CHAPTER V

FACTORS AFFECTING CD AND PB ACCUMULATION IN CULTIVATED SOIL AND VEGETABLES

This chapter aims to determine the factors affecting Pb and Cd accumulation in cultivated soil and the three selected vegetables. Four analytical models using factor analysis and multiple regression were built for this purpose. In particular, factor analysis was used to group the factors (independent variables) which are highly correlated with each other into groups (factor components). Factor score of the factor components resulting from the factor analysis used in multiple regressions to analyze and find the relation between component groups with dependent variable (accumulation of Pb and Cd in soil and vegetables). Coefficients of determination and statistical significance were used to conclude that which factors have strong influence and significance to accumulation of Pb and Cd in soil and vegetables in the research area.

5.1 Factors affecting lead (Pb) accumulation in vegetables

Factor analysis methods range from confirmatory techniques to pure exploratory procedures (Tucker and MacCallum, 1997). The exploratory factor analysis (EFA) seeks to uncover the underlying structure of a relatively large set of variables. There are no prior theory-driven hypotheses to confirm through the empirical data, and factor loadings are used to merely intuit the factor structure of the data. This is the most common form of factor analysis, where the researcher's as priori assumption is that any indicator may be associated with any factor. To conform with the notion by Costello and Osborne (2005), which concerns the most preferred type of EFA, the principal components analysis (PCA) with orthogonal (varimax) rotation was used as the method for the data analysis. Furthermore, the number of factors was decided to be retained for rotation on the basis of the Kaiser criterion, according to which all factors with Eigenvalues greater than one are feasible.

There are several views on the minimum number of cases required for the factor analysis. Hair *et al*. (2006) suggested that the minimum absolute sample size should be 50 observations and preferably the sample size should be 100 or larger. Generally, an adequate number of cases is suggested to vary between 100 and 300 (, 1983; Hatcher, 1994; Hutcheson and Sofroniou, 1999; Norušis, 2005). Also, as a general rule, the minimum is to have at least five times as many observations as the number of variables to be analyzed (Bryant and Yarnold, 1995; Hair *et al*., 2006). In this data consisting of 75 cases and 14 variables this subjects-to-variables ratio equals to 5.36. Thus, although the number of the cases is low compared to what is suggested by some authors, it seems that the results from this study based on PCA – in order to identify factors affecting lead (Pb) accumulation in vegetables – have sufficient explanatory power.

In this study, some important factors were taken into account as the independent variables to determine factors affecting Pb accumulation in vegetables. Table 5.1 showed the list of independent variables and dependent variable that used in the model of Pb accumulation in vegetables.

Table 5.1 Variables and measurements in model of lead (Pb) accumulation in vegetables

Variables		Measurement
X1=	Age of the vegetables grower	years
X2=	Schooling years	years
X3=	Family size	number of member
X4=	Total area of cultivated	m^2
X5=	Area of vegetable field	m^2
X6=	Vegetable growing experiences	number of years
X7=	Number sources of irrigation water	number
X8=	Farmer using irrigation water from public pond	Yes= 1, No= 0
X9=	Average amount of irrigation water/ m^2/day	liter/ m^2
X10=	Times of watering	times/ day
X11=	Farmer plant and produce vegetables following safe vegetable production process	Yes= 1, No= 0
X12=	Farmer know about problem accumulation of heavy metals in vegetable	Yes= 1, No= 0
X13=	Accumulation of lead (Pb) in cultivated soil	mg/kg
X14=	Accumulation of lead (Pb) in irrigation water	mg/liter
Dependent variables		
Y1=	Accumulation of lead (Pb) in vegetables	mg/kg

A factor analysis for the data (PCA with varimax rotation; values suppressed at the minimum loading 0.40) was conducted. The Kaiser-Meyer-Olkin (KMO) measure of sampling adequacy tests whether the partial correlations among variables are small. Described in other words, it is an index for comparing the magnitudes of the observed correlation coefficients to the magnitudes of the partial correlation coefficients. Large values for the KMO measure indicate that using a factor analysis for the variables is an adequate method, and a value greater than 0.50 is considered desirable (Malhotra and Birks, 2007). The Kaiser-Meyer-Olkin (KMO) measure of sampling adequacy is 0.627, and thus is greater than the recommended level. In addition, the Bartlett's test of sphericity is significant ($p<0.01$), indicating that the strength of the relationship among variables is strong. The Bartlett test of sphericity rejected the null hypothesis that the correlation matrix was an identity matrix. To conclude, Kaiser-Meyer-Olkin (KMO) measures and Bartlett's sphericity test showed that the sample met the criteria for factor analysis.

To be tested the multi-correlations among independent variables (from X1 to X14) with each other and between dependent variable (accumulation of Pb in vegetables) and independent variables, factor analysis was applied to group independent variables based on the phenomenon of multicollinearity. Factor analysis separated factor components which have higher value of variance, specifically rotation sums of squared loadings values greater than one (Table 5.2).

In order to obtain a clearer picture of factors affecting lead (Pb) accumulation in vegetables, a factor analysis of the correlation matrix was conducted. The resulting Eigenvalues for only the first five common factors were greater than unity (0.40). The same conclusion was reached by using the Scree plot (presented in the appendix) as a criterion to determine the number of group components. The proportion (percent) of variance explained by each component is also shown in Table 5.2, indicating that these five group components altogether account for about 64.45 percent of the total variance using the Principal Component method (first factor 16.64 percent, second factor 14.42 percent, third factor 12.72 percent, four factor 10.33 percent and fifth factor 10.32 percent) (Table 5.2). Specifically, component No. 1 has the highest variance which is about 16.64 percent of variance.

Results of the factor analysis after the orthogonal rotation revealed five underlying patterns that are identified as different group components of factors affecting lead (Pb) accumulation in vegetables. According to Costello and Osborne (2005) a factor with fewer than three items is generally considered weak and unstable, and five or more strongly loading items are desirable and indicate a solid factor. The factors from this empirical analysis seem to conform with this notion relatively well.

Table 5.2 Group of independent variables extracted by factor analysis

Components	Initial Eigenvalues		
	Total	% of Variance	Cumulative %
1	2.33	16.64	16.64
2	2.02	14.42	31.07
3	1.78	12.72	43.79
4	1.45	10.33	54.12
5	1.45	10.32	64.45

Result from the extraction method of principle component analysis (PCA) and rotation method of Varimax with Kaiser normalization to facilitate the interpretation of the results presented in Table 5.3.

