1999; David, 2000 ), flood warning lead time ( Penning-Rowsell *et al.*, 2000), experience of flooding ( McPherson, 1977; McBean *et al.*, 1988; Wind 1999; Krasovskaia, 2001), the preparation before disaster ( Penning-Rowsell *et al.* , 2000 ), duration of flooding (McBean *et al.*, 1988 ; Torterotot *et al.* 1992; Hubert *et al.*, 1996 ), velocity of floodwaters ( CH2M Hill, 1974; Black, 1975; Smith, 1994; Beck *et al.*, 2002 ), number of people ( McBean *et al.*, 1988 ; Shaw, 2005 ) and the location of household ( Chang, 2000; Shaw, 2005 ). Since the flood damage is affected by many factors, some recently proposed multiple regression models for establishing the flood damage function incorporate all such factors ( Shaw, 2005 ). Although this approach can incorporate all factors as the predictors and raise the R-square value, it also increases the difficulty of predictor's data collection when predicting the damage. This model is a global multiple regression method, and assumes that the regression coefficient is constant across the study region (Platt, 2004). In other words, it does not consider the spatial variation, so the residuals from the global model often exhibit spatial autocorrelation (Fotheringham *et al.*, 2002). It violates the assumption of linear regression. Thus, the aim of this study is to establish the flood damage function for one household by using the smallest possible numbers of explained variables, while also considering the spatial variation and solving the residual with spatial autocorrelation problems.

This study first establishes the applied theoretical equations. Data sources are then discussed, and the studied areas are introduced. Finally, results are discussed and conclusions are drawn.

### ***Method***

The first step in establishing the flood damage function for one household by using the smallest numbers of explained variables, considering the spatial variation and then solving the residual with spatial autocorrelation problems, are to determine the factors causing flood damage. Many flood damage factors exist as described above, but the characteristics of flood damage vary among regions. Based on case studies in Taiwan, Shaw (2005) incorporated factors including flood depth, inundating time, building type, structure, the numbers of floors, presence of a basement, area, number of people and region. He demonstrated that the flood depth is the major factor affecting flood damage functions. Grigg (1996) noted that even without considering other factors, the flood depth – damage curve was still appropriate for estimating the flood damage. Therefore, this study determines the flood damage factor according to the flood depth, which is the most commonly considered factor in previous works.

This study first applies the OLS for global regression to establish the flood damage function. The Moran's I value is then used to proceed with the test of spatial residuals to check whether the residuals have spatial autocorrelation. If the residuals have spatial autocorrelation, then the GWR is applied and then test the significance of the spatial variability in the region.

If the coefficient exhibits significant spatial variation, then spatial grouping is performed; the dummy variable is added into the original global regression model, and finally the result is modified. The theoretical models used in this study are introduced as follows.

### **Global Regression Model**

A global regression model, calculated using OLS, is adopted to establish the flood damage function. Since flood damage increases with flood depth, the following S-curve model was constructed:

$$y = e^{(\beta_0 + \beta_1 / x)} + \varepsilon \quad (1)$$

where,  $y$  is the damage ( NT dollar ) ,  $x$  is the depth ( cm ) ,

$\beta_0$ 、 $\beta_1$  are the regression coefficients,  $\varepsilon$  is the residual

Formula (2) is the natural logarithm of formula (1)

$$\ln y = \beta_0 + \beta_1 \cdot \frac{1}{x} + \varepsilon_2 \quad (2)$$

where,  $y$  is the flood damage ( NT dollar )

$x$  is the flood depth ( cm )

$\varepsilon_2$  is the residual

Then,  $\beta_0$ 、 $\beta_1$  can estimated by simple linear regression model.

A basic assumption in fitting such a model is that the observations are independent of one another. A second assumption is that the structure of the model remains constant over the study area. That is, the estimated parameters have no local variations.

### **Residual spatial autocorrelation testing**

After establishing the regression model, the spatial autocorrelation coefficient, Moran's I, is computed to detect spatial autocorrelation in the residuals. According to the definition of the researchers (Bailey & Gatrell, 1995), Moran's I value can be indicated as

$$I = \frac{n \sum_{i=1}^n \sum_{j=1}^n w_{i,j} (y_i - \bar{y})(y_j - \bar{y})}{\left( \sum_{i \neq j} \sum w_{i,j} \right) \left( \sum_{j=1}^n (y_i - \bar{y})^2 \right)} \quad (3)$$

where  $n$  is the number of points or cells,  $y_m$  is the value in zone  $m$ ,  $\bar{y}$  is the mean of attribute  $y$ , and  $w_{ij}$  is the spatial proximity of point  $i$  and  $j$ . We often use the inverse of the distance between point  $i$  and  $j$ . This assumes that attribute values of points follow the first law of geography. With the inverse of the distance, we give smaller weights to points that are far apart and larger weights to points that are closer together. For example,  $w_{ij}$  can be defined as  $1/d_{ij}$ , where  $d_{ij}$  is the distance between point  $i$  and  $j$ . The expected value of Moran's I (i.e. the value that would be obtained if there

were no spatial pattern to data) is

$$E(I) = \frac{-1}{(n-1)} \quad (4)$$

with values of  $I$  larger than this indicating positive spatial autocorrelation (similar values cluster together) and the values below this indicating native spatial autocorrelation (similar values are dispersed).

Under this assumption, the variance of  $I$  is given by

$$\text{Var}(I) = \frac{n[(n^2 - 3n + 3)S_1 - nS_2 + 3S_0] - k[(n^2 - n)S_1 - 2nS_2 + 6S_0^2]}{(n-1)(n-2)(n-3)S_0^2} - E(I)^2 \quad (5)$$

where

$$S_0 = \sum_{i=1}^n \sum_{j=1}^n w_{ij}$$

$$S_1 = \sum_{i=1}^n \sum_{j=1}^n (w_{ij} + w_{ji})^2 / 2$$

$$S_2 = \sum_{i=1}^n (w_{j\cdot} + w_{\cdot i})^2$$

$$k = \frac{\left[ \sum_{i=1}^n (y_i - \bar{y})^4 / n \right]}{\left[ \sum_{j=1}^n (y_i - \bar{y})^2 / n \right]^2}$$

The distribution of  $I$  is asymptotically normal under assumption randomization. The standardized  $Z$  scores can be calculated as

$$Z(I) = \frac{I - E(I)}{S_{E(I)}} \quad (6)$$

$$S_{E(I)} = \text{SQRT} \left[ \frac{n^2 \sum_{ij} w_{ij}^2 + 3(\sum_{ij} w_{ij})^2 - n \sum_i (\sum_j w_{ij})^2}{(n^2 - 1)(\sum_{ij} w_{ij})^2} \right] \quad (7)$$

The null hypothesis is randomly distributed.

