Chapter 4 The DLS Model and Its Application

The model of Dynamics of Land System (DLS) is a collection of programs that simulates pattern changes in land uses by conducting scenario analysis of the area of land use change. Results reveal the driving mechanisms of land use change and simulate the balance between supply and demand of land at the pixel level (Deng et al., 2008). The model analyzes causes of the dynamics of land use patterns, simulates the process of land use changes, and assists land use planning and land management decisions (Deng et al., 2010a). The DLS model can export a macroscopic pattern changes map of land uses at high spatial and temporal resolution by estimating the effects of driving factors of spatial pattern changes, formulating land use conversion rules and scenarios of land use change and simulating dynamic spatiotemporal processes of land use changes. Driving factors include natural environmental conditions, socioeconomic factors and land use management policies, all of which are closely linked to pattern changes in land uses (Lambin et al., 2001; Lambin et al., 2003; Haberl et al., 2004; Burgi et al., 2004; Aguiar et al., 2007; Turner II et al., 2007; Veldkamp and Verburg, 2004; Verburg, 2006).

4.1 Principles and Function Modules of the DLS Model

The DLS model is theoretically based on restrictions of the distribution of land use types. The model dynamically simulates the macroscopic pattern changes in land uses by classifying the driving factors that influence this pattern (Deng et al., 2008). The simulation spatially allocates the area change in land use and is based on spatial statistics, predicting the probabilities of different land use types and incorporating the probability of distribution of different land use types at the pixel level.

4.1.1 Fundamental Definition

Simulating the macroscopic pattern changes in land uses involves simulating the spatiotemporal processes of changing area and the distributions of regional land use types. This is done by quantitatively measuring flow and

X. Deng, *Modeling the Dynamics and Consequences of Land System Change*

stock in the conversion processes of various land use types (Turner Ⅱ, 1997; Veldkamp and Lambin, 2001; Burgi and Turner, 2002). To realize this goal, it is necessary to first understand the target and features of the simulated dynamics of the land use pattern.

Simulating the macroscopic pattern changes in land uses is targeted at the human-modified land system, which is closely related to land use. The land system is an open, complex system consisting of two subsystems, the geographic environment and human activities, inside which are certain structures and functions (Verburg et al., 2002; Wang et al., 2010). At the core of simulated land use pattern changes are interactions among the natural environment, human society and area of land use change. Accordingly, to produce a macroscopic simulation of a land use pattern, it is necessary to explore new approaches to simulate the dynamics of land system spatial distribution, temporal processes, change in organization, bulk effect and complementary synergies (Veldkamp and Verburg, 2004; Liu et al., 2005).

Changes in the area of regional land use types are closely related to other factors at different scales in the land system. The relationship between them generally includes features such as mechanisms, feedbacks, complexities and systematizations, which are specifically represented as follows.

(i) Natural controlling factors, represented by terrain, climate, soil and vegetation, play a dominant role in changing the regional land use pattern in the long-term and in controlling the direction and degree of change in the regional land use pattern. (ii) Socioeconomic driving factors, including population change, economic development, technical progress and institutional changes, interact with the area of land use change and play a decisive role in the pattern changes in regional land uses in the short-term. (iii) Various nonlinear relationships exist between natural controlling factors and socioeconomic factors, which often conceal the real reasons for the pattern changes in land uses.

Many limitations still exist in current research on simulating the pattern changes in land uses. Systematic analysis and expression of mechanisms of pattern changes in land uses are difficult to conduct (Dai et al., 2005; Pontius et al., 2007). Pattern changes in land uses are closely related to land use decisions, and therefore, simulating pattern changes in land uses needs to comprehensively consider factors such as socioeconomic development, cultural traditions, natural conditions and historic trends in pattern changes in land uses to improve the reliability and accuracy of simulation results.

4.1.2 Features of the DLS Model

Recent research has made progress in the analysis of driving forces behind pattern changes in land uses with economic models and empirical statistical methods (Liu, 2002; Veldkamp and Verburg, 2004; Li et al., 2005; Liu et al., 2005). Shi et al. (2000) analyzed natural and human factors driving land use change in Shenzhen with regression analysis; Chen et al. (2000) built a multiple regression model of land use change using a multi-scale statistical method. Researchers at the Institute of Agricultural Resource and Planning, Chinese Academy of Agricultural Sciences cooperated with scholars at Wageningen University, the Netherlands, to build a model of land use change in China with assistance from a geographic information system (GIS). This was a good attempt at creating a comprehensive evaluation model of land use change (Verburg et al., 2000). Simulations of pattern changes in land uses have focused on regional and microscopic aspects; however, in-depth research has involved the utility of using the models mentioned above with these two aspects, but many limitations still exist.

Conventional models capable of simulating the macroscopic pattern changes in land uses are limited to simulation of only one or several land use types (Ge and Dai, 2005); however, the DLS model differs from these models because it comprehensively simulates the spatiotemporal pattern of all kinds of land use types at the regional scale. It has solved the problem of discriminating between endogenous and exogenous driving factors of land use changes. In addition, the DLS model quantitatively analyzes the effects of different driving factors by building a spatially-explicit statistical model of the distribution of land use types and driving factors at the pixel level, and it sees the pattern changes in land uses as a dynamic spatiotemporal process. Also, different scenarios of changing area of regional land use types are designed in the DLS model based on comprehensive measurements of factors such as features of regional socioeconomic development, cultural traditions, natural conditions and history of land use. Thus, the DLS model has improved the scientific and rational nature of predicted and estimated results.

4.1.3 Framework of the DLS Model

The DLS model fully considers the links among related models of nature, ecology and economy. It also extracts decision-making reference information used in land use planning, environmental planning and management of natural resources by designing different scenarios of changing regional land use area. Users of the DLS model can input nonlinear demand change, different conversion rules and driving factors at different pattern changes in land uses to simulate and analyze the complex changes in regional land use patterns. The DLS model also considers the influence of macroscopic factors such as topography, environment, trade and institutional arrangement and land management policies to more accurately simulate possible scenarios of pattern changes in land uses.

The DLS model presumes that land use pattern change is influenced by both historic pattern changes in land uses and driving factors within the pixel and neighboring pixels. Decisions of land use planners have important influence on the pattern changes in land uses, especially at the regional level (Fig. 4.1).

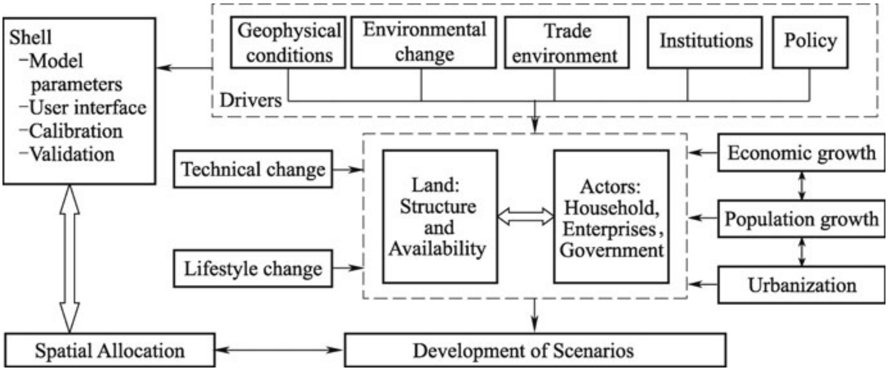


Fig. 4.1 Modeling framework of the DLS model.

In addition, the DLS model considers regional restrictions on the distribution of land use types. For example, the model sets regions where it is impossible for a certain land use type to appear as restricted regions and removes these regions so that they are not input into the model. Moreover, the input parameters and exogenous variables may change with time due to influences from conversion rules of regional land use types and nonlinear demand change. Therefore, it is still necessary to consider uncertainties of simulation results in the DLS model.

4.1.4 Application

The DLS model effectively simulates spatiotemporal pattern changes in land uses. Regarding data integration, the DLS model represents various land use types in a grid format, which spatially expresses the characteristics of the distribution of regional land use types with a high resolution. The DLS model takes information from the basic grid unit as observed data and performs a spatiotemporal simulation of the pattern changes in land uses at the pixel level.