Table 5.3 Rotated factors loading of independent variables

Variables	Rotated factor loadings				
	Component No.1	Component No.2	Component No.3	Component No.4	Component No.5
Pb in cultivated soil (X13)	0.81	-	-	-	-
Average amount irrigation water (X9)	0.78	-	-	-	-
Time of watering (X10)	0.72	-	-	-	-
Pb in irrigation water (X14)	0.60	-	-	-	-
Total areas (X4)	-	0.85	-	-	-
Field area (X5)	-	0.80	-	-	-
Family size (X3)	-	0.58	-	-	-
Age (X1)	-	-	0.83	-	-
Experiences (X6)	-	-	0.72	-	-
Schooling years (X2)	-	-	-0.67	-	-
Use public pond (X8)	-	-	-	-0.78	-
Following safe vegetable production process (X11)	-	-	-	-	0.73
Number source water (X7)	-	-	-	-	0.60
Farmer know heavy metal accumulation in vegetables (X12)	-	-	-	-	0.521

Note. Extraction Method: Principal Component Analysis.

Rotation Method: Varimax with Kaiser Normalization.

The first factor has relatively large weight for four variables in component No.1 which includes Pb in cultivated soil (X13), average amount irrigation water (X9), time of watering (X10), Pb in irrigation water (X14) which has the coefficients of 0.81, 0.78, 0.72 and 0.60 respectively (Table 5.3). They are variable relate to practices in farming using irrigation water and characteristic content of Pb in cultivated soil. Component No.2 consisted of total areas (X4), field area (X5), and family size (X3) which has the coefficients of 0.85, 0.80, and 0.58 respectively. Component No.3 consisted of age (X1), experiences (X6), schooling years (X2) which has the coefficients of 0.83, 0.72, and -0.67 respectively. Component No.4 consisted of use public pond (X8) which has the coefficient of -0.77. Component No.5 consisted of following safe vegetable production process (X11), number source water (X7), and farmer's understanding about heavy metal accumulation in vegetables (X12) which has the coefficients of 0.73, 0.60, and 0.52 respectively.

To determine the effect of the group components on Pb accumulation in vegetables, correlation matrix analysis was firstly used and multiple regression analysis was applied as the second step of analysis.

Correlation matrix analysis was used to investigate the multicollinearity of all independent variables (group components). Result from the analysis showed that Pb accumulation in vegetables (dependent variable) has a very strong positive correlation with component No.1 with coefficient 0.84 (Table 5.4), and the weak positive correlations with independent variables were component No.2, component No.3, component No.4 with the coefficients of 0.02, 0.01, and 0.14 respectively. Moreover, component No.5 was not correlated with independent variable with the coefficient 0.00.

Table 5.4 Correlation matrix of each component

Variables	Pb in vegetables	Component No.1	Component No.2	Component No.3	Component No.4	Component No.5
Pb in vegetables	1.00					
Component No.1	0.84	1.00				
Component No.2	0.02	0.00	1.00			
Component No.3	0.01	0.00	0.00	1.00		
Component No.4	0.14	0.00	0.00	0.00	1.00	
Component No.5	0.00	0.00	0.00	0.00	0.00	1.00

The relationship between accumulation of Pb in vegetables (Y1) (dependent variable) and all components (independent variable) showed in Table 5.5.

Table 5.5 Summary of the regression results of Pb accumulation model in vegetables

Variables	Three kind vegetables (n=75)			
	Coefficients	Standard Error	t-radio	Pro
Constant	1.54	0.08	19.93**	0.00
Component No.1	0.64	0.05	12.54**	0.00
Component No.2	0.05	0.05	1.06 ^{ns}	0.29
Component No.3	0.02	0.04	0.53 ^{ns}	0.60
Component No.4	0.07	0.05	1.37 ^{ns}	0.17
Component No.5	0.01	0.05	0.14 ^{ns}	0.89

R^2= 0.73 (73%) Adjusted R^2= 0.71 (71%)

F value 36.71 ** Durbin-Watson 2.19

Note: *, ** indicate the level of significance at 5% and 1% respectively.

Result from Table 5.5 showed that the constant and component No.1 were significant at 1 percent of level of significance while the other components were not statistical significant.

The coefficient of determination (R^2) statistic, adjusted for the number of parameters in the equation and the number of data observations. Analysis of multiple regression indicated that the coefficient of determination (R^2) value 0.71 can be regarded as a good fit in view of the cross-sectional data, indicating that about 71% of the total variation of Pb accumulation in vegetables explained by the independent variables and the remaining 29% may be due to error and other factors omitted in the model such as fertilizer factors. Therefore, the results of Pb accumulation in vegetables can be explained by five components about 71%. The significance of the model could reflect almost factors that affecting to Pb accumulation in vegetables.

The F value and probability statistics test the overall significance of the regression model. According to the F test, the independent variables in the estimated regression equation were highly significant with the dependent variable (Pb accumulation in vegetables) at 1% probability level.

Durbin-Watson statistic is used to detect the significant autocorrelation among the independent variables. Autocorrelation does not appear to be a problem as shown by the Durbin–Watson statistics with a coefficient of 2.19 (The 5% critical values for Durbin– Watson test when N= 75, K= 5 are d_L= 1.49 and d_u= 1.77. Thus d_u=1.77< d=2.19 <4-d_L=2.51). The null hypothesis stated that no autocorrelation was accepted, therefore, the autocorrelation does not appear in this estimation.

It can be concluded that Pb accumulation in vegetables is strongly and positive influenced by four variables in component No.1 which includes Pb in cultivated soil

(X13), average amount irrigation water (X9), time of watering (X10), Pb in irrigation water (X14). They are variables related to practices in farming using irrigation water and characteristic content of Pb in cultivated soil and Pb in irrigation water.

The component No.1 coefficient was statistically significant at the 1% level and had positive signs, which indicate that Pb accumulation in vegetables increase with the increase of four variables in component No.1. Results indicate a very strong relationship between Pb accumulation in vegetables and four variables in component No.1 with coefficient 0.64 (Table 5.5). This indicates that four independent variables of component No.1 will increase Pb accumulation in vegetables. Specifically, if four independent variables of component No.1 are increased by one unit, the Pb accumulation in vegetables will increase by 0.64 units. In other word, this result showed that an increase of Pb in cultivated soil (X13), average amount irrigation water (X9), time of watering (X10), Pb in irrigation water (X14) (component No.1) would lead to increase the Pb accumulation in vegetables.

The survey and research results in the year 2010 on the status of vegetable production and Pb accumulation in three studied vegetables showed that vegetables in this area were contaminated Pb with high level of accumulation (average 3.1 times more than the permitted standard). The main cause leading to the Pb accumulation in vegetable highly due to vegetables absorbed Pb from farming environment, especially from soil environment. Specifically, cultivated soil environment in the study area affected by the supplement of substances containing Pb into cultivated soil environment, including water contaminated by wastewater source from sewage sludge, residential waste water, industrial waste water, etc with frequency and large amounts of irrigation water etc (Survey result, 2010). In addition is the mobility of heavy metals (such as Pb) in the acidic cultivated soil environment.

