IOP Conf. Series: Earth and Environmental Science **169** (2018) 012106 doi:10.1088/1755-1315/169/1/012106

# Combination of spatial logistic regression and geographical information systems in modelling wetland changes in Setiu basin, Terengganu

**Lindah Roziani Jamru[1] and Zullyadini A Rahaman[2]**

[1]Geosciences & Digital Earth Centre (INSTeG), Research Institute of Sustainable Environment Universiti Teknologi Malaysia, 81310 UTM Johor Bahru, Malaysia
[2]Faculty of Human Sciences, Sultan Idris Education University, 35900 Tanjong Malim, Perak

lroziani2@live.utm.my

**Abstract.** Wetlands played an important role in protecting our natural ecosystem. However, recently there were some issues has been highlighted that affected the natural function of wetlands led to the decline of the biodiversity of flora and fauna. Therefore, this study was carried out to examine the changes occurred in the wetland area of the Setiu watershed, Terengganu and to identify the factor that influence and cause the wetland changes. Integration between Geographic Information System (GIS) and spatial logistic regression has been used to identify the spatial and non-spatial factors that contribute to the wetland changes and to evaluate the significance of the respective factors. Factors that have been used in this study are mean annual rainfall, distance from slope, soil suitability, distance from roads, distance from streams, population density and distance from agricultural areas. Systeme Pour l'Observation de la Terre (SPOT) 5 images have been used to extract data of the wetland changes from 1990 to 2000 for the development of the model, while data from 2000 to 2010 was used to evaluate the accuracy of the model developed. The result shows that the independent variable that significantly influences the wetland changes is the distance from agricultural areas, while the independent variable that has less significant influence is population density. These generated maps show the spatial extent of the Setiu wetlands was change 35% in 1990- 2000 and increased 19% to 54% in 2000 – 2010. The results show that the wetlands area in Setiu district is decreased and should get attention from local authorities for conservation action. Overall, this study has proven that the integration of GIS and statistical analysis become a great tool that has ability to model the wetland change in relation to time and space.

## 1. Introduction

Setiu wetlands are a natural resource that acts as a natural filter of pollutants, becomes flood control and habitat for the flora and fauna. It is vital for the fishing sector and other related economic activities that contribute benefit to the local communities as well as the state of Terengganu. The state government has introduced the High Impact Project that known as the sharp aquaculture to support the economy of Setiu. This project will contribute a lot of benefit to the local community. However, to execute the project, it will sacrifice the natural wetland area that converted to the other land use, such as pond aquaculture.

Recently, rapid changes in economic and development in the past century cause various environmental stresses to the wetlands area. Anthropogenic factors such as residential, industrial and agriculture contributed to this problem. Some cases in developing country, they were completely eliminated the wetland area due to consider it as an unproductive areas and major sources of malaria

[1]. According to [2] the total peat land in peninsular Malaysia in 1981 is approximately 0.67 million hectares. But, it is approximately 0.37 million hectares left in 1991 [3]. However, there still a lack of comparative data of the mapping wetlands change and the impact of the disturbance and development to the ecosystem. Thus, it is necessary to identify the imbalance on the biodiversity of wetland areas in order to ensure the sustainable management [4]. Hence, this study is necessary to monitor the wetlands change and identified the main factors that contribute to the changes.

In this study, logistic regression was selected to develop the wetland change model. Previous research has used this statistical method for wetlands studies such as to specify the hazard of habitat conversion [5], to recognize areas with a high threat of anthropogenic wildfires [6] to identify land use changes [7], to indicate variants in plant inhabitants [8] and modeling of wetland changes [9]. This method is suitable to evaluate the pattern of land use change and will predict the land development in the future. This pattern resulted from the historical previous land use information and the causative factor is highly valued [10]. This statistical approach could identify the most influence variables and provided a degree of confidence, according to the contribution of factor [11]. Thus, this method has been chosen due to the dependent variable has been used is nominal data and an independent variable as range data. Apart from that, logistic regression has the ability to identify the most dominant independent variables that influence the wetland changes. This could help the local district to identify the critical area that has been converted and for a better management of the wetland area.