The DLS model fully considers the complexity of the driving mechanisms of pattern changes. It reveals dynamic spatiotemporal rules of land use changes by considering regional restrictions on changing land use area based on a comprehensive analysis of factors that influence land use changes. Researches have indicated that as an auxiliary tool to analyze changes in regional land use area, natural environmental effects of these changes, land use planning and land management decision-making, the DLS model has truly realized the dynamic simulation of pattern changes in land uses with scenarios of changing land use area at the regional scale. It also analyzes mechanisms driving the distribution of land use types at the grid scale (Deng et al., 2008; Deng et al., 2010b).

4.1.5 Function Modules of the DLS Model

The DLS model is based on quantitative analysis of land use pattern changes at the pixel level, interactions among driving factors and spatiotemporal distribution of land use pattern change. It simulates the pattern changes in regional land uses by analyzing the driving forces of land distribution at the grid scale and allocation of changing land use areas. Analyzing both the driving forces behind land distribution and the spatial allocation of land use change is the most important component of the DLS model (Deng et al., 2008).

Mechanism analysis of the DLS model aims to estimate the statistical relationship between the pattern changes in land uses and its driving factors. Theoretically, mechanistic analysis provides a reaction function of each land use type. Corresponding weights are given to all driving factors according to principles that can be assumed to be fixed for a short period, but driving factors change over time. With the reaction function determined, reasons for differences between simulated and observed distribution of land use types can be summarized as follows: values of some driving factors have changed, such as population growth or temperature; competition exists among different land use types; and restrictions occur between local historic conditions and current demand. Driving factors behind land use patterns can be analyzed with the explanatory linear model of land use pattern (ELMLUP) and explanatory nonlinear model of land use pattern (ENMLUP) built at the pixel level. The two models can be used to research restrictions on the distribution of land use types at the pixel level with different backgrounds and goals, and they can be used flexibly to reveal in-depth driving mechanisms of pattern changes in land uses at the pixel level.

4.1.6 Explanatory Linear Model of Land Use Pattern

Linear regression is the model most commonly used in researching the driving mechanisms of land use patterns as it explores driving factors at wide ranges and with high spatial resolution (Verburg et al., 2002; Zhang et al., 2003). The explanatory linear model of land use patterns at the pixel level, or ELMLUP, is introduced in this chapter.

4.1.6.1 Model Hypothesis

The ELMLUP contains a demanding and a distribution module. The target variable of the ELMLUP is the proportion of the area of land use type $k(k = 1, 2, \cdots, M)$ in grid $i(i = 1, 2, \cdots, n)$ at time t abbreviated as Q_i^{kt}. The explanatory variable of the model is a covariant vector of driving factors composed of a series of natural environmental conditions and socioeconomic factors that are tightly related to the pattern changes in land uses (with a significance level of 5%).

$$X_i^t = (x_{i1}^t, x_{i2}^t, \cdots, x_{il}^t, \cdots, x_{iL}^t)^T \tag{4.1}$$

To measure the impact of spatially autocorrelated land use types, several variables, including $\hat{Q}_i^{kt}$ and $\hat{X}_i^t$, are defined in the ELMLUP. Let $\hat{Q}_i^{kt} = \sum_{j \neq i} w_{ij}^k Q_i^{kt}$, where w_{ij} is the spatial weight function of the impact of grid j on grid i. The definition of $\hat{X}_i^t$ is similar to that of $\hat{Q}_i^{kt}$, which is the weighted average of X_i^t. According to the first law of geography, the spatial weight function is usually defined as the reciprocal of the distance between grid j and grid i.

$$W_{ij}^k = \begin{cases} 1/D_{ij} \\ 0 \end{cases} \tag{4.2}$$

where D_{ij} can be the Euclidean distance, the absolute distance, or the Minkowski distance (Tobler, 1970).

4.1.6.2 Model Inference

Spatial autocorrelation

The quantitative relationship between $\hat{Q}_i^{kt}$ and $\hat{X}_i^t$ is developed through the following multiple linear regression model.

$$\hat{Q}_i^{kt} = a_0^k + a^k \hat{X}_i^t \tag{4.3}$$

where $a^k = (a_1^k, a_2^k, \cdots, a_L^k)$ is the coefficient matrix of $\hat{X}_i^t$, and a_0^k is a constant term. Regarding grid i at time t, the result $reg(\hat{Q}_i^{kt})$ estimated by the model is naturally employed to reflect the average proportion of area of land use type k under natural and socioeconomic conditions $\hat{X}_i^t$.

Apparently, $reg(\hat{Q}_i^{kt})$ does not equate with deviation of the actual observed value $real(\hat{Q}_i^{kt})$. If the demanding area of land use type k changes in the demanding module, the relative stability of Land use pattern will be broken. Therefore, we can hypothesize that a certain relationship exists between the land pattern change and the value difference between $reg(\hat{Q}_i^{kt})$ and $real(\hat{Q}_i^{kt})$: when estimated value, $reg(\hat{Q}_i^{kt})$, is smaller than the observed value, $real(\hat{Q}_i^{kt})$, the area proportion of land use type k will increase; when estimated value, $reg(\hat{Q}_i^{kt})$, is larger than the observed value, $real(\hat{Q}_i^{kt})$, the area proportion of land use type k will decrease.

When the demanding module of the ELMLUP requires the area proportion of land use type k to change to $DEMAND^{k(t+1)}$ at time $t + 1$ in grid

i, the area proportion of land use type k will vary.

$$\hat{Q}_i^{k(t+1)} = real(\hat{Q}_i^{kt}) + [reg(\hat{Q}_i^{kt}) - real(\hat{Q}_i^{kt})] \cdot F_i \tag{4.4}$$

where $\hat{Q}_i^{k(t+1)}$ is the area proportion of land use type k in grid i at time $t+1$; F_i is the changing coefficient of the area of the land use type regulated by the general demanding equation:

$$DEMAND^{k(t+1)} = \sum_{i=1}^{n} \{real(\hat{Q}_i^{kt}) + [reg(\hat{Q}_i^{kt}) - real(\hat{Q}_i^{kt})] \cdot F_i\} \tag{4.5}$$

An iteration adjustment is then needed until the proportion of the area of the land use type k increases to $DEMAND^{k(t+1)}$.

Conversely, if the demanding area of land uses type decreases, the trends for the changes in the land uses types will be dissimilar.

ELMLUP

In the same way, the quantitative relationship between $\hat{Q}_i^{kt}$ and $\hat{X}_i^t$ is developed through the following regression equation:

$$Q_i^{kt} = b_0^k + b^k X_i^t \tag{4.6}$$

where $b^k = (b_1^k, b_2^k, \cdots, b_L^k)$ is the regression coefficient matrix of X_i^t, and b_0^k is a constant term. $reg(\hat{Q}_i^{kt})$ is the area proportion of land use type k at time t, estimated by the multiple linear regression model; $real(\hat{Q}_i^{kt})$ is the actual observed value of the area proportion of land use type k at time t. When the demanding module requires the proportion of the area of land use type k to change to $DEMAND^{k(t+1)}$ at time $t+1$ in grid $\underline{i}$, the area proportion of land use type k will vary correspondingly.

$$Q_i^{k(t+1)} = real(Q_i^{kt}) + [reg(Q_i^{kt}) - real(Q_i^{kt})] \cdot F_i' \cdot R_i^{kt} \tag{4.7}$$

where the definition of F' is similar to that of F_i; R_i^{kt} is the influence function that stands for the influence of the spatial autocorrelation factors on grid i that change with $\hat{Q}_i^{kt}$. If the change in area of land use type k in grids near grid i is frequent, there will be correspondingly great changes in Q_i^{kt} and R_i^{kt} due to spatial autocorrelation.

By contrast, if land use type k in grids near grid i are relatively steady, the change in Q_i^{kt} would be correspondingly little. This is the explanatory model for land use patterns in linear form, or the ELMLUP, which considers the effect of spatial autocorrelation.