This is also consistent with other studies (Devkota and Schmidt, 2000; Frost and Ketchum, 2000; Mangwayana, 1995) which concluded that vegetables uptake heavy metals contamination from factors such as using irrigation water contaminated by sewage sludge, residential waste water, industrial waste water etc; and depend on other factors such as concentrations of heavy metals in soils, nature of the soil on which vegetables are grown (Lake *et al*., 1984; Scott *et al*., 1996; Voutsa *et al*., 1996). Specifically, according to research result of Kachenko and Singh (2004), the heavy metal concentrations in soil have influenced the uptake of heavy metal in vegetables. Specifically, there is a significant correlation which was observed between soluble lead (Pb) from soil and lead (Pb) in some kinds of vegetables such as tomatoes, carrots.

In addition to, concentration of heavy metals accumulation in plants also depends on the ability to assimilate heavy metals of the plants, on pH of environment, amount of heavy metals in soil and irrigation water as well as plant species and different types of heavy metals, on organic matter in soil, ion exchange capacity, and clay composition (Lake *et al*., 1984; Scott *et al*., 1996; Voutsa *et al*., 1996). Concentration of heavy metals in plants also depends on the type of their compounds in soil and irrigation water (Jung, 2008).

5.2 Factors affecting lead (Pb) accumulation in cultivated soil

Factor analysis methods range from confirmatory techniques to pure exploratory procedures (Tucker and MacCallum, 1997). The exploratory factor analysis (EFA) seeks to uncover the underlying structure of a relatively large set of variables. There are no prior theory-driven hypotheses to confirm through the empirical data, and factor loadings are used to merely intuit the factor structure of the data. This is the most common form of factor analysis, where the researcher's as priori assumption is that any indicator may be associated with any factor. To conform with the notion by Costello and Osborne (2005), which concerns the most preferred type of EFA, the principal components analysis (PCA) with orthogonal (varimax) rotation was used as the method for the data analysis. Furthermore, the number of factors was decided to be retained for rotation on the basis of the Kaiser criterion, according to which all factors with Eigenvalues greater than one are feasible.

There are several views on the minimum number of cases required for the factor analysis. Hair *et al*. (2006) suggested that the minimum absolute sample size should be 50 observations and preferably the sample size should be 100 or larger. Generally, an adequate number of cases is suggested to vary between 100 and 300 (Gorsuch, 1983; Hatcher, 1994; Hutcheson and Sofroniou, 1999; Norušis, 2005). Also, as a general rule, the minimum is to have at least five times as many observations as the number of variables to be analyzed (Bryant and Yarnold, 1995; Hair *et al*., 2006). In this data consisting of 75 cases and 13 variables this subjects-to-variables ratio equals to 5.77. Thus, although the number of the cases is low compared to what is suggested by some authors, it seems that the results from this study based on PCA – in order to identify factors affecting lead (Pb) accumulation in cultivated soil – have sufficient explanatory power.

In this study, some important factors were taken into account as the independent variables to determine factors affecting Pb accumulation in cultivated soil. Table 5.6 showed the list of independent variables and dependent variable that used in the model of Pb accumulation in cultivated soil.

Table 5.6 Variables and measurements in model of lead (Pb) accumulation in cultivated soil

Independent variables	Measurement
X1= Age of the vegetables grower	years
X2= Schooling years	years
X3= Family size	number of member
X4= Total area of cultivated	m^2
X5= Area of vegetable field	m^2
X6= Vegetable growing experiences	number of years
X7= Number sources of irrigation water	number
X8= Farmer using irrigation water from public water pond	Yes= 1, No= 0
X9= Average amount of irrigation water/ m^2/day	liter/ m^2
X10= Times of watering	times/ day
X11= Farmer plant and produce vegetables following safe vegetable production process	Yes= 1, No= 0
X12= Farmer's understanding about heavy metal accumulation in vegetables	Yes= 1, No= 0
X13= Accumulation of lead (Pb) in irrigation water	mg/liter
Dependent variables	
Y2= Accumulation of lead (Pb) in cultivated soil	mg/kg

A factor analysis for the data (PCA with varimax rotation; values suppressed at the minimum loading 0.40) was conducted. The Kaiser-Meyer-Olkin (KMO) measure of sampling adequacy tests whether the partial correlations among variables are small. Described in other words, it is an index for comparing the magnitudes of the observed correlation coefficients to the magnitudes of the partial correlation coefficients. Large values for the KMO measure indicate that using a factor analysis for the variables is an adequate method, and a value greater than 0.50 is considered desirable (Malhotra and Birks, 2007). The Kaiser-Meyer-Olkin (KMO) measure of sampling adequacy is 0.582, and thus is greater than the recommended level. In addition, the Bartlett's test of sphericity is significant ($p<0.01$), indicating that the strength of the relationship among variables is strong. The Bartlett test of sphericity rejected the null hypothesis that the correlation matrix was an identity matrix. To conclude, Kaiser-Meyer-Olkin (KMO) measures and Bartlett's sphericity test showed that the sample met the criteria for factor analysis.

To be tested the multi-correlations among independent variables (from X1 to X13) with each other and between dependent variable (accumulation of Pb in cultivated soil) and independent variables, factor analysis was applied to group independent variables based on the phenomenon of multicollinearity. Factor analysis separated factor components which have higher value of variance, specifically rotation sums of squared loadings values greater than one (Table 5.7).

In order to obtain a clearer picture of factors affecting lead (Pb) accumulation in cultivated soil, a factor analysis of the correlation matrix was conducted. The resulting Eigenvalues for only the first five common factors were greater than unity (0.40). The same conclusion was reached by using the Scree plot (presented in the appendix) as a criterion to determine the number of group components. The proportion (percent) of variance explained by each component is also shown in Table 5.7, indicating that these five group components altogether account for about 65.30 percent of the total variance using the Principal Component method (first factor 15.26 percent, second factor 15.06 percent, third factor 13.57 percent, four factor 10.83 percent and fifth factor 10.58 percent). Specifically, component No. 1 has the highest variance which is about 15.26 percent of variance.

Results of the factor analysis after the orthogonal rotation revealed five underlying patterns that are identified as different group components of factors affecting lead (Pb) accumulation in cultivated soil. According to Costello and Osborne (2005) a factor with fewer than three items is generally considered weak and unstable, and five or more strongly loading items are desirable and indicate a solid factor. The factors from this empirical analysis seem to conform with this notion relatively well.