A complete database containing information of the wetland area could help the local district to identify the critical area that has been converted and required to monitor the status of the wetland in relation to human activities and conservation programs. GIS software capable provided a database and analyzes the influence of the driving factors to the wetland changes. The aim of this study is to examine the changes occurred in the wetland area of the Setiu watershed, Terengganu and to identify the factor that influence and cause the wetland changes. The independent variables that have been used as input model are population density, mean annual rainfall, distance from the slope, suitability of soil, distance from road, distance from the river and distance from the agricultural area.

## 2. Material and Methods

### *2.1. Study Area*

Setiu is located at the east of peninsular Malaysia in the state of Terengganu. Setiu wetland have been identified as a " Kawasan Sensitif Alam Sekitar" that start from Kampung Penarik located about 300 meters from the coast. The Setiu river systems with a diverse array of marine and terrestrial habitats support various wetland types, natural and man-made (Figure 1). The natural wetland consists of lagoon, freshwater swamps, peat land and mangrove, while the man-made wetland consists of paddy field, dam, pool, lake and aquaculture. The Setiu wetland area is approximately 880 hectares. It provides benefit to locals and become habitat for flora and fauna as well. The wetland has distinctive natural attractions that consists various ecosystems such as beaches, sea, mud flux, lagoon, estuary, river, island and mangrove forest. In general, Setiu wetland can be divided into two distinctive seasons, which are dry season and monsoon (wet season). The monsoon that carries more rain into the area start from November and continues until February. During this period, Setiu recorded a mean annual rainfall of 757.01 ± 422.39 mm. The average daily temperature was low during the monsoon period, ranging between $22^0$C - $19^0$C, compared to $23^0$C - $32^0$C during normal period (dry season).

### *2.2. Data Acquisition*

The source of data to extract the dependent and independent variables are listed in Table 1. Topographic maps series L7030 has been used to extract the land use data for 1990 with sheet number 4165, 4166, 4167 and 4266. Meanwhile, the land use of 2000 and 2010 have been extracted and updated based on SPOT 5 image.

**Table 1**. Data types and sources

| Data type | Scale | Sources |
|---|---|---|
| Land use map | 1: 10 000 | Department of Survey and Mapping Malaysia (1990) |
| SPOT 5 image | 30 m x 30 m | Malaysia Remote Sensing Agency (2009) |
| Rainfall | 1: 10 000 | Department of irrigation & drainage (1990, 2000, 2010) |
| Digital Elevation Model (DEM) | 1: 10 000 | Department of Survey and Mapping Malaysia (1990) |
| Soil type | 1: 10 000 | Malaysia Remote Sensing Agency (1990) |
| Road | 1: 10 000 | Rancangan Tempatan Daerah (1990) |
| District | 1: 10 000 | Malaysia Remote Sensing Agency (1990) |
| River | 1: 10 000 | Malaysia Remote Sensing Agency (1990) |
| Population density | 1: 10 000 | Population census of Malaysia (1990, 2000, 2010) |