If  $Z(I) > 1.96$  or  $Z(I) < -1.96$ , then the residual is statistically significant with a statistical significance level of 5%. The residual pattern is clustered when  $Z(I) > 1.96$ . Conversely, the residual pattern is dispersed when  $Z(I) < -1.96$ . Alternatively, if  $-1.96 < Z(I) < 1.96$ , then the residual patterns not statistically significantly different from a random pattern, even if it looks somewhat clustered or visually dispersed.

### GWR Model

If the residual has spatial autocorrelation, then GWR can be utilized to modify and solve the problem (Brunsdon et al., 1996, 1998a,b, Fotheringham et al., 1996, 1997a,b, 1998, 2000, 2002, Platt, 2004). The modification of Formula (2) is

$$\ln y_i = \beta_0(u_i, v_i) + \beta_1(u_i, v_i) \cdot \frac{1}{x_i} + \varepsilon_i \quad (8)$$

Where

$y_i$  is the flood damage of point  $i$

$x_i$  is the flood depth of point  $i$

$(u_i, v_i)$  denotes the coordinates of the  $i$ th point in space

$\beta_0(u_i, v_i), \beta_1(u_i, v_i)$  is a realization of the continuous function at point  $I$

$\varepsilon_i$  is the residual of point  $(u_i, v_i)$

In simple linear regression model, a parameter is estimated for the relationship between each independent variable and dependent by OLS and the relationship is assumed to be constant across the study area. The estimator for it is

$$\beta = (X^T X)^{-1} X^T Y \quad (9)$$

The GWR model recognizes that spatial variations in relationships might exist. The GWR estimator is

$$\beta = (X^T W X)^{-1} X^T W Y \quad (10)$$

Where  $X$  is the matrix of the independent variable's observation value, which is the matrix of  $n \times 1$

$$X = \begin{bmatrix} x_1(u_1, v_1) \\ x_1(u_2, v_2) \\ \vdots \\ x_1(u_n, v_n) \end{bmatrix}$$

$\beta$  is the matrix of the regression coefficient, which is the matrix of  $n \times 2$

$$\beta = \begin{bmatrix} \beta_0(u_1, v_1) & \beta_1(u_1, v_1) \\ \beta_0(u_2, v_2) & \beta_1(u_2, v_2) \\ . & . \\ . & . \\ \beta_0(u_n, v_n) & \beta_1(u_n, v_n) \end{bmatrix}$$

$W$  is an  $n \times n$  matrix whose off-diagonal elements are zero and whose diagonal elements denote the geographical weighting of observed data for point  $i$ . That is

$$W = \begin{bmatrix} w_{i1}(u_1, v_1) & 0 & \dots & \dots & 0 \\ 0 & w_{i2}(u_2, v_2) & \dots & \dots & 0 \\ . & . & & & . \\ . & . & & & . \\ 0 & 0 & \dots & \dots & w_{in}(u_n, v_n) \end{bmatrix}$$

The weighting of each observed data is

$$w_{ij}(u_i, v_i) = \exp(-d_{ij}/h)^2 \quad (11)$$

$d_{ij}$  is the Euclidean distance between observed data  $i$  and  $j$

$h$  is the constant value of bandwidth

\*The bandwidth may be either supplied by the user, or estimated using a technique such as crossvalidation.

The parameter estimated with GWR is then plotted onto the map to determine the parameter variation within the region. Similarly, the standard error can be plotted onto the map to derive its variation over space. Finally, the Monte Carlo test is employed to determine whether any of the local parameter estimates are significantly non-stationary. If the test result is insignificant, then the parameter variations are due to chance.

### Modified global regression model

GWR analysis can not only modify the residuals of traditional regression with spatial autocorrelation, but also consider the spatial variation. Nevertheless, the result of GWR can obtain  $n$ th regressions more complicated than traditional global regression.

Therefore, the result of the GWR model is adopted to modify the traditional



regression model in three steps. First, plot the histogram of parameters with significant variation over space, and observe the distribution of each parameter. Next, perform spatial grouping based on the spatial distribution of each parameter. Finally, incorporate dummy variables into the original regression model. Determine the main variable by stepwise regression, and eliminate the insignificant variable. The model is thus modified as follows:

$$\ln y = \beta_0 + \beta_1 \cdot \frac{1}{x} + \sum_{i=2}^{m+1} \left( \beta_i \cdot GP_i + \beta_{i+1} \cdot \frac{1}{x} \cdot GP_i \right) + \varepsilon \quad (12)$$

where

$y$  is the flood damage ( NT dollar )

$x$  is the flood depth (cm)

GP is the dummy variable

$m$  is the numbers of spatial grouping

$\beta_0$ 、 $\beta_1$ 、 $\beta_i$  are the regression coefficients,  $\varepsilon$  is the residual.

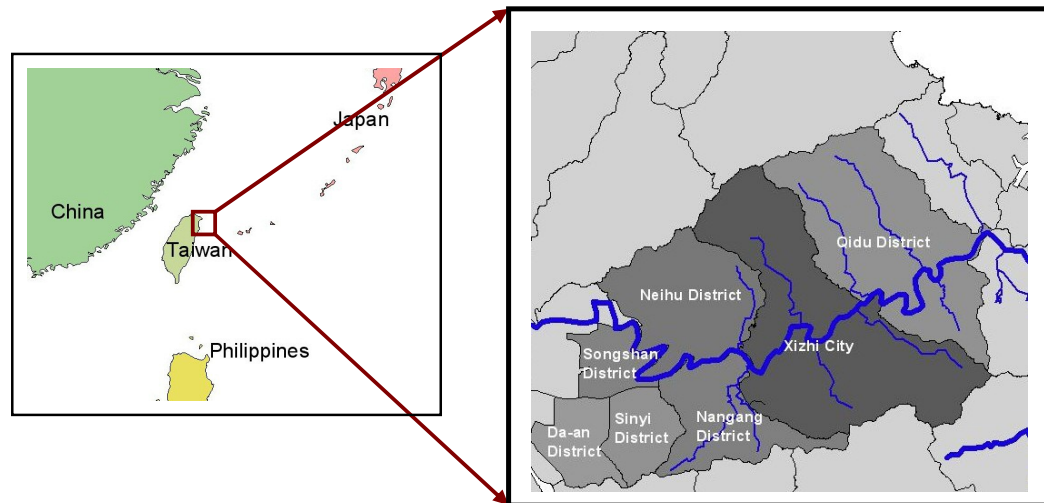
The modified regression model can not only consider the spatial variation of each parameter, but can also avoid residuals with spatial autocorrelation and too many regression equations.