Table 5.7 Group of independent variables extracted by factor analysis

Components	Initial Eigenvalues		
	Total	% of Variance	Cumulative %
1	1.98	15.26	15.26
2	1.96	15.06	30.32
3	1.76	13.57	43.89
4	1.41	10.83	54.72
5	1.38	10.58	65.30

Result from the extraction method of principle component analysis (PCA) and rotation method of Varimax with Kaiser normalization to facilitate the interpretation of the results presented in Table 5.8. The first factor has relatively large weight for three variables in component No.1 which includes total areas (X4), field area (X5), family size (X3) which has the coefficients of 0.87, 0.79, and 0.56 respectively (Table 5.8). They are variables relate to size of vegetables field, size of total farming area owned by farmers and size of farmer's family. Component No.2 consisted of average amount irrigation water (X9), time of watering (X10), Pb in irrigation water (X13) which has the coefficients of 0.78, 0.69, and 0.68 respectively. They are variable relate to practices in farming using irrigation water and characteristic content of Pb in irrigation water. Component No.3 consisted of age (X1), experiences (X6), schooling years (X2) which has the coefficients of 0.84, 0.75, and -0.69 respectively. Component No.4 consisted of use public pond (X8), farmer know about problems of heavy metal accumulation in vegetables (X12) which has the coefficient of -0.64, and 0.64 respectively. Component No.5 consisted of following safe vegetable production process (X11), number source water (X7) which has the coefficients of 0.70, and 0.67 respectively.

Table 5.8 Rotated factors loading of independent variables

Variables	Rotated factor loadings				
	Component No.1	Component No.2	Component No.3	Component No.4	Component No.5
Total areas (X4)	0.87	-	-	-	-
Field area (X5)	0.79	-	-	-	-
Family size (X3)	0.56	-	-	-	-
Average amount irrigation water (X9)	-	0.78	-	-	-
Time of watering (X10)	-	0.69	-	-	-
Pb in irrigation water (X13)	-	0.68	-	-	-
Age (X1)	-	-	0.84	-	-
Experiences (X6)	-	-	0.75	-	-
Schooling years (X2)	-	-	-0.69	-	-
Use public pond (X8)	-	-	-	-0.64	-
Farmer heard heavy metal accumulation in vegetables (X12)	-	-	-	0.64	-
Following safe vegetable production process (X11)	-	-	-	-	0.70
Number source water(X7)	-	-	-	-	0.67

Note. Extraction Method: Principal Component Analysis.

Rotation Method: Varimax with Kaiser Normalization.

To determine the effect of the group components on Pb accumulation in cultivated soil, correlation matrix analysis was firstly used and multiple regression analysis was applied as the second step of analysis.

Five components and dependent variable (accumulation of Pb in cultivated soil (Y2)) was used to analyze by multiple regression analysis. Result from the analysis showed that Pb in cultivated soil (dependent variable) has a very strong correlation with component No.2 with coefficient 0.57 (Table 5.9). The weak positive correlations with independent variables were component No.1, component No.4, component No.5 with the coefficients of 0.08, 0.11, and 0.11 respectively. And component No.4 was a weak negative correlation with independent variable with the coefficients of -0.09. Therefore, from correlation matrix can concluded that component No.2 is the most favorable independent variables and very strong correlation with the dependent variable which is Pb accumulation in cultivated soil with coefficient 0.57.

Table 5.9 Correlation matrix of each component

Variables	Pb in cultivated soil	Component No.1	Component No.2	Component No.3	Component No.4	Component No.5
Pb in cultivated soil	1.00					
Component No.1	0.08	1.00				
Component No.2	0.57	0.00	1.00			
Component No.3	-0.09	0.00	0.00	1.00		
Component No.4	0.11	0.00	0.00	0.00	1.00	
Component No.5	0.11	0.00	0.00	0.00	0.00	1.00

The relationship between accumulation of Pb in cultivated soil (Y2) (dependent variable) and all components (independent variable) showed in Table 5.10.

Table 5.10 Summary of the regression results of Pb accumulation model in cultivated soil

Variables	Three kind vegetables (n=75)			
	Coefficients	Standard Error	t-radio	Pro
Constant	65.46	3.21	20.39 **	0.00
Component No.1	5.49	2.03	2.71**	0.00
Component No.2	11.91	2.11	5.65**	0.00
Component No.3	-1.67	1.91	-0.94^{ns}	0.35
Component No.4	3.33	2.02	1.74^{ns}	0.08
Component No.5	1.77	0.05	0.87^{ns}	0.38

R^2= 0.37 (37%) Adjusted R^2= 0.32 (32%)

F value 8.03 *** Durbin-Watson 2.11

Note: *, ** indicate the level of significance at 5% and 1% respectively.

Result from Table 5.10 showed that the constant, component No.1 and component No.2 were significant at 1 percent of level of significance while the other components were not statistical significant.

The coefficient of determination (R^2) statistic, adjusted for the number of parameters in the equation and the number of data observations. Analysis of multiple regression indicated that the coefficient of determination (R^2) value 0.32 can be regarded as a good fit in view of the cross-sectional data, indicating that about 32% of the total variation of Pb accumulation in cultivated soil explained by the independent variables and the remaining 68% may be due to error and other factors omitted in the model such as pH factor. Therefore, the results of Pb accumulation in cultivated can be explained by five components about 32%. The model has low significance, therefore the significance of the model could not reflect almost factors that affecting to Pb accumulation in cultivated soil.

The F value and probability statistics test the overall significance of the regression model. According to the F test, the independent variables in the estimated regression equation were highly significant with the dependent variable (Pb accumulation in cultivated soil) at 1% probability level.

Durbin-Watson statistic is used to detect the significant autocorrelation among the independent variables. Autocorrelation does not appear to be a problem as shown by the Durbin–Watson statistics with a coefficient of 2.11 (The 5% critical values for Durbin–Watson test when N= 75, K= 5 are d_L= 1.49 and d_u= 1.77. Thus $d_u=1.77 < d=2.11 < 4-d_L=2.51$). The null hypothesis stated that no autocorrelation was accepted, therefore, the autocorrelation does not appear in this estimation.

It can be concluded that Pb accumulation in cultivated soil is strongly and positively influenced by three variables in component No.2, which includes average amount irrigation water (X9), time of watering (X10), Pb in irrigation water (X13). They are variables related to practices in farming using irrigation water and characteristic content of Pb in irrigation water.

In addition, from the analysis also concluded that Pb accumulation in cultivated soil is very weak positive influenced by three variables in component No.1 which includes total areas (X4), field area (X5), family size (X3). They are variables relate to size of vegetables field, size of total farming area owned by farmers and size of farmer's family.

The component No.2 coefficient was statistically significant at the 1% level and had positive signs, which indicates that Pb accumulation in cultivated soil increase with the increase of three variables in component No.2. Results indicate a very strong relationship between Pb accumulation in cultivated soil and three variables in component No.2 with coefficient 0.57 (Table 5.9). This indicates that three independent variables of component No.2 will increase Pb accumulation in cultivated soil. Specifically, if three independent variables of component No.2 are increased by one unit, the Pb accumulation in cultivated soil will increase by 11.91 units. In other word, this result showed that an increase of average amount irrigation water (X9), time of watering (X10), Pb in irrigation water (X13) (component No.2) would lead to increase the Pb accumulation in cultivated soil.