4.1.6.3 Model Estimation

Many approaches exist to estimate the coefficient of the two multiple linear regression functions in ELMLUP. In this chapter, we introduce one of the most commonly used methods, the least squares method.

The least squares method produces a line that has the minimum sum of the deviations squared (least square error) from a given set of data (Gao et al., 2005). For the two multiple linear regression functions in the ELMLUP, the minimum sums of the deviations squared are respectively defined by the following two equations:

$$\hat{Q}(a_0^k, a^k) = \sum_{i=1}^{n} [real(\hat{Q}_i^{kt}) - reg(\hat{Q}_i^{kt})]^2 = \sum_{i=1}^{n} [real(\hat{Q}_i^{kt}) - (a_0^k + a^k \hat{X}_i^t)]^2 \tag{4.8}$$

$$Q(b_0^k, b^k) = \sum_{i=1}^{n} [real(Q_i^{kt}) - reg(Q_i^{kt})]^2 = \sum_{i=1}^{n} [real(Q_i^{kt}) - (b_0^k + b^k \hat{X}_i^t)]^2 \tag{4.9}$$

4.1.6.4 Model Test

It is still unknown whether linear relationships exist between the change in land use pattern and the natural and socioeconomic factors after the regression coefficient of the multiple linear functions is obtained in the ELMLUP. Therefore, a significance test is needed for the estimated multiple linear regression function. Here, we introduce the approach of using variance analysis to test the significance of the regression function. In this approach, the total variance is decomposed into two parts.

$$\begin{aligned}\sum_{i=1}^{n} (Q_i^{kt} - \bar{Q}^{kt})^2 &= \sum_{i=1}^{n} [Q_i^{kt} - reg(Q_i^{kt})]^2 + \sum_{i=1}^{n} [reg(Q_i^{kt}) - \bar{Q}^{kt}]^2 \\ &= ESS + MSS\end{aligned} \tag{4.10}$$

where $\bar{Q}^{kt}$ is the average of Q_i^{kt}; *ESS* is the sum of variance of regression function; *MSS* is the sum of variances of the errors. Then, a new statistic is defined as follows:

$$F = \frac{MSS/f}{ESS/g} = \frac{MMS}{EMS} \tag{4.11}$$

where statistic F has a distribution of $F(f, g)$, and f and g are the degrees of freedom of the regression function and error, respectively.

By calculating statistic F and the significance probability, we can judge the significance of the multiple regression function. If the value of the significance probability is relatively small, or smaller than the significance level (for instance 0.01), we can conclude that the regression function accurately simulates the relationships between land use patterns and their driving factors.

4.1.7 Explanatory Nonlinear Model of Land Use Pattern

The driving force analysis model for land use patterns in nonlinear form is built based on land use area percentage grid data.

4.1.7.1 Grid Area Percentage Data

Percentage data were first proposed by Ferrers (1866) and is becoming increasingly important in statistical analysis. It is usually expressed as the following vector set:

$$S = \left\{ (s_1,\ s_2,\ \cdots,\ s_m)^T \in R^m \middle| \sum_{i=1}^{m} s_i = 1,\ 0 < s_i < 1 \right\} \tag{4.12}$$

$$s_i = S_i \Big/ \sum_{j=1}^{m} S_j \tag{4.13}$$

where s_i is the ith element of the percentage data, and S_i is the original observed value of s_i, or the area of cultivated land and the area of developed land.

Area percentage data are derived from grid data at a certain grid pixel scale. Area percentage data are constrained by two restriction conditions as follows:

$$\sum_{i=1}^{m} s_i = 1, \quad 0 < s_i < 1 \tag{4.14}$$

$$\sum_{i=1}^{m} S_i = \Omega \tag{4.15}$$

where Ω is constant and represents the area of the grid pixel.

Three main problems must be solved before regression analysis can be conducted using area percentage data. One problem is that the range of values of area percentage data should be located in the interior (between 0 and 1). However, one or several elements usually exist that have values equal to zero. Another problem is that the existence of perfect multicollinearity among variables of area percentage data indicates that the ordinary least squares method is invalid. The final problem is that the regression model must account for the restriction conditions, Eqs. (4.14) and (4.15). We have designed a scheme to overcome these three problems using methods of zero suppression handling and symmetric log-ratio transformation.

Theoretically, it is impossible for the area of some land use types to equal zero if areas are counted at a high enough resolution, that is, areas of some types of land use categories are too small to be detected (Bacon-Shone, 2003). Thus, if the area of one land use category is equal to zero, the area of this type of land use category is assigned a minimal value. Consequently, the sample vector is in the following form

$$\left(s_1',\ s_2',\ \cdots,\ s_p'\right)^T \in [0,\ 1]^p \tag{4.16}$$

where s'_j is the area proportion of the jth land use category in the total area of grid pixels. If $s'_j = 0$, it is assigned a new minimal value $s'_j = \varepsilon$ and $y_i = s'_j \Big/ \sum\limits_{i=1}^{p} s'_i$. The grid area percentage data are then obtained for land use categories:

$$Y = \left\{ (y_1,\ y_2,\ \cdots,\ y_p)^T \in R^p \middle| \sum_{j=1}^{p} y_j = 1,\ 0 < y_j < 1 \right\} \tag{4.17}$$

where y_j is the area proportion of the jth land use category in the total area of grid pixels.

Symmetric log-ratio transformation is conducted after the grid percentage data are treated to stretch the values of area percentage data from (0, 1) to $(-\infty,\ +\infty)$ (4.18).

$$Z = (z_1,\ z_2,\ \cdots,\ z_p)^T, \quad z_j = \ln\left(y_j \Big/ \sqrt[p]{\prod_{i=1}^{p} y_i} \right), \quad j = 1,\ 2,\ \cdots,\ p \tag{4.18}$$

where $z_j \in (-\infty, +\infty)$. Let $s_j = z_j - z_p$, $j = 1, 2, \cdots, p-1$, and through the inverse transformation we can get Eq. (4.19).

$$y_j = \frac{e^{s_j}}{1 + \sum\limits_{i=1}^{p-1} e^{s_i}}, \quad y_p = \frac{1}{1 + \sum\limits_{i=1}^{p-1} e^{s_i}}, \quad j = 1,\ 2,\ \cdots,\ p-1 \tag{4.19}$$

Symmetric log-ratio transformation not only solves the essential zero problem and the problem of constrained total land area, but also linearizes the non-linear relationships between land use patterns and their driving factors. In addition, the transformation retains the symmetry of the original percentage data, and the newly generated variables can be used directly to explore characteristics of the percentage data, making estimation results easily explainable (Paustian et al., 1997; Wang et al., 2008).

4.1.7.2 Partial Least Squares Analysis

Multicollinearity among variables in regression analysis is a problem that must be addressed, as is analysis of driving mechanisms of land use pattern changes. Since its discovery in the 1930s by Frisch (1934), multicollinearity has received increasing attention. Multicollinearity among independent variables always causes deviations of regression estimates, preventing accurate and robust estimations of the coefficients. Without exception, analysis of driving mechanisms of land use patterns faces the same problem. Wold et al. (1983) proposed partial least squares (PLS) regression to tackle the problem. This approach, based on factor analysis, maximizes the covariance between the predicted matrix and the independent matrix composed of factors in the reductive space.