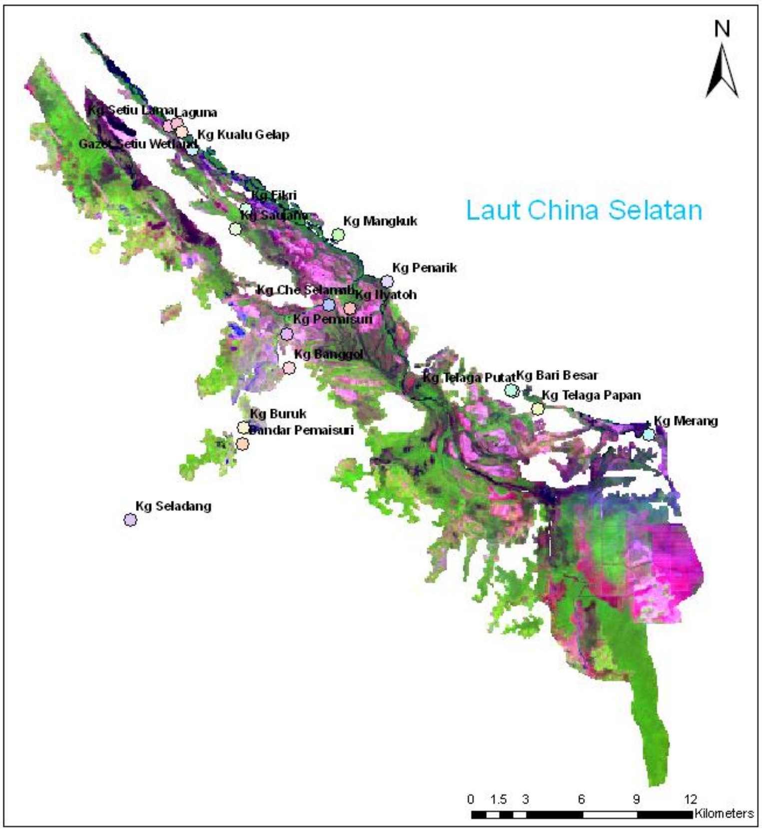



**Figure 1**. Setiu basin

*2.2.1 Remote Sensing Data*

SPOT 5 has a second instrument known as SPOT vegetation that very useful for observant and evaluate the change of land over large areas and mapping the vegetation at different level of scales [12]. Previous work by [13] and [14] used SPOT to map and analyze wetland and found that object based method can clearly identify land cover types. In this study, SPOT 5 satellite imagery was acquired on 13 March 2009 provide by Agency of Remote Sensing Malaysia (ARSM). The scene number 271339 with 10 m resolution has been used in this study to extract the wetland features. The images were processed using ERDAS IMAGINE 9.1 software and were registered to Rectified Skew Orthomorphic (RSO) projection. A total of nine ground control points is selected with the maximum SMS error of rectification was 15 m. There were seven classes of wetland types in the study area; namely aquaculture, *Melalueca* swamp, freshwater swamps, peat swamps, mangrove, paddy and grassland. Maximum Likelihood classification was performed, since this classification was widely used to classify wetlands [11] [12] [15]. Figure 2 shows the wetland information as obtained from

Maximum Likelihood analysis.

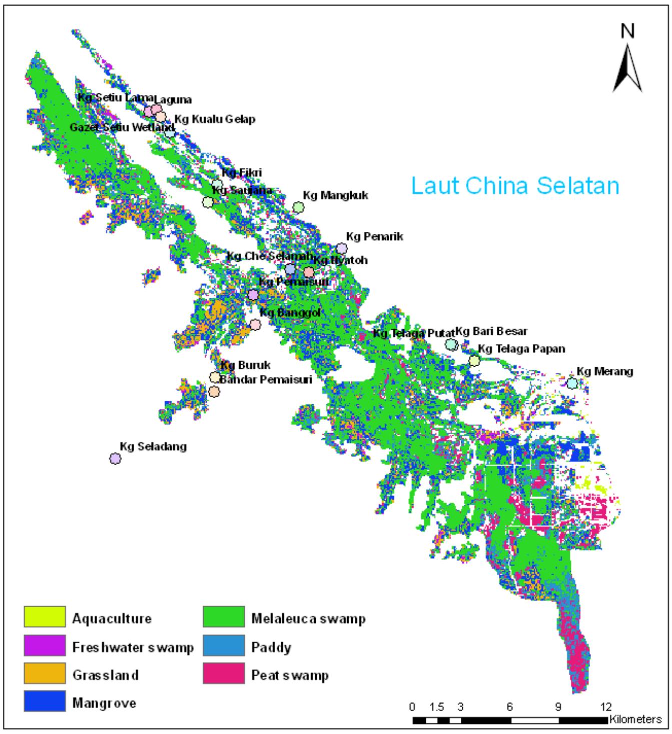



**Figure 2**. Wetland information obtained from Maximum Likelihood classification based on 2006 satellite imagery

*2.3. Logistic regression analysis*

In this study, logistic regression was used to estimate the probability of the occurrence of wetland changes in the study area based on the relationship with independent variables that contribute to the changes. The equation (1) used in this study was adapted from [16] as follows:

$$p = (E(Y) = \exp(\beta o+\beta 1X+\varepsilon) / (1 + \exp(\beta o+\beta 1X+\varepsilon))\ p \quad (1)$$

where,
P = Probability occurrence of wetland change
E (Y) = Expected values of binary dependent variables Y (wetland changes)
Z = Logistic regression function for independent variable
$a_0$ = Constant
a1..an = Coefficient
x1..xn = Independent variables (metric & non-metric)
e = Stochastic disturbance

Logistic regression was used to estimate the probability based on the hypothesis null between presence as 1 or absence as a 0. The dependent variables are a binary variable. This method has been used to predict the wetland area base on the influence of independent variables towards the dependent variables. The stepwise analysis method has been applied as the best elimination criteria. All the variables in the model will be tested and the variable not significant will be eliminated.