### ***Data Collection and Studied Area***

To establish the flood damage function for one household in a residential area, the basin of the Keelung River, where flood disasters occur frequently, was selected as the studied area. The data of questionnaires from the flood damage caused by Nari Typhoon in 2001 were collected. Households with previous flood experience were explored. The investigated areas included Xizhi City, Qidu District, Nangang District, Neihs District, Songshan District, Sinyi District and Da-an District (as shown in Figure 1).

The questionnaire included questions on disaster scale (the flood depth and inundated time), level of damage (the damage of household furniture, decoration, and vehicles etc.), basic household information (the characteristics of the building like numbers of floors and area) and the risk perception factors (the flood experience, risk information, scale of the risk, influence of mass media, whether one is willing to take the risk or not, whether the risk is controllable or not, fear of

the risk). A total of 302 completed questionnaires were received. All data were geocoded to the map.



**Figure 1. Geographic distribution of study area in Taiwan**

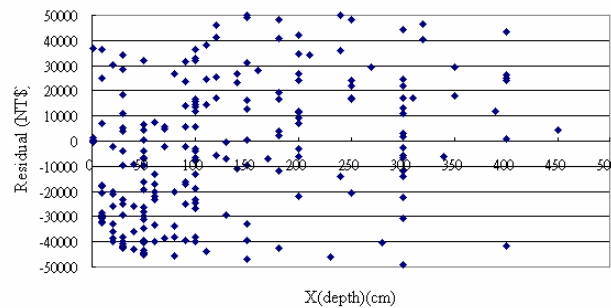
## **Result**

### **Global regression model**

The regression result of Formula (2) is shown in Table 1. The coefficient of determination  $R^2=0.15$ , the regression coefficient of intercept and flood depth were significantly different from zero (at 0.05 level). Figure 2 plots the residuals versus the predictor values. Because the points appear to scatter randomly about the line the mean of the residuals, all fundamental assumptions are correct. Initially, the residual was mapped to determine whether the residuals had spatial autocorrelation (as shown in Figure 3). The figure reveals that the residual spatial pattern was visually clustered, so Moran's I test was employed to test whether the residuals had spatial autocorrelation. The testing result demonstrates that the Moran's  $I = 0.6118$ , and  $(Z(I) = 4.936 > 1.96)$ , implying that the residuals had spatial autocorrelation, violating the assumption of linear regression. Therefore, GWR was applied to modify the model.

**Table 1. Global regression parameter estimates ( n=302 )**

Parameter	Estimate	Std Estimate	Std Err	T	P-Value
Intercept	10.825		0.092	118.041	0.000
1/X	-4.288	-0.386	0.591	-7.259	0.000

**Figure 2. Global model residual plot**

### GWR Model

The GWR model result indicates that  $R^2$  increased from 0.15 (OLS) to 0.26 (GWR), demonstrating that GWR provides a better explanatory ability than OLS. Figure 4 and 5 show the histogram and map of the intercept term from the GWR model. Figure 4 reveals that the values of the intercept can be divided into three groups, high, medium and low. Figure 5 depicts the spatial distribution of these three groups. The groups with high values were located on the northeast area. The groups with medium values were located on the middle area. The groups with low values were located on the western area. These findings indicate that the spatial pattern was clustered significantly. The intercept term indicates that flooding leads to fundamental flood damage. Fundamental flood damage is rising gradually from West to Northeast, according to the spatial distribution of the intercept.

Figure 6 displays the histogram of the inverse of flood depth variable-form GWR model, indicating that the coefficients can be divided into high and low. Figure 7 depicts the regression coefficients of the inverse flood depth variable. The figure shows that the groups with high values were located on the medium and western areas, and the groups with low values were located on the northeast area. These findings indicate that the spatial pattern was also clustered significantly.