Suppose $X = (X_1, X_2, \cdots, X_q)^T$ is the independent vector variable, and $Z = (Z_1, Z_2, \cdots, Z_q)^T$ is the dependent vector variable. The standardized observed data matrixes of the dependent and independent vector variables after log-ratio transformation are respectively as follows:

$$Z_0 = \begin{bmatrix} z_{11} & \cdots & z_{1p} \\ z_{21} & \cdots & z_{2p} \\ \vdots & & \vdots \\ z_{n1} & \cdots & z_{np} \end{bmatrix}, \quad X_0 = \begin{bmatrix} x_{11} & \cdots & x_{1q} \\ x_{21} & \cdots & x_{2q} \\ \vdots & & \vdots \\ x_{n1} & \cdots & x_{nq} \end{bmatrix} \tag{4.20}$$

The first pair of PLS components

The first pair of components is defined as U_1 and V_1, where U_1 is a linear combination of the independent vector variable X:

$$U_1 = \omega_{11} X_1 + \cdots + \omega_{1q} X_q = \omega_1^T X \tag{4.21}$$

and V_1 is a linear combination of the independent vector variable Z:

$$V_1 = \upsilon_{11} Z_1 + \cdots + \upsilon_{1p} Z_p = \upsilon_1^T Z \tag{4.22}$$

The score-vector of the first pair of components U_1 and V_1 can be calculated and denoted as u_1 and υ_1, respectively.

$$u_1 = X_0 \omega_1 = \begin{bmatrix} x_{11} & \cdots & x_{1q} \\ x_{21} & \cdots & x_{2q} \\ \vdots & & \vdots \\ x_{n1} & \cdots & x_{nq} \end{bmatrix} \begin{bmatrix} \omega_{11} \\ \omega_{12} \\ \vdots \\ \omega_{1q} \end{bmatrix} = \begin{bmatrix} u_{11} \\ u_{21} \\ \vdots \\ u_{n1} \end{bmatrix} \tag{4.23}$$

$$\nu_1 = Z_0 \upsilon_1 = \begin{bmatrix} z_{11} & \cdots & z_{1p} \\ z_{21} & \cdots & z_{2p} \\ \vdots & & \vdots \\ z_{n1} & \cdots & z_{np} \end{bmatrix} \begin{bmatrix} \upsilon_{11} \\ \upsilon_{12} \\ \vdots \\ \upsilon_{1p} \end{bmatrix} = \begin{bmatrix} \nu_{11} \\ \nu_{21} \\ \vdots \\ \nu_{n1} \end{bmatrix} \tag{4.24}$$

The covariance of the first pair of components U_1 and V_1 can be calculated by the inner-product of the score-vectors u_1 and ν_1. Thus the constrained extremum problem, Eq. (4.25), is used to calculate the unit vectors ω_1 and υ_1, which satisfy the qualification that: (i) the first pair of PLS components U_1 and V_1 extracts as much information from the standard observed data matrix as possible; and (ii) the covariate between U_1 and V_1 receives a maximum value.

$$\begin{cases} \max\{\langle u_1,\ \upsilon_1 \rangle\} = \max\left\{\omega_1^T X_0^T Z_0 \upsilon_1\right\} \\ \omega_1^T \omega_1 = \|\omega_1\|^2 = 1,\ \upsilon_1^T \upsilon_1 = \|\upsilon_1\|^2 = 1 \end{cases} \tag{4.25}$$

The constrained extremum problem is solved by calculating the eigenvalue and its corresponding eigenvector of the matrix $Q = X_0^T Z_0 Z_0^T X_0$. The eigenvector of maximum eigenvalue $(\omega_1^T X_0^T Z_0 \upsilon_1)^2$ is the solution of vector ω_1. Vector υ_1 is calculated using Eq. (4.26).

$$\upsilon_1 = \frac{1}{\omega_1^T X_0^T Z_0 \upsilon_1} Z_0^T X_0 \omega_1 \tag{4.26}$$

Regression equations based on the first pair of PLS components

Suppose the regression model with independent variable U_1 and dependent variables X_0 and Z_0 is defined as follows:

$$\begin{cases} X_0 = u_1 \alpha_1^T + S_1 \\ Z_0 = u_1 \beta_1^T + T_1 \end{cases} \tag{4.27}$$

where u_1 is the nth dimension score vector of U_1; $\alpha_1^T = (\alpha_{11}, \alpha_{12}, \cdots, \alpha_{1q})$, and $\beta_1^T = (\beta_{11}, \beta_{12}, \cdots, \beta_{1q})$ are the parameter vectors of the regression model; S_1 and T_1 are the residual matrices. Therefore, the least squares estimates of the regression coefficient vectors α_1 and β_1 are calculated according to Eq. (4.28).

$$\begin{cases} \alpha_1^T = \left(u_1^T u_1\right)^{-1} u_1^T X_0 \\ \beta_1^T = \left(u_1^T u_1\right)^{-1} u_1^T Z_0 \end{cases} \tag{4.28}$$

Final regression equation

Let $X_0' = u_1 \alpha_1^T$ and $Z_0' = u_1 \beta_1^T$, then the residual matrices are illustrated as $S_1 = X_0 - X_0'$ and $T_1 = Z_0 - Z_0'$. Replacing the standard observed data matrices X_0 and Z_0 with S_1 and T_1, respectively, and repeating the above mathematical operation, the weights of the second pair of PLS components, S_2 and T_2, are obtained:

$$\omega_2 = (\omega_{21}, \ \cdots \ , \ \omega_{2q})^T, \quad \upsilon_2 = (\upsilon_{21}, \ \cdots \ , \ \upsilon_{2p})^T \tag{4.29}$$

Then, $\upsilon_2 = T_1 \upsilon_2$ and $u_2 = S_1 \omega_2$ are the score vectors of the second pair of components S_2 and T_2, respectively, and can be calculated. The load capacity of the second pair of PLS components can be calculated with Eq. (4.30).

$$\begin{cases} \alpha_2^T = \left(u_2^T u_2\right)^{-1} u_2^T S_1 \\ \beta_2^T = \left(u_2^T u_2\right)^{-1} u_2^T T_1 \end{cases} \tag{4.30}$$

The generic form of the area percentage data model can then be written as Eq. (4.31)

$$\begin{cases} X_0 = u_1 \alpha_1^T + u_2 \alpha_2^T + S_2 \\ Z_0 = u_1 \beta_1^T + u_2 \beta_2^T + T_2 \end{cases} \tag{4.31}$$

Suppose the rank of data matrix $n \times q$ of X_0 is r, which satisfies $r \leqslant \min(n-1, q)$. The r components and the standardized observation data matrices X_0 and Z_0, can be obtained and further disassembled as shown in Eq. (4.32).

$$\begin{cases} X_0 = u_1\alpha_1^T + \cdots + u_r\alpha_r^T + S_r \\ Z_0 = u_1\beta_1^T + \cdots + u_r\beta_r^T + T_r \end{cases} \tag{4.32}$$

Given that $X_i^*(i = 1, 2, \cdots, q)$ and Z_j^* $(j = 1, 2, \cdots, p)$ are the standardized variables, the values of U_k and Z_j^* can easily be obtained according to Eqs. (4.33) and (4.34).

$$U_k = \omega_{k1}X_1^* + \cdots + \omega_{kq}X_q^* \,, \quad k = 1, \cdots, r \tag{4.33}$$

$$Z_j^* = \beta_{1j}U_1 + \beta_{2j}U_2 + \cdots + \beta_{rj}U_r \,, \quad j = 1, \cdots, p \tag{4.34}$$

After substituting Eq. (4.33) into Eq. (4.34), the PLS regression equation, Eq. (4.35), is obtained.

$$\hat{Z}_j^* = a_{j1}^*X_1^* + \cdots + a_{jq}^*X_q^* \,, \quad j = 1, \cdots, p \tag{4.35}$$

The PLS regression model of original variables, which are included in Eq. (4.36), can be generated by replacing the standardized variables X_i^* and Z_j^* with the original variables X_i and Z_j in Eq. (4.34).

$$\hat{Z}_j = a_{j0} + a_{j1}X_1 + \cdots + a_{jq}X_q \,, \quad j = 1, \cdots, p \tag{4.36}$$

The PLS regression model can also be verified to follow restriction conditions, Eqs. (4.14) and (4.15), and Eq. (4.37) should always hold for any $i = 0, 1, 2, \cdots, q$

$$\sum_j^p a_{ji} = 0 \tag{4.37}$$

Determining of the number of PLS components

Generally, it is not always necessary to obtain all PLS components, which is time consuming when establishing the PLS regression model. The first several PLS components are always enough to explain the regression model. Approaches including leave-one-out, batch-wise cross-validation, split-sample cross-validation and random sample cross-validation are widely used to ascertain the number of obtained components. These methods differ from each other in cross-validation datasets.

The leave-one-out approach leaves ith $(i = 1, 2, \cdots, n)$ observations as the validation data, and the remaining $n - 1$ observations are used to build the PLS regression model.