The significant value for the independent variable is $\alpha = 0.1$ and below $\alpha = 0.1$ considered not significant. The significant variables included in the macro model developed using IDRISI software. An input data like raster and vector will be linked using this model to produce an output. The developed model then will be evaluated to access the level of accuracy. The accuracy assessment is important to ensure that the predicted wetland changes are similar to the actual wetland changes.

*2.4 Dependent variable*

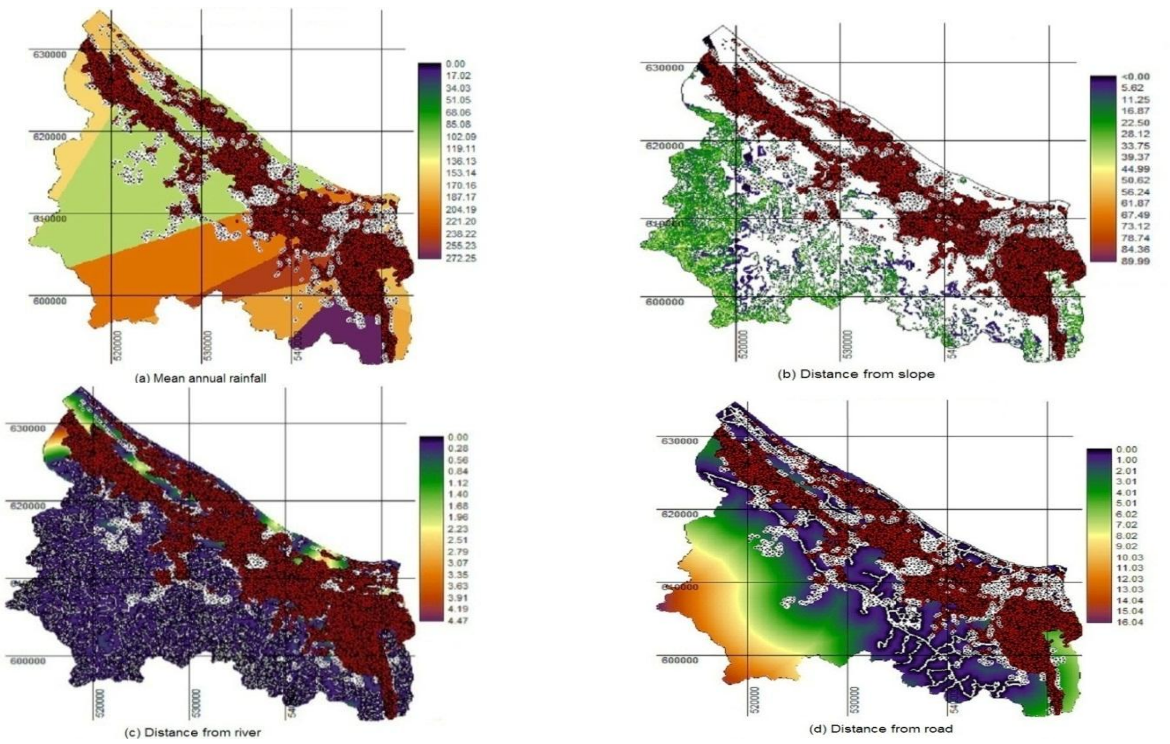
The wetland changes detection was processed using ArcGIS 9.3. Raster calculator analysis tools used to calculate the wetland area that has been changed. The wetland changes occurred between the years 1990 to 2000 has been used as a reference to the model development and changes between the years 2000 to 2010 has been used to validate the accuracy of the model. The dependent variable classified to 0 and 1 that represent no occurrence wetland change and wetland change respectively. The results were calculated to analyze the percentage and total area of the wetland has been changed.

*2.5 Independent variables*

The purpose of this model is to evaluate the contribution spatial and non-spatial factors towards wetland changes. There are seven factors were selected in this model. The seven independent variables are classified to geographical variables and characteristics. Four independent variables, namely DEM, soil suitability, rivers and rainfall are classified as characteristic and the remaining three, namely the distance from road, distance from agricultural areas and population density are classified as geographical parameters (Table 2 and Figure 3). The soil suitability factor is classified based on the "Key to Soil Identification by the Department of Agriculture". All independent variables were processed in ArcGIS 9.3 using spatial analysis, such as classification, interpolation, surface analysis and distance analysis.