According to the survey and research results in 2010 on the status of vegetable production and Pb accumulation in cultivated soil in the research area showed that the cultivated soil in this area had signs of contaminated Pb with high level of Pb content in many soil samples although average Pb accumulation in cultivated soil still within the permitted standard. Pb accumulation in cultivated has increase trend by the years as results of comparing results of the survey (2010) with research results of Hang (2007).

According to research results of the Hang (2007), vegetable growing soil in Thai Nguyen city in general and vegetable land at Tuc Duyen ward in particular heavy metal concentrations in soil are still eligible for the production. In many cultivated soil samples, Pb content in soil samples are very low and lower than the permitted standard. The concentration of Pb in cultivated soil is average from 0.024-0.9672 mg/kg (Hang, 2007). Comparing with study results 2010, Pb content in cultivated soil was close to the threshold content (average content of Pb is 65.49 mg/kg), therefore it seem that Pb accumulation in cultivated had tended to increase rapidly. Pb accumulation in cultivated soils has increased after many years of cultivation.

The main cause leading to the Pb accumulation in cultivated soil is due to cultivated soil absorbed and accumulated Pb from substances added into cultivated soil environment, especially substances from irrigation water sources. Specifically, cultivated soil environment in the study area affected by the supplement of substances containing Pb into cultivated soil environment, including irrigation water contaminated by wastewater sources from sewage sludge, residential waste water, industrial waste water, etc with frequency and large amounts of irrigation water etc (Survey result, 2010). In addition is the mobility of heavy metals (such as Pb) in the acidic cultivated soil environment.

5.3 Factors affecting cadmium (Cd) accumulation in vegetables

Factor analysis methods range from confirmatory techniques to pure exploratory procedures (Tucker and MacCallum, 1997). The exploratory factor analysis (EFA) seeks to uncover the underlying structure of a relatively large set of variables. There are no prior theory-driven hypotheses to confirm through the empirical data, and factor loadings are used to merely intuit the factor structure of the data. This is the most common form of factor analysis, where the researcher's as priori assumption is that any indicator may be associated with any factor. To conform with the notion by Costello and Osborne (2005), which concerns the most preferred type of EFA, the principal components analysis (PCA) with orthogonal (varimax) rotation was used as the method for the data analysis. Furthermore, the number of factors was decided to be retained for rotation on the basis of the Kaiser criterion, according to which all factors with Eigenvalues greater than one are feasible.

There are several views on the minimum number of cases required for the factor analysis. Hair *et al*. (2006) suggested that the minimum absolute sample size should be 50 observations and preferably the sample size should be 100 or larger. Generally, an adequate number of cases is suggested to vary between 100 and 300 (Gorsuch, 1983; Hatcher, 1994; Hutcheson and Sofroniou, 1999; Norušis, 2005). Also, as a general rule, the minimum is to have at least five times as many observations as the number of variables to be analyzed (Bryant and Yarnold, 1995; Hair *et al*., 2006). In this data consisting of 75 cases and 14 variables this subjects-to-variables ratio equals to 5.36. Thus, although the number of the cases is low compared to what is suggested by some authors, it seems that the results from this study based on PCA – in order to identify factors affecting lead (Pb) accumulation in vegetables – have sufficient explanatory power.

In this study, some important factors were taken into account as the independent variables to determine factors affecting Cd accumulation in vegetables. Table 5.11 showed the list of independent variables and dependent variable that used in the model of Cd accumulation in vegetables.

Table 5.11 Variables and measurements in model of cadmium (Cd) accumulation in vegetables

	Variables	Measurement
X1=	Age of the vegetables grower	years
X2=	Schooling years	years
X3=	Family size	number of member
X4=	Total area of cultivated	m^2
X5=	Area of vegetable field	m^2
X6=	Vegetable growing experiences	number of years
X7=	Number sources of irrigation water	number
X8=	Farmer using irrigation water from public pond	Yes= 1, No= 0
X9=	Average amount of irrigation water/ m^2/day	liter/ m^2
X10=	Times of watering	times/ day
X11=	Farmer plant and produce vegetables following safe vegetable production process	Yes= 1, No= 0
X12=	Farmer know about problem accumulation of heavy metals in vegetable	Yes= 1, No= 0
X13=	Accumulation of cadmium (Cd) in cultivated soil	mg/kg
X14=	Accumulation of cadmium (Cd) in irrigation water	mg/liter
Dependent variables		
Y3=	Accumulation of cadmium (Cd) in vegetables	mg/kg

A factor analysis for the data (PCA with varimax rotation; values suppressed at the minimum loading 0.40) was conducted. The Kaiser-Meyer-Olkin (KMO) measure of sampling adequacy tests whether the partial correlations among variables are small. Described in other words, it is an index for comparing the magnitudes of the observed correlation coefficients to the magnitudes of the partial correlation coefficients. Large values for the KMO measure indicate that using a factor analysis for the variables is an adequate method, and a value greater than 0.50 is considered desirable (Malhotra and Birks, 2007). The Kaiser-Meyer-Olkin (KMO) measure of sampling adequacy is 0.547, and thus is greater than the recommended level. In addition, the Bartlett's test of sphericity is significant ($p<0.01$), indicating that the strength of the relationship among variables is strong. The Bartlett test of sphericity rejected the null hypothesis that the correlation matrix was an identity matrix. To conclude, Kaiser-Meyer-Olkin (KMO) measures and Bartlett's sphericity test showed that the sample met the criteria for factor analysis.

To be tested the multi-correlations among independent variables (from X1 to X14) with each other and between dependent variable (accumulation of Cd in vegetables) and independent variables, factor analysis was applied to group independent variables based on the phenomenon of multicollinearity. Factor analysis separated factor components which have higher value of variance, specifically rotation sums of squared loadings values greater than one (Table 5.12).

In order to obtain a clearer picture of factors affecting cadmium (Cd) accumulation in vegetables, a factor analysis of the correlation matrix was conducted. The resulting Eigenvalues for only the first five common factors were greater than unity (0.40). The same conclusion was reached by using the Scree plot (presented in the appendix) as a criterion to determine the number of group components. The proportion (percent) of variance explained by each component is also shown in Table 5.12, indicating that these five group components altogether account for about 60.32 percent of the total variance using the Principal Component method (first factor 15.37

percent, second factor 12.80 percent, third factor 12.41 percent, four factor 10.14 percent and fifth factor 9.61 percent) (Table 5.12). Specifically, component No. 1 has the highest variance which is about 15.37 percent of variance.

Results of the factor analysis after the orthogonal rotation revealed five underlying patterns that are identified as different group components of factors affecting cadmium (Cd) accumulation in vegetables. According to Costello and Osborne (2005) a factor with fewer than three items is generally considered weak and unstable, and five or more strongly loading items are desirable and indicate a solid factor. The factors from this empirical analysis seem to conform with this notion relatively well.