The batch-wise cross-validation approach follows the same strategy as the leave-one-out approach, except that it uses a sequence of j ordinal observations as the validation dataset. When $j = 1$, the batch-wise cross-validation approach is retrogressed to the leave-one-out approach.

The split-sample cross-validation approach follows the same strategy as the batch-wise cross-validation approach, except that it does not strictly require the observation to be set in an ordinal sequence but is separated by a certain span in the original observation sequence.

In the random sample cross-validation approach, the validation data are randomly chosen.

The estimation result $z'_{j(i)}(k)(i \in I)$ of observation $Z_j(j = 1, 2, \cdots, p)$ can be obtained when the cross-validation dataset $I \subset$ is included in the PLS regression equation with k components. Regarding I, after repeating the above operations, the predictive residual error sum of square of the jth independent variable $Z_j(j = 1, 2, \cdots, p)$ can be calculated using Eq. (4.38) when the kth components have been extracted.

$$PRESS_j(k) = \sum_{I \subset \{1, \cdots, n\}} \sum_{i \in I} \left(z_{ij} - \hat{z}_{j(i)}(k)\right)^2 \tag{4.38}$$

Furthermore, the predicted residual error sum of square of $Z = (Z_1, Z_2, \cdots, Z_p)^T$ is calculated from Eq. (4.39).

$$PRESS(k) = \sum_{j=1}^{p} PRESS_j(k) \tag{4.39}$$

The k that minimizes the predicted residual error sum of square of Z is the number of components to obtain.

4.1.7.3 Neighborhood Effect

Neighborhood enrichment reflects the relative enrichment of one certain land use type in neighbor grids, which can be calculated with the following formula:

$$F_{i,k,d} = \frac{P_{i,k,d}}{P_k} \tag{4.40}$$

where $F_{i,k,d}$ represents the neighborhood enrichment factors; i is the grid number; k is the number of land use types; d stands for the radius of the neighborhood, which is determined with prior knowledge; $P_{i,k,d} = n_{i,k,d}/n_{i,d}$ is the percent of the grid number of the kth land use type in the total grid number in the neighborhood of grid i; $P_k = N_k/N$ is the percent of the grid number of the kth land use type in the total grid number of the study area; $n_{i,k,d}$ is the grid number of the kth land use type in the neighborhood of grid i; $n_{i,d}$ represents the total grid number in the neighborhood of grid i; N_k is the total grid number of the kth land use type in the whole study area; and N is the total grid number in the region

When $F_{i,k,d} = 1$, $F_{i,k,d} < 1$ and $F_{i,k,d} > 1$, the grid enrichment of the kth land use type in the neighborhood with a radius of d of grid i is equal to, or smaller than or larger than that of the whole region, respectively.

Average neighborhood enrichment is an indicator that quantitatively represents the mutual promotion or inhibition effects of different land use types in different neighborhood ranges and is calculated with the following formula:

$$G_{l,k,d} = \frac{1}{N_l} \sum_{i \in l} F_{i,k,d} \tag{4.41}$$

where $G_{l,k,d}$ indicates the average neighborhood enrichment of the lth and kth land use type; N_l is the total grid number of the lth land use type in this region; $i \in l$ indicates the grids that belong to the range of the lth land use type; and $\sum F_{i,k,d}$ represents the sum of neighborhood enrichment of the kth land use type within the domain of the lth land use type. When $G_{l,k,d} > 1$, there are mutual promotion effects in the spatially statistical sense between the lth and kth land use types within the neighborhood with a radius of d, indicating the presences of inhibition effects. The interaction between two land use types in different neighborhood ranges can be quantitatively analyzed by regulating the neighborhood radius d and calculating the average neighborhood enrichment of the lth and kth land use types in different neighborhood ranges.

4.1.8 Spatial Allocation of the Changing Area of Land Uses

4.1.8.1 Decision Rules

In the process of spatially allocating changes in land use area, it is necessary for the DLS model to set certain decision rules for the various land use types with different degrees of stability to restrict the actions of land use change in the model according to historical changes in various land use types and land use planning. If historical data and future land use planning indicate that a certain land use type is prone to converting to another land use type, it is set to be easily converted to other land uses; otherwise, it can be set to relatively stable or hard to be converted to other land uses. Spatial allocation of changes in land use area simulates the actual stability of various land use types by assigning proper values to the stability parameters, which constitute the following stability parameter matrix.

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1i} & \cdots & a_{1p} \\ a_{21} & a_{22} & \cdots & a_{2i} & \cdots & a_{2p} \\ \vdots & \vdots & & \vdots & & \vdots \\ a_{j1} & a_{j2} & \cdots & a_{ji} & \cdots & a_{jp} \\ \vdots & \vdots & & \vdots & & \vdots \\ a_{p1} & a_{p2} & \cdots & a_{pi} & \cdots & a_{pp} \end{pmatrix} \tag{4.42}$$

where a_{ji} is the stability parameter of conversion of land use type j to land use type i; a_{ii} is the stability parameter of the area of land use type i that remains unchanged. Three conditions are usually met when values are assigned to stability parameters. First, for the conversion of one land type to another that has been historically stable and unrelated to land planning, the corresponding stability parameter in the model is assigned to 1. Second, for land use types that are prone to conversion, the corresponding stability parameters in the model are assigned to 0. Third, for most land use types, the possibility of conversion is between the two extremes, and the stability parameters in the model are assigned values between 0 and 1. The larger the stability parameter, the more stable the corresponding land use type and the lower the possibility of mutual conversion. In addition, it is necessary to further verify and regulate the process of checking the model because the configuration of stability parameters in the model mainly depends on experts' experience and researchers' understanding of actual conditions of the study area.

In the process of spatially allocating changes in land use area, the stability of certain land use types defined with stability parameters in the model is restricted by two aspects, i.e., when one land use type is not prone to convert to other land use types and when other land use types are not prone to convert to this land use type. Many uncertainties exist in land use change. In fact, in model simulations, it may be very difficult or impossible for land use change to occur in a certain direction, but conversion in the reverse direction may be very common (Deng et al., 2008). For example, it is costly and uncommon to convert water areas to cultivated land, but in southern China, where the demand for aquatic produce is high and the price continues to rise, many paddy fields are converted into fish or shrimp ponds. Under these conditions, the spatial allocation of changing land use area defines the decision rules of these conversions by constructing the land use conversion rule matrix.

4.1.8.2 Allocation Steps

The input parameters used in the module for spatially allocating changes in land use area reflect local, regional and historical characteristics of the pattern changes in regional land uses (Fig. 4.2). Specific steps are shown in the following figure.

The spatial allocation module of land use first calculates the number of grids to allocate according to the conversion rules set for each pixel and the land areas to be allocated over space. It then calculates the allocation probability L_{ik} of different land use types for the grid to allocate. Finally, it allocates the land use pattern with the obtained allocation probability L_{ik} and obtains the change rules for the regional land use pattern.

Generally, the allocation probability of different land use types L_{ik} is determined based on the following three situations:

(i) If a certain land use type existed in the previous simulation year, and its stability is less than 1, the spatial allocation module will calculate the distribution probability, sum of the compensation factor and stability factor,

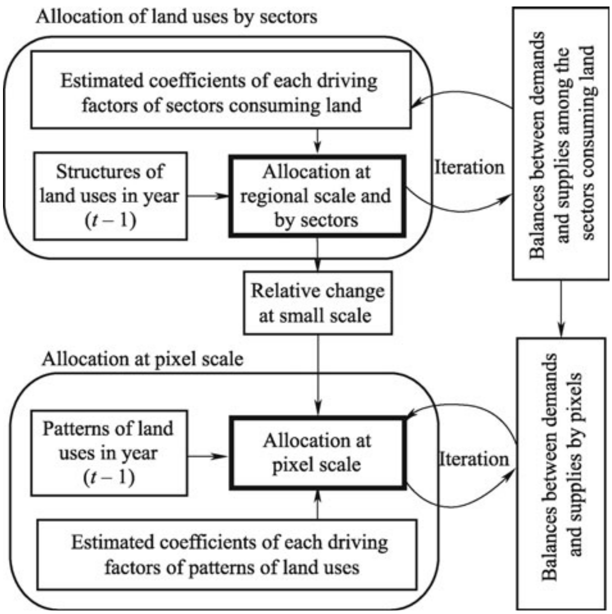


Fig. 4.2 Steps of spatial allocation of area change in land use in the DLS model.

which are used as the allocation probability:

$$L_{i,k} = P_{i,k} + C_k + S_k \tag{4.43}$$

where L_{ik} is the allocation probability of the kth land use type in grid i; P_{ik} is the distribution probability of the kth land use type in grid i; C_k and S_k are the compensation and stability factors, respectively.