**Table 2**. The list of dependent and independent variables

| Variables | Content | Type of variables |
|---|---|---|
| **Dependent** | | |
| Y | 0 – non wetland, 1- wetland | Dicotomos |
| **Independent** | | |
| $X_1$ | Mean annual rainfall | constant |
| $X_2$ | Distance from slope | constant |
| $X_3$ | Suitability of soil | design |
| $X_4$ | Distance from road | constant |
| $X_5$ | Distance from stream | constant |
| $X_6$ | Density of resident | constant |
| $X_7$ | Distance from agriculture area | constant |



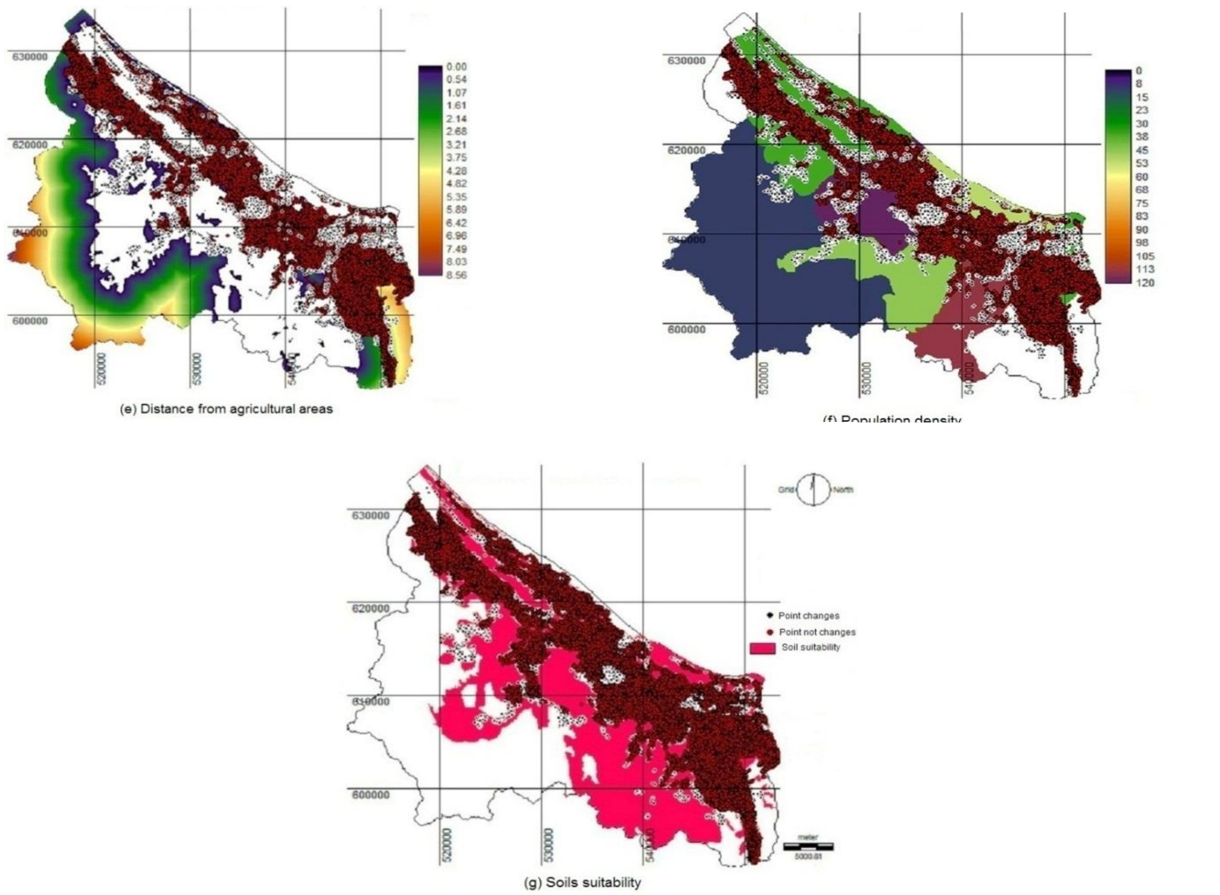



**Figure 3**. The spatial map of independent variables.

*2.6 Wetland change modeling*

Logistic regression analysis was carried out using SPSS Version 16 to identify the influence of driving factor towards wetland changes. The dependent and independent variables were converted into polygons and overlaid in the data base. Then, it was exported as an X and Y variables into SPSS for statistical analysis. The independent variables that are significant ($\alpha$ = 0.1) will be combined to produce wetland change model. Next, the significant independent variables exported into the models and run using IDRISI Klimanjaro software via LOGISTICREG. The output is a map showing the wetland changes. The logistic regression equation (2) obtained is as follows;

$$WC = -0.094 \times DR - 0.372 \times DS + 0.078 \times DEM - 0.429 \times SS - 1.023 \times DAA \quad (2)$$

where;

WC = Wetland changes
DR = Distance from road
DS = Distance from stream
DEM = Digital Elevation Model
DAA = Distance from agriculture area

**3. Result and Discussions**

The wetland change model that has been developed in terms of mathematical equations and translated into spatial layer maps. It shows the probability of wetland changes in the study area. The variables that have significant influence to the wetland changes were identified thus enabling to predict the probability of wetland changes in the future. The results of logistic regression analysis show that the distance from agricultural areas have a strong effect on the probability of wetland changes with a