Table 5.12 Group of independent variables extracted by factor analysis

Components	Initial Eigenvalues		
	Total	% of Variance	Cumulative %
1	2.15	15.37	15.37
2	1.79	12.80	28.16
3	1.74	12.41	40.57
4	1.42	10.14	50.71
5	1.35	9.61	60.32

Result from the extraction method of principle component analysis (PCA) and rotation method of Varimax with Kaiser normalization to facilitate the interpretation of the results presented in Table 5.13.

Table 5.13 Rotated factors loading of independent variables

Variables	Rotated factor loadings				
	Component No.1	Component No.2	Component No.3	Component No.4	Component No.5
Total areas (X4)	0.88	-	-	-	-
Field area (X5)	0.77	-	-	-	-
Age (X1)	-	0.82	-	-	-
Experiences (X6)	-	0.77	-	-	-
Schooling years (X2)	-	-0.62	-	-	-
Average amount irrigation water (X9)	-	-	0.74	-	-
Time of watering (X10)	-	-	0.63	-	-
Cd in cultivated soil (X14)	-	-	0.49	-	-
Cd in irrigation water (X13)	-	-	0.48	-	-
Number source water (X7)	-	-	-	0.76	-
Following safe vegetable production process (X11)	-	-	-	0.59	-
Farmer know about heavy metal accumulation in vegetables (X12)	-	-	-	-	0.68
Use public pond (X8)		-	-	-	-0.57
Family size (X3)		-	-	-	-0.54

Note. Extraction Method: Principal Component Analysis.

Rotation Method: Varimax with Kaiser Normalization.

The first factor has relatively large weight for four variables in component No.1 which includes total areas (X4), field area (X5) which has the coefficients of 0.88, and 0.77 respectively (Table 5.13). Component No.2 consisted of age (X1), experiences (X6), schooling years (X2) which has the coefficients of 0.82, 0.77, and -0.62 respectively. Component No.3 consisted of average amount irrigation water (X9), time of watering (X10), Cd in cultivated soil (X14), Cd in irrigation water (X13) which has the coefficients of 0.74, 0.63, 0.49, and 0.48 respectively. They are variable relate to practices in farming using irrigation water and characteristic content of Cd in cultivated soil and Cd in irrigation water. Component No.4 consisted of number source water (X7), following safe vegetable production process (X11) which has the coefficient of 0.76, and 0.59 respectively. Component No.5 consisted of farmer's understanding about heavy metal accumulation in vegetables (X12), use public pond (X8), family size (X3) which has the coefficients of 0.68, -0.57, and -0.54 respectively.

To determine the effect of the group components on Cd accumulation in vegetables, correlation matrix analysis was firstly used and multiple regression analysis was applied as the second step of analysis.

Correlation matrix analysis was used to investigate the multicollinearity of all independent variables (group components). Result from the analysis showed that Cd accumulation in vegetables (dependent variable) has a strong positive correlation with component No.3 with coefficient 0.47, and the weak positive correlations with independent variables were component No.2 and component No.5 with the coefficients of 0.11, and 0.13 respectively (Table 5.14). Besides, Cd in vegetables (dependent variable) has a weak negative correlations with independent variables were component No.1 and component No.5 with the coefficients of -0.39, and -0.23 respectively.

Table 5.14 Correlation matrix of each component

Variables	Cd in vegetables	Component No.1	Component No.2	Component No.3	Component No.4	Component No.5
Cd in vegetables	1.00					
Component No.1	-0.39	1.00				
Component No.2	0.11	0.00	1.00			
Component No.3	0.47	0.00	0.00	1.00		
Component No.4	-0.23	0.00	0.00	0.00	1.00	
Component No.5	0.13	0.00	0.00	0.00	0.00	1.00

The relationship between accumulation of Cd in vegetables (Y3) (dependent variable) and all components (independent variable) showed in Table 5.15.

Table 5.15 Summary of the regression results of Cd accumulation model in vegetables

Variables	Three kind vegetables (n=75)			
	Coefficients	Standard Error	t-radio	Pro
Constant	0.13	0.01	25.72 **	0.00
Component No.1	-0.02	0.01	-4.37**	0.00
Component No.2	0.01	0.01	1.22 ns	0.23
Component No.3	0.03	0.01	5.33 **	0.00
Component No.4	-0.01	0.01	-2.54 **	0.01
Component No.5	0.01	0.01	1.49 ns	s0.14

R^2= 0.46 (46%) Adjusted R^2 = 0.42 (42%)

F value 11.523 ** Durbin-Watson 2.03

Note: *, ** indicate the level of significance at 5% and 1% respectively.

Result from Table 5.15 showed that the constant, component No.1, component No.3 and component No.4 were significant at 1 percent of level of significance while the other components were not statistical significant.

The coefficient of determination (R^2) statistic, adjusted for the number of parameters in the equation and the number of data observations. Analysis of multiple regression indicated that the coefficient of determination (R^2) value 0.42 can be regarded as a good fit in view of the cross-sectional data, indicating that about 42% of the total variation of Cd accumulation in vegetables explained by the independent variables and the remaining 58% may be due to error and other factors omitted in the model such as fertilizer factors. Therefore, the results of Cd accumulation in vegetables can be explained by five components about 42%. The significance of the model could not reflect factors that affecting to Cd accumulation in vegetables.

The F value and probability statistics test the overall significance of the regression model. According to the F test, the independent variables in the estimated regression equation were highly significant with the dependent variable (Cd accumulation in vegetables) at 1% probability level.

Durbin-Watson statistic is used to detect the significant autocorrelation among the independent variables. Autocorrelation does not appear to be a problem as shown by the Durbin–Watson statistics with a coefficient of 2.03 (The 5% critical values for Durbin– Watson test when N= 75, K= 5 are d_L= 1.49 and d_u= 1.77. Thus d_u=1.77< d=2.03 <4-d_L=2.51). The null hypothesis stated that no autocorrelation was accepted, therefore, the autocorrelation does not appear in this estimation.

It can be concluded that Cd accumulation in vegetables is strongly and positively influenced by four variables in component No.3 which includes average amount irrigation water (X9), time of watering (X10), Cd in cultivated soil (X13), Cd in irrigation water (X14). They are variables related to practices in farming using irrigation water and characteristic content of Cd in cultivated soil and Cd in irrigation water.

The component No.3 coefficient was statistically significant at the 1% level and had positive signs, which indicates that Cd accumulation in vegetables increase with the increase of four variables in component No.3. Results indicate a strong relationship between Cd accumulation in vegetables and four variables in component No.3 with coefficient 0.47 (Table 5.14). This indicates that four independent variables of component No.3 will increase Cd accumulation in vegetables. Specifically, if four independent variables of component No.3 are increased by one unit, the Cd accumulation in vegetables will increase by 0.03 units. In other word, this result showed that an increase of average amount irrigation water (X9), time of watering (X10), Cd in cultivated soil (X13), Cd in irrigation water (X14) (component No.3) would lead to increase the Cd accumulation in vegetables.