(ii) When the compensation factor C_k is nearly 0, L_{ik} consists of the distribution probability P_{ik} and stability factor S_k as follows:

$$L_{i,k} = P_{i,k} + S_k \tag{4.44}$$

(iii) Within each spatial allocation step, the DLS model excludes those pixels with a decreasing trend for a certain kind of land use type from obtaining new areas of that kind of land use type. If the spatial allocation module does not allow the configuration of stability, then the land use type with the largest L_{ik} is allocated to the grids without enough area of land use types (Fig. 4.3).

When the area of the study site is small and the geophysical conditions are relatively consistent, the area change can be directly allocated with the method mentioned above. Conversely, if the area of the study site is large and there is significant spatial difference in regional geophysical conditions, it is more feasible to first zone and then allocate the area change based on the spatial distribution of geophysical conditions (Gao and Deng, 2002; Deng et al., 2008).

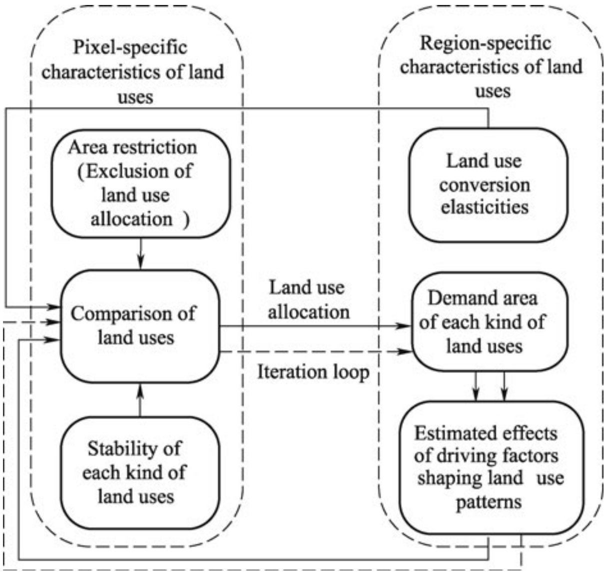


Fig. 4.3 Schematic representation of the spatial allocation of land use changes at the regional and pixel scales.

In summary, the DLS model is an effective tool to identify various factors that influence the distribution of regional land use types, reveal driving mechanisms of land use pattern changes and reflect the pattern changes in land uses in grids at certain scales.

4.2 DLS Installation and Configuration

Installation and configuration of DLS software are two main steps before development of the DLS model.

4.2.1 DLS Installation

DLS Software is a software tool for the dynamic simulation of the land use pattern which was developed based on the DLS model. The latest version of DLS was released in 2007. It has developed into a program package that can be installed independently. DLS in version 2007 can be used in Microsoft Windows and was developed for Windows XP Professional. However, the software has passed testing in Windows 9x, Me, 2000/NT, 2003 Server and XP Home environments.

In the Windows operating system, insert the installation program CD in the CD-ROM, click "DLS MODEL setup program" in the root directory, enter the installation directory, double-click the installation file "Setup.exe" and then enter the installation interface. Follow the instructions to complete the installation.

4.2.2 Configuring the DLS Operating Environment

After installation, the DLS model interface (Fig. 4.4) can be opened by clicking "Start" → "All Programs" → "DLS MODEL." The DLS user interface is user friendly, and the user can directly configure the model operating environment and input the parameters and variables in the main interface.

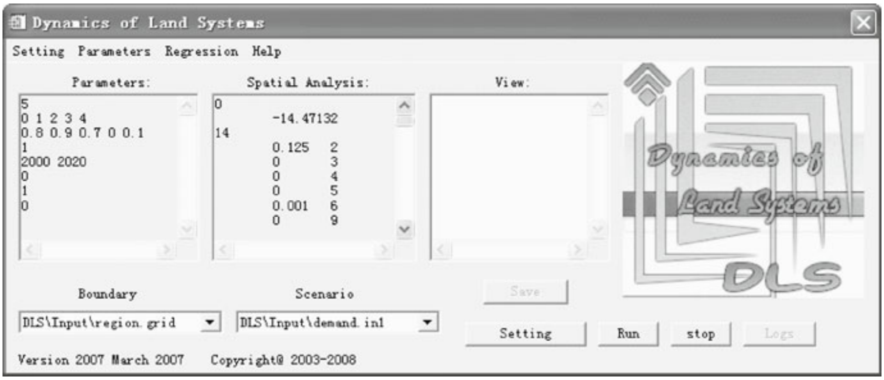


Fig. 4.4 User interface of DLS.

The operating system can be configured either in the Settings pull-down menu or the Settings button in the main interface. After starting the DLS model, set the file allocation running path and the name of the control file (Fig. 4.5). Select the control file allocation input parameters, the storage path of the output results and the names of the restricted zone data and those of land demand scenario data (Fig. 4.6).

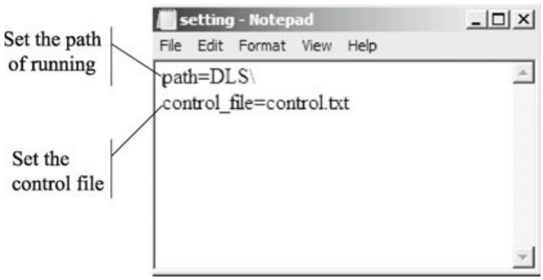


Fig. 4.5 Setting of running environment of DLS.

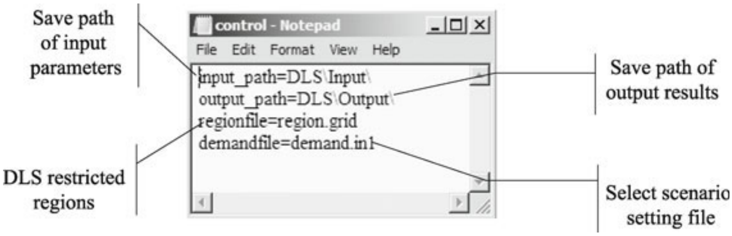


Fig. 4.6 Configurations of input and output parameters for running the DLS model.

If DLS is installed in the Linux operating system, it is necessary to use "/" when setting the file path, while "\" is used in a Windows system. All configuration files are saved in the file folder in the installation directory.

4.3 DLS Input Parameter Preparation

Six kinds of parameters are needed for DLS, i.e., simulation condition setting parameters, driving factor data, spatial analysis parameters, restricted region code, land demand scenario data and binary data of land types. The main DLS interface is convenient for the user to input and change the parameters (Fig 4.5). Meanwhile, to make the DLS program automatically read the related files, the user can use an ASCII data file prepared in another software environment and directly copy it to the DLS installation directory or the simulation directory set by the user.

4.3.1 Simulation Condition Setting Parameters

Parameters in the "Parameters" window at the far left can be edited through the "Input" menu under the "Parameters" menu. Once the user begins to edit the parameters, the "Parameters" window will change from a grey inactive window to an active window where the user can edit the parameters. After editing parameters, the user can store changes in the parameters in the simulated pattern changes in land uses by clicking the "Save" button.

Parameters in the "Parameters" window, which the user can also edit with Notepad software or a text editor, are included in the "main.1" file under the installation directory "DLS\Input." The user can also create a new file in Notepad software or a text editor, save it as a file in "main.1" format and put it under the installation directory "DLS\Input." The main parameters are included in the following figure (Fig. 4.7).