coefficient value of -1.130 (Table 3). Meanwhile, population density has the less significant among the independent variables. Based on the statistical analysis, six out of seven independent variables are significant. Mean annual rainfall has been eliminated because of low coefficient value. Based on the trend of land use changes in the study area showed that the agricultural areas increased from 296 $km^2$ to 383 $km^2$ in the year 2010. The wetland areas have declined from 321.54 $km^2$ in the year 1990 to 286.23 $km^2$ (2010). Setiu district has been included in the East Coast Economic Region (ECER) and this area has been identified to generate economic sources for agricultural and fisheries sectors. The agricultural area expanded progressively and threatening the wetland areas.

**Table 3**. Spatial logistic regression of wetland changes models of the year 1990-2000 and 2000-2010

| Variables | Development Models (1990-2000) | Validation Models (2000-2009) |
|---|---|---|
| | Coefficient | Coefficient |
| Mean annual rainfall | 0.00005764 | -0.00288550 |
| Distance from slope | 0.00001425 | -0.00033995 |
| Distance from river | -0.37981185 | 0.45623770 |
| Distance from road | -0.06334020 | -1.28436628 |
| Distance from agricultural areas | -1.13069661 | -0.01051496 |
| Population density | -0.00753848 | 0.01487736 |
| Soil suitability | 0.33905100 | 2.32249147 |
| Constant | -0.34065008 | -4.56002514 |

Figure 4 shows the probability map of wetland changes in the year 1990 to 2000. The values of wetland changes classified from 0.00 to 0.49 as a no occurrence of wetland changes, while 0.5 to 0.1 classified as a wetland change. According to the wetland pixel area, 7106 sample points have been used in this model. The result shows that 64.60 percent or 4591 points predicted to have no occurrence of wetland change and 35.35 percent or 2515 of sample points has predicted changes of wetland areas for the year 2000 to 2010. For the accuracy model (Figure 5), 3250 sample points or 49.53 percent recorded no wetland changes, while 3586 or 54.46 percent has predicted changes of wetland areas.