Besides, Cd accumulation in vegetables is weak negative influenced by component No.1 and component No.4 with coefficient -0.39, -0.23 respectively. Specifically, component No.1 consists of total areas (X4), field area (X5). They are variables relate to size of vegetables field, size of total farming area owned by farmers and size of farmer's family. And component No.4 consists of number source water (X7), following safe vegetable production process (X11). They are variables relate to number irrigation water sources using by farmers and decision of them following safe vegetable production process.

Results from the laboratory analysis of Cd in vegetables showed that the vegetables samples in the study area were contaminated Cd with high level of pollution (average 4.37 times more than the permitted standard). The main cause which leading to the Cd pollution in vegetables in the study area due to the selected vegetables absorbed Cd from farming environment, especially from soil environment. Specifically, cultivated soil environment in the study area affected by the supplement of substances containing Cd into cultivated soil environment, including irrigation water contaminated by wastewater sources from sewage sludge, residential waste water, industrial waste water, etc with frequency and large amounts of irrigation water (Survey result, 2010). In addition, the mobility of heavy metals (especially Cd) in the acidic cultivated soil environment may enhanced uptake into the vegetables. This is also consistent with other studies (Devkota *et al*., 2000; Frost *et al*., 2000; Mangwayana, 1995) which suggested that vegetables had heavy metal contamination from such factors such as using irrigation water contaminated by sewage sludge, concentrations of heavy metals in soils, and in nature of the soil on which vegetables are grown etc ((Lake *et al*., 1984; Scott *et al*., 1996; Voutsa *et al*., 1996). In addition, absorption of Cd is about 10-fold more rapid than that of Pb (Curtis *et al*., 2002; Fritioff *et al*., 2007).

5.4 Factors affecting cadmium (Cd) accumulation in cultivated soil

Factor analysis methods range from confirmatory techniques to pure exploratory procedures (Tucker and MacCallum, 1997). The exploratory factor analysis (EFA) seeks to uncover the underlying structure of a relatively large set of variables. There are no prior theory-driven hypotheses to confirm through the empirical data, and factor loadings are used to merely intuit the factor structure of the data. This is the most common form of factor analysis, where the researcher's as priori assumption is that any indicator may be associated with any factor. To conform with the notion by Costello and Osborne (2005), which concerns the most preferred type of EFA, the principal components analysis (PCA) with orthogonal (varimax) rotation was used as the method for the data analysis. Furthermore, the number of factors was decided to be retained for rotation on the basis of the Kaiser criterion, according to which all factors with Eigenvalues greater than one are feasible.

There are several views on the minimum number of cases required for the factor analysis. Hair *et al*. (2006) suggested that the minimum absolute sample size should be 50 observations and preferably the sample size should be 100 or larger. Generally, an adequate number of cases is suggested to vary between 100 and 300 (Gorsuch, 1983; Hatcher, 1994; Hutcheson and Sofroniou, 1999; Norušis, 2005). Also, as a general rule, the minimum is to have at least five times as many observations as the number of variables to be analyzed (Bryant and Yarnold, 1995; Hair *et al*., 2006). In this data consisting of 75 cases and 13 variables this subjects-to-variables ratio equals to 5.77. Thus, although the number of the cases is low compared to what is suggested by some authors, it seems that the results from this study based on PCA – in order to identify factors affecting cadmium (Cd) accumulation in cultivated soil – have sufficient explanatory power.

In this study, some important factors were taken into account as the independent variables to determine factors affecting Cd accumulation in cultivated soil. Table 5.16 showed the list of independent variables and dependent variable that used in the model of Cd accumulation in cultivated soil.

Table 5.16 Variables and measurements in model of Cd accumulation in cultivated soil

Independent variables	Measurement
X1= Age of the vegetables grower	years
X2= Schooling years	years
X3= Family size	number of member
X4= Total area of cultivated	m^2
X5= Area of vegetable field	m^2
X6= Vegetable growing experiences	number of years
X7= Number sources of irrigation water	number
X8= Farmer using irrigation water from public water pond	Yes= 1, No= 0
X9= Average amount of irrigation water/ m^2/day	liter/ m^2
X10= Times of watering	times/ day
X11= Farmer plant and produce vegetables following safe vegetable production process	Yes= 1, No= 0
X12= Farmer's understanding about heavy metal accumulation in vegetables	Yes= 1, No= 0
X13= Accumulation of cadmium (Cd) in irrigation water	mg/liter
Dependent variables	
Y4= Accumulation of cadmium (Cd) in cultivated soil	mg/kg

A factor analysis for the data (PCA with varimax rotation; values suppressed at the minimum loading 0.40) was conducted. The Kaiser-Meyer-Olkin (KMO) measure of sampling adequacy tests whether the partial correlations among variables are small. Described in other words, it is an index for comparing the magnitudes of the observed correlation coefficients to the magnitudes of the partial correlation coefficients. Large values for the KMO measure indicate that using a factor analysis for the variables is an adequate method, and a value greater than 0.50 is considered desirable (Malhotra and Birks, 2007). The Kaiser-Meyer-Olkin (KMO) measure of sampling adequacy is 0.537, and thus is greater than the recommended level. In addition, the Bartlett's test of sphericity is significant ($p<0.01$), indicating that the strength of the relationship among variables is strong. The Bartlett test of sphericity rejected the null hypothesis that the correlation matrix was an identity matrix. To conclude, Kaiser-Meyer-Olkin (KMO) measures and Bartlett's sphericity test showed that the sample met the criteria for factor analysis.

To be tested the multi-correlations among independent variables (from X1 to X13) with each other and between dependent variable (accumulation of Cd in cultivated soil) and independent variables, factor analysis was applied to group independent variables based on the phenomenon of multicollinearity. Factor analysis separated factor components which have higher value of variance, specifically rotation sums of squared loadings values greater than one (Table 5.17).

In order to obtain a clearer picture of factors affecting cadmium (Cd) accumulation in cultivated soil, a factor analysis of the correlation matrix was conducted. The resulting Eigenvalues for only the first five common factors were greater than unity (0.40). The same conclusion was reached by using the Scree plot (presented in the appendix) as a criterion to determine the number of group components. The proportion (percent) of variance explained by each component is also shown in Table 5.17, indicating that these five group components altogether account for about 62.57 percent of the total variance using the Principal Component method (first factor 15.74 percent, second factor 13.74 percent, third factor 12.41

percent, four factor 10.64 percent and fifth factor 10.04 percent). Specifically, component No. 1 has the highest variance which is about 15.74 percent of variance.