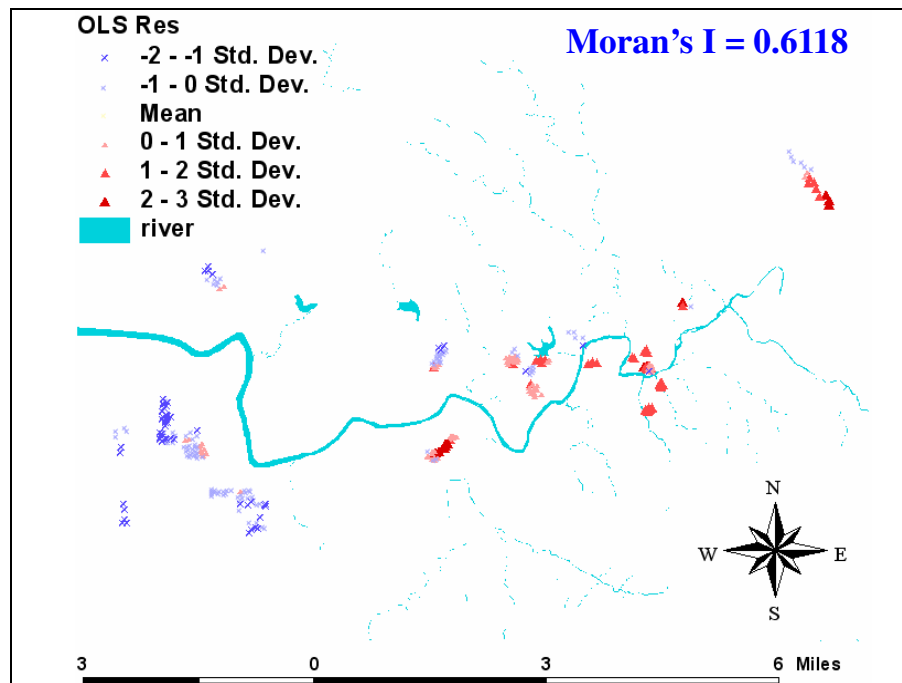


Figure 3. Global model residual surface

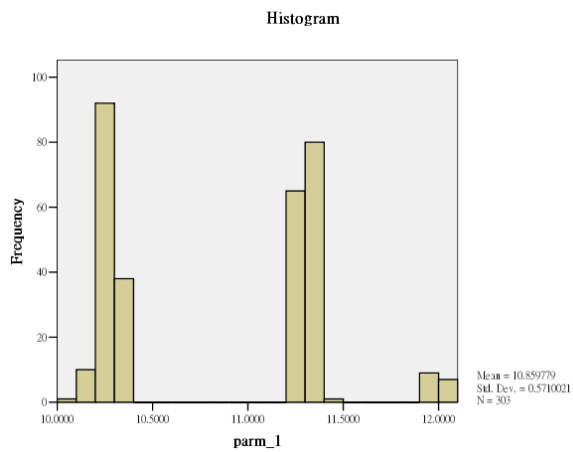


Figure 4. Histogram of the Intercept term from GWR model

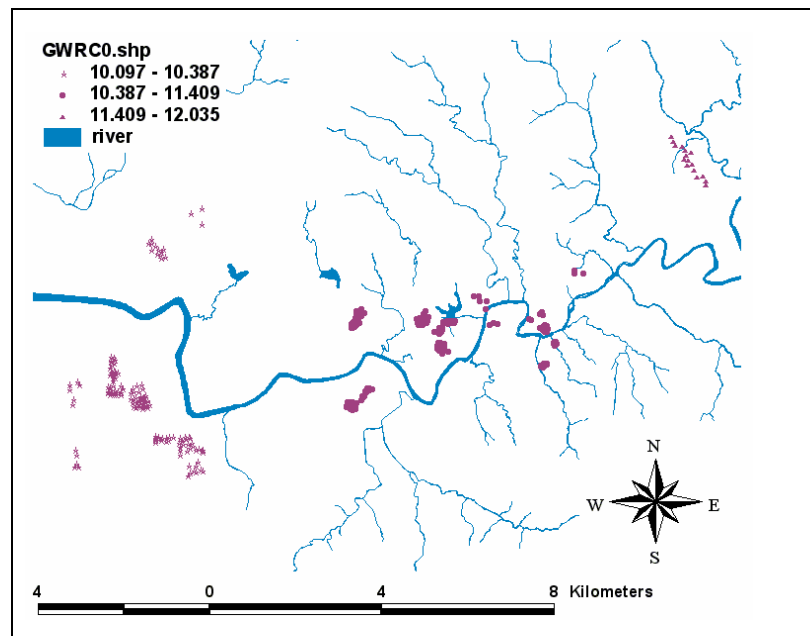


Figure 5. Map of the Intercept term from GWR model

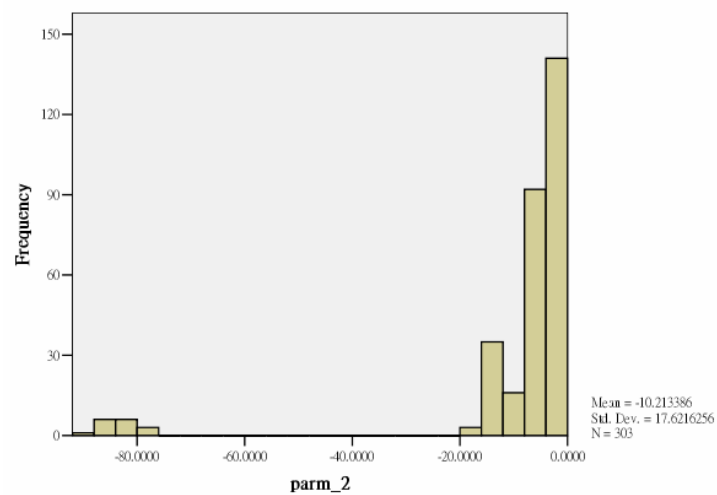
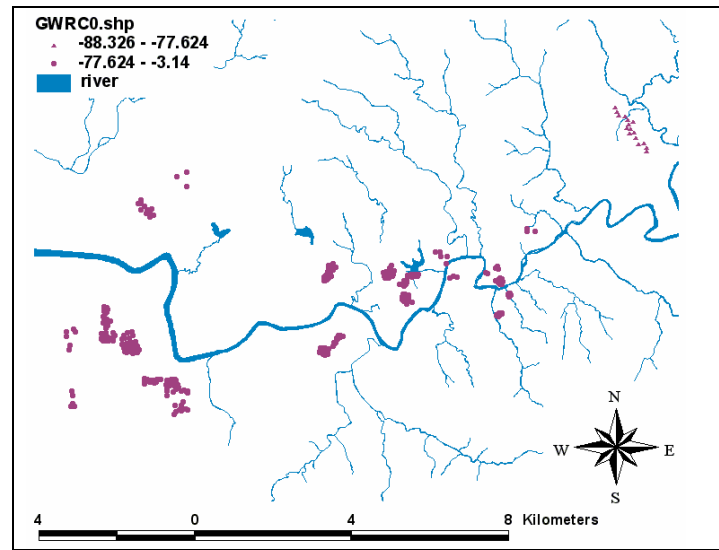


Figure 6: Histogram of the regression coefficients of the inverse of flood depth variable from GWR model



**Figure 7. Map of the Inverse of Flood Depth term from GWR model**

The regression coefficients of the inverse of flood depth variable indicate the change of the flood damage per flood depth. A greater value indicates a greater change in flood damage with increasing flood depth. The spatial distribution of the coefficients reveals that the value increased gradually from Northeast to West.