Line 1: Number of land use types;

Line 2: Land use type codes, starting from 0;

Line 3: Decision rules corresponding with the land use type;

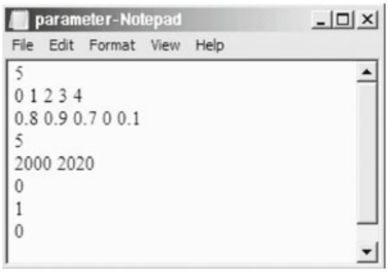


Fig. 4.7 Compilation of main file for running DLS.

Line 4: Condition for convergence: the permissible error when the land demand changes and actual allocated land converge;

Line 5: Initial and ending year of simulation;

Line 6: Number of driving factors for the distribution of land use types;

Line 7: Record format mark of the main file: 1 indicates that the ArcView header file will be exported in the output file; 0 indicates that it will not.

4.3.2 Spatial Analysis Parameters

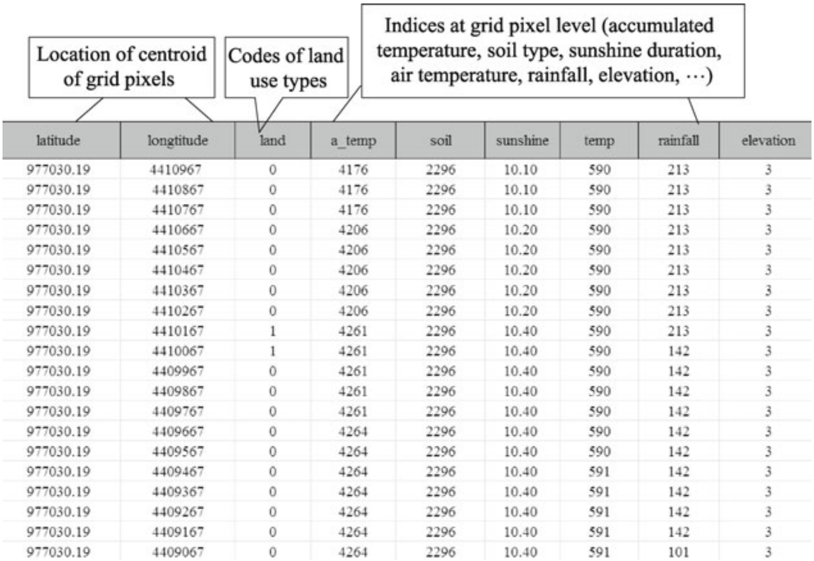
The main component of the user interface is the Spatial Analysis window, which contains regression equations of different land use types and driving factors. The user can edit this file by clicking "Modify" under the "Regression" menu. Another option is to open and modify the "alloc1.reg" file under the installation directory "DLS\Input."

The regression coefficients between different land types and driving factors are calculated with the following steps.

(i) Resample the GRID data of land types and driving factors at a certain spatial scale with the "SAMPLE" order in the Arc module of ArcGIS software. Save the result in a text file, the first line of which records the variable names of the corresponding land type and driving factors (Fig. 4.8).

(ii) Calculate the coefficients of determination between the land type and driving factors with statistical software capable of logistic regression analysis (e.g., SPSS software) and ENMLUP with the following procedures: open the logistic regression dialog box from the menu Analyze → Regression → Binary Logistic. Select the independent variables and related parameters, and perform the calculation to obtain the regression coefficients between the land type and driving factors (Table 4.1). Estimate the correlation between the land types and driving factors according to "S.E.". An S.E. larger than 0.2 indicates that the correlation is not strong, and the coefficient of determination can be deleted.

(iii) Delete the parameters with weak correlation, and create a file named "alloc1.reg" in the following format shown in Fig. 4.9. Save it under the installation directory "DLS\Input."



latitude	longtitude	land	a_temp	soil	sunshine	temp	rainfall	elevation
977030.19	4410967	0	4176	2296	10.10	590	213	3
977030.19	4410867	0	4176	2296	10.10	590	213	3
977030.19	4410767	0	4176	2296	10.10	590	213	3
977030.19	4410667	0	4206	2296	10.20	590	213	3
977030.19	4410567	0	4206	2296	10.20	590	213	3
977030.19	4410467	0	4206	2296	10.20	590	213	3
977030.19	4410367	0	4206	2296	10.20	590	213	3
977030.19	4410267	0	4206	2296	10.20	590	213	3
977030.19	4410167	1	4261	2296	10.40	590	213	3
977030.19	4410067	1	4261	2296	10.40	590	142	3
977030.19	4409967	0	4261	2296	10.40	590	142	3
977030.19	4409867	0	4261	2296	10.40	590	142	3
977030.19	4409767	0	4261	2296	10.40	590	142	3
977030.19	4409667	0	4264	2296	10.40	590	142	3
977030.19	4409567	0	4264	2296	10.40	590	142	3
977030.19	4409467	0	4264	2296	10.40	591	142	3
977030.19	4409367	0	4264	2296	10.40	591	142	3
977030.19	4409267	0	4264	2296	10.40	591	142	3
977030.19	4409167	0	4264	2296	10.40	591	142	3
977030.19	4409067	0	4264	2296	10.40	591	101	3

Fig. 4.8 Snapshot of the variables used by the DLS model to estimate the land use conversion elasticity.

Table 4.1 Output results of the ENMLUP estimation

Variables in the equation							
		B	S.E.	Wald	df	Sig.	Exp(B)
Step 1[a]	*Ctemp*	.002	.001	3.860	1	.049	1.002
	Erosion	−.009	.007	1.493	1	.222	.992
	shining_hr	−1.052	.336	9.801	1	.002	.349
	Temp	.046	.002	704.854	1	.000	1.047
	Rainfall	−.018	.000	3.047E3	1	.000	.982
	Elevation	.758	.017	2.041E3	1	.000	2.134
	Loam	−.464	.018	653.227	1	.000	.629
	Splain	−.003	.000	3.360E3	1	.000	.997
	d2expway	.005	.009	.232	1	.630	1.005
	Landform	.009	.001	72.306	1	.000	1.009
	Organic	−1.763	.029	3.799E3	1	.000	.172
	d2hwy	.081	.005	295.466	1	.000	1.084
	d2pvcap	−.046	.008	36.404	1	.000	.955
	d2pvway	.188	.013	222.260	1	.000	1.207
	d2road	.019	.006	10.530	1	.001	1.019
	Popdensity	.064	.008	66.277	1	.000	1.066
	GDP	−.001	.000	479.715	1	.000	.999
	Constant	−31.984	14.830	4.651	1	.031	.000

a. Variable(s) entered in step 1: *ctemp*, *erosion*, *shining_hr*, *temp*, *rainfall*, *elevation*, *loam*, *splain*, *d2expway*, *landform*, *organic*, *d2hwy*, *d2pvcap*, *d2pvway*, *d2road*, *popdensity*, *GDP*.

Line 1: Codes of each land type;

Line 2: Constant of spatial regression results of each land type;

Line 3: Number of driving factors for the distribution of land types;

Line 4 (and subsequent lines): Codes of the regression coefficients and driving factors.

The next land types after the regression coefficients of one land type are listed, and the regression coefficients of all land types that need to be calculated will be listed in this file.