Results of the factor analysis after the orthogonal rotation revealed five underlying patterns that are identified as different group components of factors affecting cadmium (Cd) accumulation in cultivated soil. According to Costello and Osborne (2005) a factor with fewer than three items is generally considered weak and unstable, and five or more strongly loading items are desirable and indicate a solid factor. The factors from this empirical analysis seem to conform with this notion relatively well.

Table 5.17 Group of independent variables extracted by factor analysis

Components	Initial Eigenvalues		
	Total	% of Variance	Cumulative %
1	2.05	15.74	15.74
2	1.79	13.74	29.48
3	1.61	12.41	41.89
4	1.38	10.64	52.53
5	1.31	10.04	62.57

Result from the extraction method of principle component analysis (PCA) and rotation method of Varimax with Kaiser normalization to facilitate the interpretation of the results presented in Table 5.18.

Table 5.18 Rotated factors loading of independent variables

Variables	Rotated factor loadings				
	Component No.1	Component No.2	Component No.3	Component No.4	Component No.5
Total areas (X4)	0.88	-	-	-	-
Field area (X5)	0.79	-	-	-	-
Family size (X3)	0.52	-	-	-	-
Cd in irrigation water (X13)	-0.44	-	-	-	-
Age (X1)	-	0.83	-	-	-
Experiences (X6)	-	0.77	-	-	-
Schooling years (X2)	-	-0.62	-	-	-
Average amount irrigation water (X9)	-	-	0.74	-	-
Time of watering (X10)		-	0.71	-	-
Number source water (X7)	-	-	-	0.72	-
Following safe vegetable production process (X11)	-	-	-	0.66	-
Use public pond (X8)	-	-	-	-	-0.67
Farmer know heavy metal accumulation in vegetables X12)	-	-	-	-	0.67

Note. Extraction Method: Principal Component Analysis.

Rotation Method: Varimax with Kaiser Normalization.

The first factor has relatively large weight for three variables in component No.1 which includes total areas (X4), field area (X5), family size (X3), Cd in irrigation water (X13) which has the coefficients of 0.88, 0.79, 0.52, and -0.44 respectively (Table 5.18). They are variables relate to size of vegetables field, size of total farming area owned by farmers and size of farmer's family and Cd content in irrigation water . Component No.2 consisted of age (X1), experiences (X6), schooling years (X2) which has the coefficients of 0.83, 0.77, and -0.62 respectively. Component No.3 consisted of average amount irrigation water (X9), time of watering (X10) which has the coefficients of 0.74, and 0.71 respectively. Component No.4 consisted of number source water (X7), following safe vegetable production process (X11) which has the coefficient of 0.72, and 0.66 respectively. Component No.5 consisted of use public pond (X8), farmer heard heavy metal accumulation in vegetables (X12) which has the coefficients of -0.67, and 0.67 respectively.

To determine the effect of the group components on Cd accumulation in cultivated soil, correlation matrix analysis was firstly used and multiple regression analysis was applied as the second step of analysis.

Five components and dependent variable (accumulation of Cd in cultivated soil (Y4)) was used to analyze by multiple regression analysis. Result from the analysis showed that Cd in cultivated soil (dependent variable) has a weak positive correlations with independent variables which were in component No.2, component No.3 and component No.5 with the coefficients of 0.02, 0.23, and 0.16 respectively (Table 5.19). And component No.1 and component No.4 has a weak negative correlation with independent variable with the coefficients of -0.29, and -0.14 respectively. So, from correlation matrix can concluded that there is no component which has strong correlation with the dependent variable which is Cd accumulation in cultivated soil.

Table 5.19 Correlation matrix of each component

Variables	Cd in cultivated soil	Component No.1	Component No.2	Component No.3	Component No.4	Component No.5
Cd in cultivated soil	1.00					
Component No.1	-0.29	1.00				
Component No.2	0.02	0.00	1.00			
Component No.3	0.23	0.00	0.00	1.00		
Component No.4	-0.14	0.00	0.00	0.00	1.00	
Component No.5	0.16	0.00	0.00	0.00	0.00	1.00

The relationship between accumulation of Cd in cultivated soil (Y4) (dependent variable) and all components (independent variable) showed in Table 5.20.

Table 5.20 Summary of the regression results of Cd accumulation model in cultivated soil

Variables	Three kind vegetables (n=75)			
	Coefficients	Standard Error	t-radio	Pro
Constant	1.09	0.08	14.44 **	0.00
Component No.1	-0.14	0.06	-2.19 *	0.03
Component No.2	0.02	0.06	0.28 ns	0.78
Component No.3	0.11	0.06	1.79 ns	0.07
Component No.4	-0.09	0.06	-1.47 ns	0.14
Component No.5	0.11	0.06	1.74 ns	0.08

R^2 = 0.18 (18%) Adjusted R^2 = 0.12 (12%)

F value : 2.99 ** Durbin-Watson 2.14

Note: *, ** indicate the level of significance at 5% and 1% respectively.

Result from Table 5.20 showed that the constant were significant at 1 percent of level of significance, component No.1 were significant at 5 percent of level of significance, while the other components were not statistical significant.

The coefficient of determination (R^2) statistic, adjusted for the number of parameters in the equation and the number of data observations. Analysis of multiple regression indicated that the coefficient of determination (R^2) value 0.12 can be regarded as a good fit in view of the cross-sectional data, indicating that about 12% of the total variation of Cd accumulation in cultivated soil explained by the independent variables and the remaining 88% may be due to error and other factors omitted in the model such as fertilizer factors. Therefore, the results of Cd accumulation in cultivated can be explained by five components about 12%. The model has very low significance, therefore the significance of the model could not reflect almost factors that affecting to Cd accumulation in cultivated soil.

The F value and probability statistics test the overall significance of the regression model. According to the F test, the independent variables in the estimated regression equation were highly significant with the dependent variable (Cd accumulation in cultivated soil) at 1% probability level.

Durbin-Watson statistic is used to detect the significant autocorrelation among the independent variables. Autocorrelation does not appear to be a problem as shown by the Durbin–Watson statistics with a coefficient of 2.14 (The 5% critical values for Durbin–Watson test when N= 75, K= 5 are d_L= 1.49 and d_u= 1.77. Thus d_u=1.77< d=2.14 <4-d_L=2.51). The null hypothesis stated that no autocorrelation was accepted, therefore, the autocorrelation does not appear in this estimation.

It can be concluded that Cd accumulation in cultivated soil is very weakly and negatively influenced by four variables in component No.1 (with coefficient -0.29) which includes total areas (X4), field area (X5), family size (X3). They are variables relate to size of vegetables field, size of total farming area owned by farmers and size of farmer's family.

Conclusion, model of factors affecting cadmium (Cd) accumulation in cultivated soil is non statistical significant and could not reflect the factors that have influenced the accumulation of Cd in cultivated soil.