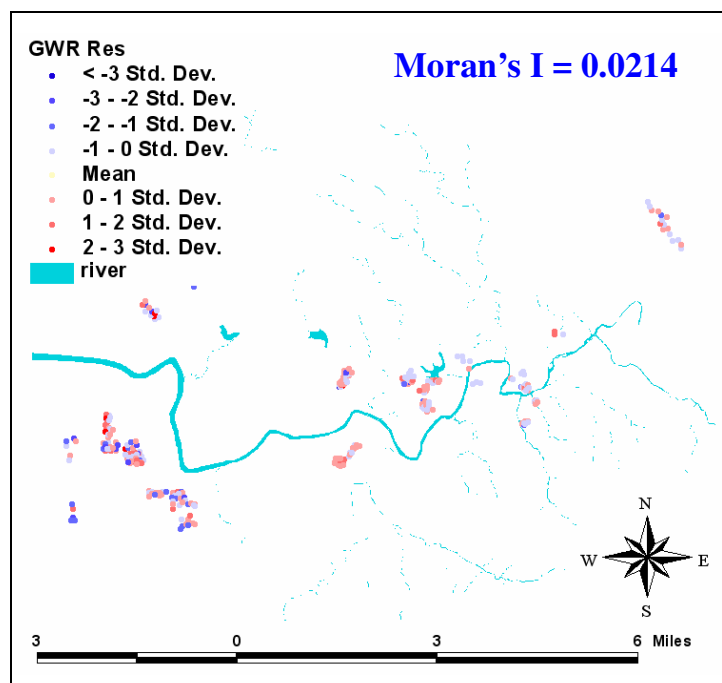
The residual of the GWR was mapped to identify any relationship between residual and spatial autocorrelation (as shown in Figure 8). The spatial pattern of residual is not clustered over space in visual. Further Moran's I was calculated to test whether the residuals had spatial autocorrelation. According to the test result,  $I = 0.0214$  (  $Z(I) = 0.216 < 1.96$  ), demonstrating that the residual with spatial autocorrelation was already modified.

Monte Carlo simulation was then used to determine whether each regression coefficient was spatially non-stationary (Table 2). Simulation results show that the regression coefficients of Intercept and  $1/X$  had significant spatial variation at the 1% level, meaning the two variables, Intercept and  $1/X$ , spatially affect the flood damage. This finding implies that the regression coefficients were not constant in the study region. Therefore, the GWR model is well suited to modifying the traditional regression model.

### **Modified global regression model**

The constant values in Fig. 5 were split into high, medium and low, and the

inverse flood depth's regression coefficients in Fig. 7 were divided into high and low respectively. In Table 3, 「\*」 denotes data, and 「N/A」 represents non-data. Analytical results show that all of the data could be divided into three groups. These three groups were then mapped, revealing that these three groups were spatially clustered (as shown in Figure 9). Therefore, spatial grouping was utilized to modify the original traditional model (OLS).



**Figure 8. Residuals from GWR model**

**Table 2. Results of Monte Carlo test for spatial non-stationary a (n=302 )**

	P-Value
Intercept	0.000***
1/X	0.000***

<sup>a</sup> Tests if regression coefficients change over space in a way that is unlikely to occur at random

\*\*\* = significant at .1% level

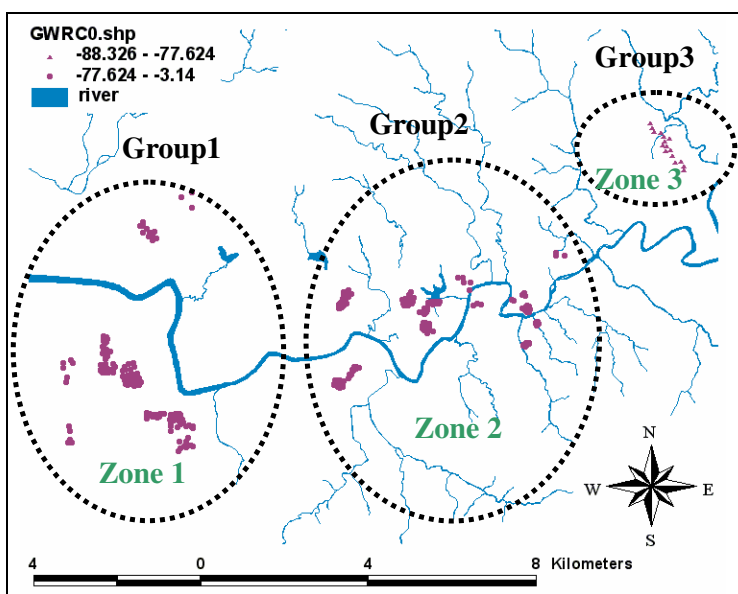
\*\* = significant at 1% level

\* = significant at 5% level

**Table 3. The Distribution of GWR's Regression Coefficient Values**

Parameter\Intercept	Low	Middle	High
Low	N/A	N/A	*Group3
High	*Group1	*Group2	N/A

「\*」 denotes data and 「N/A」 denotes non-data

**Figure 9. The Spatial Clustering**

The original traditional model (OLS) was modified according to the grouping result by adding two dummy variables, GP1 and GP2. The dummy variable GP1 = 1 when the location of household was within the Zone 1, and GP1 = 0 otherwise. The dummy variable GP2 = 1 when the household was within Zone 2, and GP2 = 0 otherwise.

The null hypothesis is that the regression coefficients of the constant and inverse of the flood depth will not change with different areas

The alternative hypothesis is that the regression coefficients of the constant and



inverse of the flood depth will change with different areas

Therefore, the equation can be modified as the follows.

$$\ln y = \beta_0 + \beta_1 \cdot \frac{1}{x} + \beta_2 \cdot GP1 + \beta_3 \cdot GP2 + \beta_4 \cdot \frac{1}{x} \cdot GP1 + \beta_5 \cdot \frac{1}{x} \cdot GP2 + \varepsilon \quad (13)$$

Where

$y$  is the flood damage

$x$  is the flood depth

$GP1 = 1$  when the location of household was within the Zone 1

$GP1 = 0$  otherwise

$GP2 = 1$  when the household was within Zone 2

$GP2 = 0$  otherwise

$\beta_0$ 、 $\beta_1$ 、 $\beta_2$ 、 $\beta_3$ 、 $\beta_4$ 、 $\beta_5$  are the regression coefficients,  $\varepsilon$  is the residual.