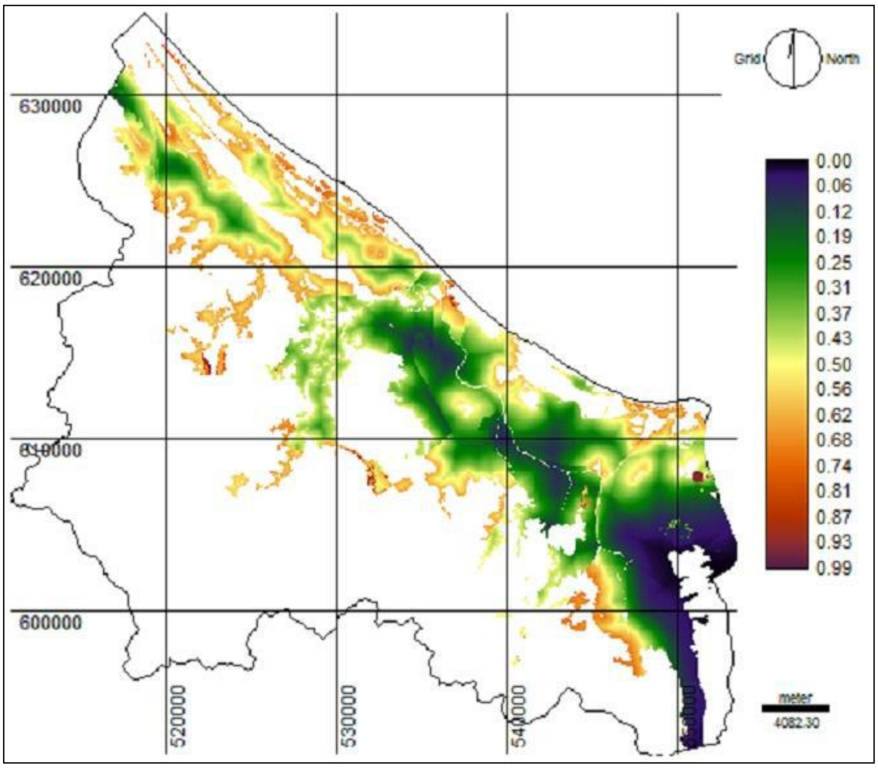



**Figure 4.** Map of the probability wetland changes 1990 - 2000

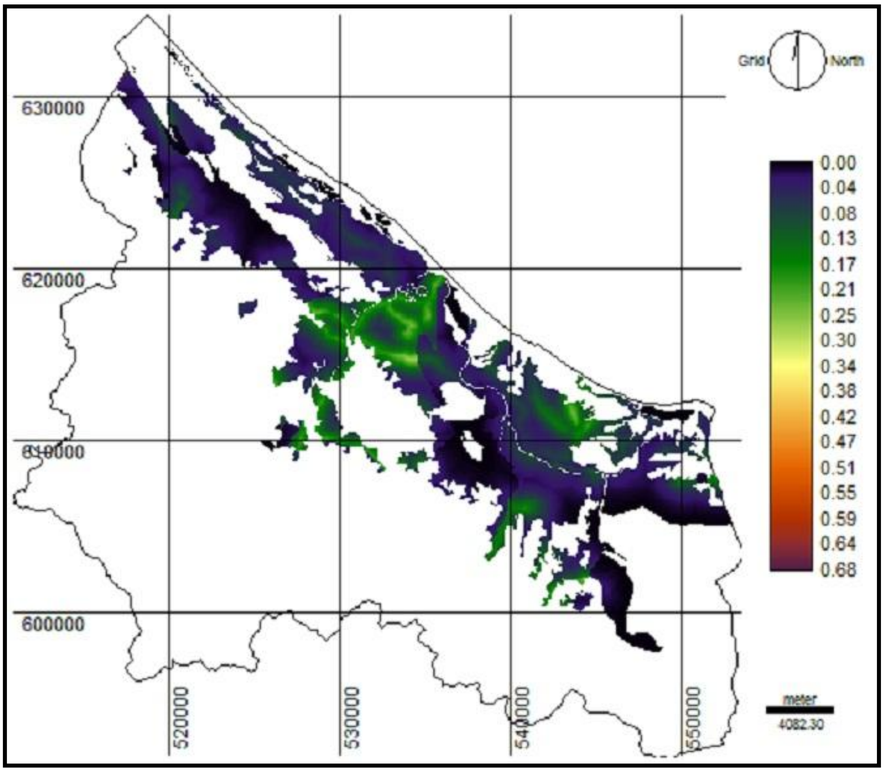



**Figure 5**. Map of the probability wetland changes 2000 – 2010

## 4. Conclusions

The spatial logistic regression and GIS in this study is a great tool to predict the wetland changes in the future. This model was able to identify the location of areas that have a high probability of changes in the future. Also, the influence of the independent variables to the wetland changes has been evaluated. From this model, the highest impact of independent variables to the wetland changes is distance from agricultural areas. Wetland areas assumes as an unproductive areas to the developer and local people. Most of the wetland areas were converted to agricultural area and more profitable. This achieved the aim of the ECER to boost the economic sources through agricultural and fisheries sector. However, the relevant authority needs to monitor and manage the wetland area from extinction. The relationship between wetland changes and the influencing factors that analyzed using an empirical technique is a complex process. However, this complexity can be simplified using the logistic regression model and repeatable.