Stepwise regression was first adopted to determine the main variable. The calculation result reveals that only  $\frac{1}{x}$  and  $GP1$  were significant, meaning that the equation could be changed as

$$\ln y = \beta_0 + \beta_1 \cdot \frac{1}{x} + \beta_2 \cdot GP1 + \varepsilon \quad (14)$$

Equation (14) indicates that if the household was located in Zone 1, where  $GP1 = 1$ , then the equation could be presented as

$$\ln y = (\beta_0 + \beta_2) + \beta_1 \cdot \frac{1}{x} + \varepsilon \quad (15)$$

When the household was not located in Zone 1, where  $GP1 = 0$ , then the equation could be presented as

$$\ln y = \beta_0 + \beta_1 \cdot \frac{1}{x} + \varepsilon \quad (16)$$

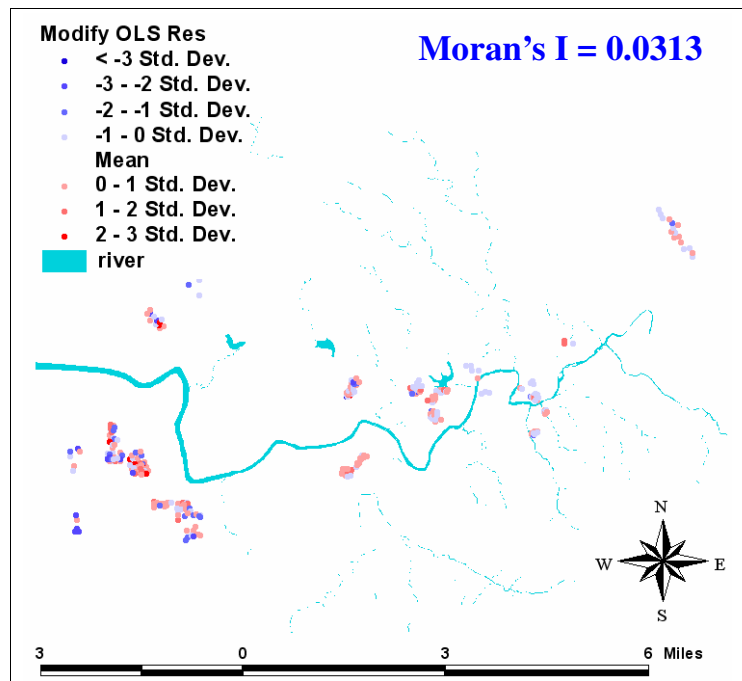
Table 4 presents the result of the modified regression model. Then focusing on the main variables and proceeds with regression analysis. The regression coefficients were statistically significant at a statistical significance level of 5%. The coefficient of determination  $R^2$  of the modified model increased from 0.15 (OLS) to 0.26 (modified OLS). This result was similar to that of the GWR model.

**Table 4. Result of modified global a regression model (n=302)**

Parameter	Estimate	Std Estimate	Std Err	T	P-Value
Intercept	1.327		0.113	100.457	0.000
1/X	-3.353	-3.302	0.567	-5.910	0.000
GP1	-1.161	-3.348	0.170	-6.819	0.000

<sup>a</sup> The average regression result of the whole studied area

To test if for spatial autocorrelation, the residual of the modified OLS was mapped to the map, and the Moran's I value was obtained (as shown in Fig. 10). Test results demonstrate that the Moran's I = 0.0313, and (  $Z(I) = 0.231 < 1.96$  ), indicating that the spatial pattern of residual was random. Therefore, the modified OLS did not violate the assumption of linear regression.

**Figure 10. Residual spatial distribution from modified regression model**

### Discussion

Figure 11 illustrates the result of the global regression equation, which demonstrates that the damage variation was greatest at low depths. The average total damage was \$50,000 per household.

Substituting the analytical result of modified global regression into Equations

(15) (16) yields then the flood damage functions for households is located in and outside Zone 1, respectively.

When the household is located in Zone 1, the flood damage function is:

$$\ln y = 0.166 - 3.353 \cdot \frac{1}{x} + \varepsilon \quad (17)$$

When the household is outside Zone 1, the flood damage function is:

$$\ln y = 1.327 - 3.353 \cdot \frac{1}{x} + \varepsilon \quad (18)$$

Where

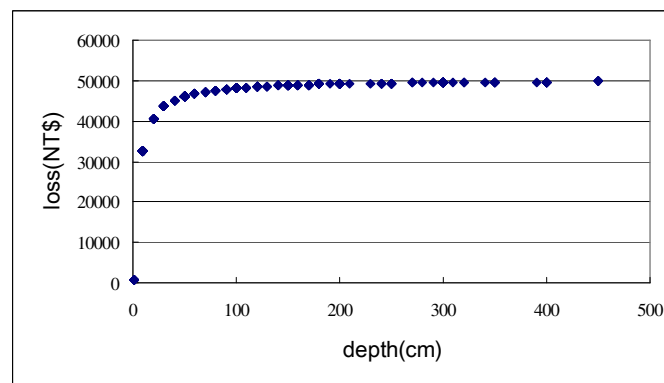
$y$  is the flood damage

$x$  is the flood depth

$\varepsilon$  is the residual

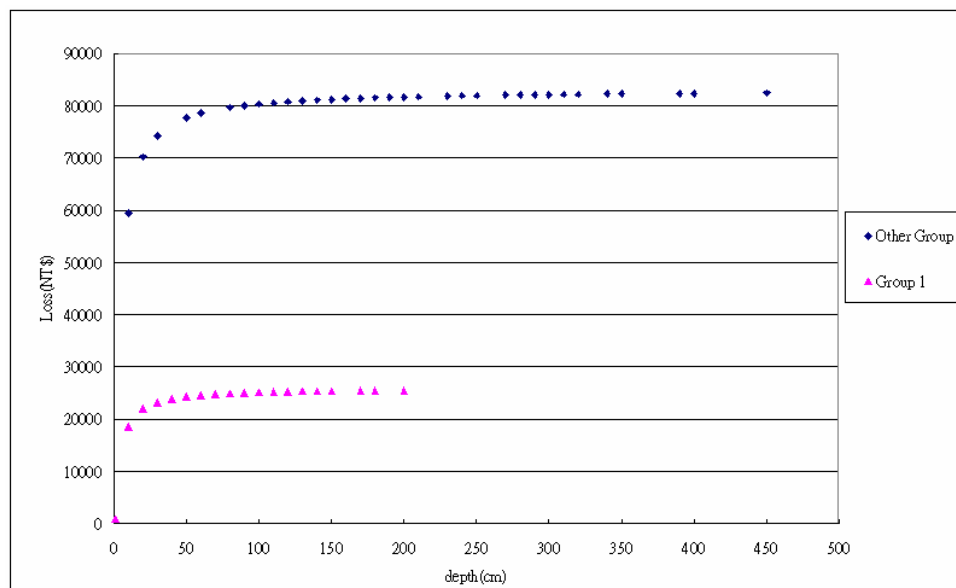
Equations (17) and (18) indicate that the flood damage function constant of a household located in Zone 1 is lower than that of a household located in the other Zones. This result shows the houses located outside Zone 1 would suffer greater fundamental sanitary damage than houses in Zone 1 when flood occurs.