## References

[1] Sa´nchez-Gonza´lez S. M., Martı´nez-Alegrı´a R., Taboada J 2015 Modeling wetland change in Spain's Tierra de Campos district .*Wetlands Ecol Manage* DOI 10.1007/s11273-015-9463-x.
[2] Shamsudin, I. & Ismail. H 1991 The impacts of the present landuse on peat swamp forests in Peninsular Malaysia. *Malaysia Forester* **54** (1): 15 – 23.
[3] Chin T.Y., Nor Akhiruddin M., Samsuanuar N., Yong T.K., Hasnuddin M.A & Mohd. Nashir S.I 1997 Report on Third National Forest Inventory Peninsular Malaysia. Forestry Deparment Headquarters. Kuala Lumpur. 211pp.
[4] Tamblyn, A., Turner, C. and Raines, P 2006 Malaysia tropical forest conservation project report of the Setiu Wetlands phase. Department of Wildlife and National Parks, Malaysia (PERHILITAN) and Coral Cay, Conservation. p. 1 – 86.
[5] Stephens SE, Walker JA, Blunck DR, Jayaraman A, Naugle DE, Ringelman JK, Smith AJ 2008 Predicting risk of habitat conservation in native temperate grasslands. *Conserv Biol*

**22**:1320–1330.

[6] Vilar L, Martı´n MP, Vega J 2008 Empleo de te´cnicas de regresio´n logı´stica para la obtencio´n de modelos de riesgo humano de incendio forestal a escala regional. Bol. Asoc. *Geo´grafos Esp* **47**:5–29.

[7] Pineda NB, Bosque J, Go´mez M, Plata W 2009 Ana´lisis de cambio del uso del suelo en el Estado De Me´xico mediante sistemas de informacio´n geogra´fica y te´cnicas de regresion multivariantes. Una aproximacio´n a los procesos de deforestacio´n. Investigaciones Geogra´ficas. Bol. Instituto Geografı´a, *UNAM* **69**:33–52.

[8] Brus DJ, Hengeveld GM, Walvoort DJ, Goedhart PW, Heidema AH, Nabuurs GJ, Gunia K 2012 Statistical mapping of tree species over Europe. *Eur J Forest Res* **131**:145–157.

[9] Verburg, P.H., de Koning, G.H.J., Kok, K., Veldkamp, A., & Bouma, J 1999 A spatial explicit allocation procedure for modeling the pattern of land use based upon actual Land use. *Ecol. Model*. **116**, 45- 61.

[10] Irwin, E. G., & Geoghegan, J 2001 Theory, data, methods: developing spatially explicit economic models of land use change. *Agriculture, Ecosystems and Environment*, **85**, 7 – 23.

[11] Lee T. M. and Yeh H. C 2009 Applying remote sensing techniques to monitor shifting wetland vegetation: A case study of Danshui River estuary mangrove communities, Taiwan. *Ecological Engineering*, **35**: 487-496.

[12] Yichuan X, Zongyao S, Mei Y 2008 Remote sensing imagery in vegetation mapping: a review. *Journal of Plant Ecology*, **1**:9 – 23.

[13] Hubert –Moy L, Michel K., Corpetti T, Clement B 2006 Object oriented mapping and analysis of wetlands using spot 5 data; Proceedings of the IEEE International Conference on Geoscience and Remote Sensing Symposium (IGARSS 2006); pp. 3447–3450.

[14] Hassan N., Hamid J.R.A., Adnan N.A., Jaafar M 2013 Delineation of wetland areas from high resolution WorldView-2 data by object-based method; Proceedings of the IOP Conference Series Earth and Environmental Science; Kuching Sarawak, Malaysia.

[15] Shanmugam P., Ahn Y. H., Shanmugam S 2006 A comparison of the classification of wetland characteristics by linear spectral mixture modeling and traditional hard classifiers on multispectral remotely sensed imagery in southern India. *Ecological modeling*, **194**: 379-394.

[16] Manard, S 2002 Applied logistic regression analysis. Series: Quantitative application in the Social Science. 2nd Ed. London: Sage Publications. Inc.