Mapping the results of Equations (15) and (16) to the plot of flood damage versus flood depth (as displayed in Fig. 12) indicates that the flood damage in Zone 1 would be smaller than that in the other areas. The low water depth's tangent slopes outside Zone 1 would be greater than those within Zone 1, meaning the damage is most variable outside Zone 1.



**Figure 11. The Curve of Flood Depth Damage**

Figure 12 demonstrates that the total flood damage with flood depths in other areas of 100cm–330cm would be approximately \$80,000. Since Zone 1 did not have high water depth's data in this torrential rain, the total flood damage at a flood depth of 50–200cm would be approximately \$26,000. The results of this global regression were represented as a chart of flood damage versus flood depth, and then compared with those of Fig. 11. The result of the global regression's average total household damage was approximately \$50,000, which is in between the \$26,000–80,000 estimated by the modified regression model, because the global regression model did not consider the spatial variation, making the differences between areas undetectable.



**Figure 12. Flood Depth-Damage Curve of Zone 1 and Other region**

Furthermore, the damage function results of this study were compared with those estimated by Shaw *et al.* ( 2005 ), who utilized the following damage function:

$$\ln(\text{TLOSS}) = 1.541 + 0.389 \ln(\text{DEPTH}) + 0.739(\text{PRE}) - 6.510(\text{INS}) - 0.270(\text{OWN}) + 0.442(\text{BUILD})$$

$$+ 0.737(\text{EXP1}) + 0.845(\text{EXP2}) + 0.069(\text{EXP3}) - 0.025(\text{LIVY}) + 0.001(\text{IC}) + 0.380(\text{PEO})$$

(1.63)
(10.39)\*\*\*
(1.64)
(-3.29)\*\*\*
(-0.39)
(0.97)

(1.66)\*
(1.22)
(0.12)
(-1.86)\*
(0.50)
(0.97)

Where

TLOSS	is Total losses of physical property incurred by a household
DEPTH	is Inundation from the flood (cm)
PRE	Dummy=1 if adopting preparedness against floods
INS	Dummy=1 if purchasing flood insurance for the house or car
OWN	Dummy=1 if having the ownership of the house
BUILD	Dummy=1 if a single house; Dummy=0 if an apartment
IC	Household income (thousands of NT dollars)
PEO	The number of people in the household
EXP1	Dummy=1 if having one flooding experience in the past
EXP2	Dummy=1 if having two flooding experiences in the past
EXP3	Dummy=1 if having three or more flooding experiences in the past
LIVY	Years of living in the area

The total flood damage of one floor estimated by Shaw *et al.* ( 2005 ) , was approximately \$43,400, which is also between \$26,000 and \$80,000 as estimated by the proposed modified regression model.

For comparison, Kang *et al.* (2005) calculated the total flood damage for one household as approximately \$200,000, and Chang ( 2000 ) which the total flood damage for one household was approximately \$208,000. The present study produced lower a estimate than the above studies, probably because it defined flood damage differently from the others, and also considered the residents' preparedness and the actions immediately after flooding. Damage can be classified into three main types: capital restoration of affected households; replacement, in which new capital is utilized to replace the damaged capital, and directly providing the affected capital with the original services. This study adopts the flood damage definition of Kang *et al.* (2005):

$$d = w + \min( x, y, z )$$

where d denotes the Flood damage to capital; w denotes the value of services that cannot be provided after damage to capital from the day of flood to the day of restoration; x denotes the value of services that cannot be provided after damage to capital from the day of restoration; y denotes the cost of restoration or replacement following this restoration, and z denotes the value of services provided after damage.

Kang *et al.* (2005) and Chang (2000) assumed that all damaged capital was replaced with new capital. Additionally, both adopted the Synthesis approach, which

does not consider the residents' preparedness and the immediate actions to against the flood damage. Both factors can cause the over-estimation of damage values.

### ***Conclusion***

The proposed approach not only uses the smallest numbers of explained variables to establish the flood damage functions for one household, but also solves the problem in traditional regression models, which cannot consider spatial variation. Additionally, the proposed method modifies the residual with spatial autocorrelation.

In the coefficient of determination equation,  $R^2 = 0.15$  in the original OLS. The GWR equation not only considers the spatial variation, but also can increase the coefficient of determination 0.26. However, the proposed method has some limitations: because the characteristic of GWR equation is its  $n$ th data points have  $n$ th regressions, only flood damage of those  $n$ th coordinates can be forecast, and the use of space is irrational. Consequently, this study applies the result of GWR considering the spatial variation to proceed with the spatial clustering, and adopts dummy variables to modify the original OLS equation. This approach not only modifies the problem of the original OLS, which cannot consider the spatial variation, but also raises the modified coefficient of determination to 0.26.

This study not only finds a quantitative equation to describe the dependent variable, flood damage, as a function of the independent variable, flood depth, but also considers the spatial variation. The final analyzing results indicate that the flood damage to a household unit is mainly a factor of the flood depth and its located Zone. That is, the region should be considered with respect to the effect of the function coefficient.

### ***Acknowledgment***

The authors would like to thank the National Science Council of the Republic of China, Taiwan for financially supporting this research under Contract No. NSC\_93-2625-Z-002-035.

### ***Reference***

- BAILEY TC and GATRELL AC (1995) *Interactive spatial data analysis*. Wiley, New York
- BECK J, METZGER R, HINGRAY B, and Must A. (2002) Flood risk assessment based on security deficit analysis. Paper presented at the 27th General Assembly of the European Geophysical Society, *Geophysical Research*, Nice France, 21– 26 April 2002

- BLACK RD (1975) *Flood proofing rural structures: a 'project agnes' report, pennsylvania*. Final Report prepared for the United States Department of Commerce, Economic Development Administration. National Technical Information Service, Springfield, VA, USA, May 1975
- BRUNSDON C, FOTHERINGHAM AS, and CHARLTON ME (1996) Geographically weighted regression: a method for exploring spatial non-stationarity. *Geographical Analysis* 28(4):281-298
- BRUNSDON C, FOTHERINGHAM AS, and CHARLTON ME (1998a) Spatial non-stationarity and autoregressive models, *Environment and Planning A* 30(6):957-993
- BRUNSDON C, FOTHERINGHAM AS, and CHARLTON ME (1998b) Geographically weighted regression - modeling spatial non-stationarity, *Journal of the Royal Statistical Society, Series D-The Statistician* 47(3):431-443
- CHANG LF (2000) *Flood damage estimation for residential area. Dissertation*, National Taiwan University (in Chinese)
- CHANG LF, and SU MD (2001). Application of spatial data to damage estimations in flood. *Journal of Chinese Agricultural Engineering* 47(1):20-28 (in Chinese)
- CH2M HILL (1974) *Potential flood damages*. Willamette River System Department of the Army Portland District, Corps of Engineers, Portland, OR, USA
- DAVID TF (2001) Flood-warning decision-support system for sacramento, California. *Water Resources planning and management* 127(4):254-260
- DU PLESSIS LA (2002). A review of effective flood forecasting, warning and response system for application in South Africa. *WATER SA* 28 (2): 129-137 APR 2002
- FEMA (1977) *Reducing flood damage through building design: a guide manual - elevated residential structures*. F. E. M. Agency, ed.
- FIA (1970) *Flood hazard factors, depth-damage curves, elevation frequency curves, standard rate tables*. U. S. Federal Insurance Administration.
- FOTHERINGHAM AS, BRUNSDON C, and CHARLTON ME (2002) *Geographically weighted regression: the analysis of spatially varying relationships*. Wiley, Chichester
- FOTHERINGHAM AS, BRUNSDON C, and CHARLTON ME (2000) *Quantitative geography*. Sage, London
- FOTHERINGHAM AS, BRUNSDON C, and CHARLTON ME (1998) Geographically weighted regression: a natural evolution of the expansion method for spatial data analysis. *Environment and Planning A* 30(11):1905-1927
- FOTHERINGHAM AS, CHARLTON ME and BRUNSDON C (1997a) Two techniques for exploring non-stationarity in geographical data. *Geographical Systems* 4: 59-82.

- FOTHERINGHAM AS, CHARLTON ME and BRUNSDON C (1997b) *Measuring spatial variations in relationships with geographically weighted regression*. In: Fischer MM and Getis A. (eds.) Recent developments in spatial analysis, spatial statistics, Behavioral Modeling and Neurocomputing. Springer-Verlag, London, Chapter 4
- FOTHERINGHAM A.S., M.E. CHARLTON and BRUNSDON C (1996) The geography of parameter space: an investigation into spatial non-stationarity. *International Journal of Geographic Information Systems* 10: 605-627
- GRIGG NS and HEIWEG OJ (1974) *Estimating direct residential flood damage in urban areas*. Colorado State University
- GRIGG NS (1985) *Water resources planning*. McGraw-Hill, New York
- GRIGG NS and HEIWEG OJ (1975) State-of-the-art of estimating flood damage in urban areas. *Water Resources Bulletin* 11(2), 379-390
- GRIGG NS (1996) *Water Resources Management*. McGraw-Hill, New York
- HUBERT G, DEUTSCH JC, and DESBORDES MC (1996) *Policy decision support systems: modeling of rainfall flood damages*. In: PENNING-ROWSELL, E. (Ed.) Improving flood hazard management across Europe. European Union Environment Program, Contract Number EV5V-CT93-0296, EURO Flood II, Chapter 3
- KANG JL, SU MD, and CHANG LF (2005) Loss functions and framework for regional flood damage estimation in residential area. *Journal of Marine Science and Technology* 13(3) :193-199
- KRASOVSKAIA I (2001) Perception of the risk of flooding: the case of the 1995 flood in Norway. *Hydrological Sciences Journal-Journal Des Sciences Hydrologiques* 46(6): 855-868
- MCBEAN EA, GORRIE J, FORTIN M, DING J, and MOULTON R (1988) Adjustment factors for flood damage curves. *Journal of Water Resources Planning and Management* 114(6):635-646
- MCPHERSON HJ and SAARINEN TF (1977) Flood plain dwellers perception of flood hazard in Tucson. Arizona, *Annals of Regional Science* 11(2): 25-40
- PLATT RV (2004) Global and local analysis of fragmentation in a mountain region of Colorado, *Agriculture Ecosystems & Environment* 101(2-3): 207-218
- PENNING-ROWSELL EC, TUNSTALL SM, TAPSELL SM, and PARKER DJ (2000) The benefits of flood warnings: Real but elusive, and politically significant. *Journal of the Chartered Institution of Water and Environmental Management* 14(1): 7-14
- SHAW DG, HUANG HH, and HO MC (2005) Modeling flood loss and risk perception: the case of typhoon Nari in Taipei. *Proceedings of the Fifth Annual IIASA-DPRI Meeting on Integrated Disaster Risk Management: Innovations in Science and Policy*, Beijing, China, 13-18 September 2005



- SMITH DI (1994) Flood damage estimation - A review of urban stage-damage curves and loss function, *Water SA* 20(3):231-239
- SU MD, KANG JL, CHANG LF, and CHEN AS (2005) A grid-based GIS approach to regional flood damage assessment. *Journal of Marine Science and Technology* 13(3): 184-192
- TAIWAN WATER RESOURCE AGENCY (1997) *National flood insurance program pilot study: A case studies and for Tang-Dee-Yang area*. Taiwan project report, Taiwan (in Chinese).
- TORTEROTOT JP, KAUARK-LEITE LA, and ROCHE PA (1992) Analysis of individual real-time responses to flooding and influence on damage to households. Paper presented at the 3rd International Conference on *Floods and Flood Management*, Florence, Italy, 24– 26 November 1992.
- WIND HG, NIEROP TM, DE BLOIS CJ, and DE KOK JL (1999) Analysis of flood damages from the 1993 and 1995 Meuse floods. *Water Resources Research* 35(11): 3459-3465
- YANG L, ZUO C, and WANG YG (2005) An effective two-stage neural network model and its application on flood loss prediction. *Advances In Neural Networks - Isnn 2005, Pt 3, Proceedings*. Vol:3498, 1010-1016