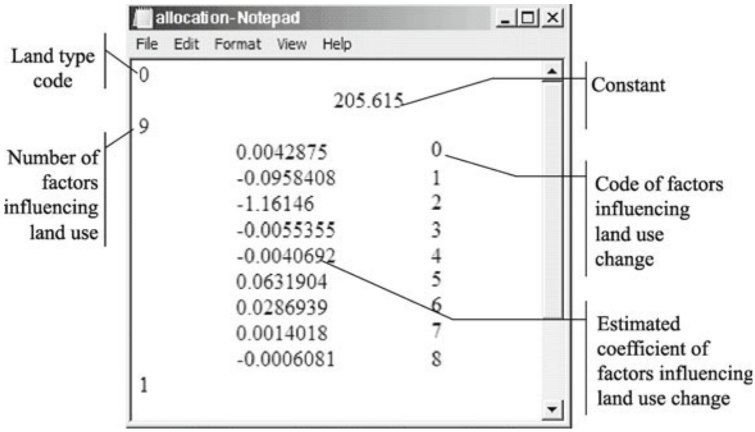


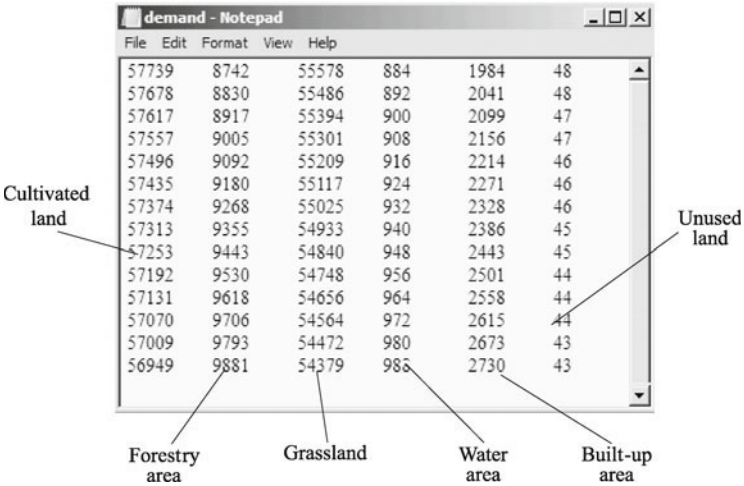
Fig. 4.9 Formatted parameters exported from the ENMLUP estimation.

4.3.3 Driving Factor Data

Driving factor data are saved in the files under the installation directory "DLS\Input." All factors that affect the distribution of land types such as natural variables, socioeconomic variables and policies are prepared with ArcGIS software according to characteristics of each variable. All data are disaggregated onto the spatial unit with a certain spatial resolution, and the final dataset is saved as ASCII data in "sc1gr*.grd" format. The "*" in different driving force file names is replaced with a number corresponding to the code numbers of the impact factors. The process of transforming the data in ArcView and ArcGIS software is as follows. (i) Transform the GRID data into text data with the File → Export Data Source tool; (ii) Select the output data format; as the output data are GRID data, the ASCII Raster format is selected; (iii) Select the GRID data file that needs to be output from the storage directory of the GRID data; (iv) Choose the save file path, input the name of the output text data and save the text file in the selected path.

4.3.4 Land Demand Scenario Data

The land demand scenario data are saved under the installation directory "DLS\Input." The land demand scenario file is named in the "demand.*" format the corresponding selection list of which is displayed in the "Scenario Design" pull-down menu in the main interface. This can be opened and edited with a text editor such as Notepad, and each line of this file records the demand of different land types in the simulation period (Fig. 4.10).



57739	8742	55578	884	1984	48
57678	8830	55486	892	2041	48
57617	8917	55394	900	2099	47
57557	9005	55301	908	2156	47
57496	9092	55209	916	2214	46
57435	9180	55117	924	2271	46
57374	9268	55025	932	2328	46
57313	9355	54933	940	2386	45
57253	9443	54840	948	2443	45
57192	9530	54748	956	2501	44
57131	9618	54656	964	2558	44
57070	9706	54564	972	2615	44
57009	9793	54472	980	2673	43
56949	9881	54379	988	2730	43

Fig. 4.10 Scenario-based land demands for running DLS.

The land demand scenario data are prepared as follows: estimate the rate of change for each land type according to land use planning data for the subsequent 20 years and statistical data of land types of recent years. Then calculate the land type data of each year with a rate of change with $N_{i+1} = N_i^*(1 + b)$.

The area sum of all land types should be equal to the regional total area prior to performing the calculation. Thus, it is necessary to make adjustments according to the regional total area and land use pattern.

4.3.5 Restricted Region Code Data

The restricted region files under the installation directory "DLS\Input" are named in the "regi*.*" format, and their names appear in the region selection box "Area restriction." The GRID files included in the restricted region files are in a rectangular format; only these active grids located in non-restricted areas can participate in simulations of land use changes, where the value "0" stands for the active grid and "−9999" null and "−9998" stand for restricted regions (Fig. 4.11). The sort order of values in this file corresponds with that of the actual grids. Preparation of the restricted region code data are similar to that of the driver data, which can be realized by saving the GRID data as a text file with the Export Data tool.

```
__++__  simulation for year   1 __+__
  0    205.6150    9
  1   -116.2159    9
  2   -36.26043    9
  3    85.94983    9
  4    153.8026    8

demand direction for cove 0 is -1; demand:32568.7
demand direction for cove 1 is 1; demand:120236.
demand direction for cove 2 is 1; demand:21463.6
demand direction for cove 3 is 1; demand:19147.6
demand direction for cove 4 is -1; demand:28510.6
```

	①	②	③	④
1	40725.50	0.5006190	25.04491	76.79225
2	40717.50	0.5010777	25.02034	76.75047
3	40701.50	0.5013463	24.97122	76.66692
4	40681.50	0.5018201	24.90982	76.54158
5	40673.50	0.5021290	24.88526	76.54158
6	40665.50	0.5026180	24.86069	76.47891
7	40665.50	0.5026758	24.86069	76.47891
8	40665.50	0.5044107	24.86069	76.45803
9	40637.50	0.5047202	24.77472	76.31180
10	40625.50	0.5051681	24.73788	76.20735
11	40629.50	0.5054268	24.75016	76.16558
12	40629.50	0.5056825	24.75016	76.16558
13	40353.52	0.5059981	23.90278	73.80512
14	40337.52	0.5070848	23.85365	73.63800
15	40317.52	0.5073349	23.79224	73.55444
16	40305.52	0.5081456	23.75540	73.47089
17	40289.53	0.5083910	23.70628	73.26199
18	40257.53	0.5086447	23.60803	73.05311
19	40249.53	0.5094516	23.58347	72.92777
20	40241.53	0.5097812	23.55891	72.90688
21	40233.53	0.5115739	23.53435	72.82333

Fig. 4.11 Representation of the null value, restricted and active grid pixels by evaluating −9999, −9998 and 0, respectively.

4.3.6 Land Use Type Data

Many methods can be used to prepare the land-type binary data. The method used in the Workstation program in the ArcGIS software environment will be introduced in Chapter 6. Another method with ArcView software is introduced as follows.

First, open the grid data for the land type and boundary. Select Analysis → Property to open the dialog box, and select the boundary grid data in Analysis Mask.

Then, open the dialog box through Analysis → Map Calculation, and input the calculation formula to extract the binary grid data with the land type codes 1, 2, 3, 4, 5 and 6, respectively.

Save the calculation results as follows: open Theme → Storage Data Set, select the storage path and input the storage file name.

Finally, transform the binary data into text data named in the "cov1_#.0" format, and then save it under the installation directory "DLS\Input."

4.4 DLS Operation and Results Output

The model must be reconfigured when it is used in a new region and the model results are saved in two formats according to the year.

4.4.1 DLS Operation

After reconfiguration, select the restricted region and scenario files and click the "Run" button. The model will run automatically, and the results will be listed one by one in the view window. These can be saved in the file "DLS\Output," and the running process will be saved in the "Log" file under the output directory, which includes not only the running information but also the iteration parameters listed in every step of the operation. All these parameters are listed in the operation steps. The implication of the codes annotated is as follows. (i) Number of iterations; (ii) Allocated quantity of each land use type; (iii) Iteration parameters of each land use type; (iv) Gaps between the demanded quantity and allocated quantity; (v) Maximum gaps between the demanded quantity and allocated quantity.

4.4.2 Results Output

The results of each land type are saved as an "*out1_#.**" file where "#" is the land type code and "*" is the year to calculate. The results of all land types every year are saved as "out all.*" files, in which a single value of each land type is saved. The new values of all grids are also saved in these result files, which can be read by ArcView GIS software. ArcView GIS reads data as follows. (i) If the seventh line of the "main.1" file is 1, the ArcView title will be output in the output file, which can be directly saved in ASCII-GRID format. This kind of data can be imported through File (Import Data Source) in ArcView GIS software; (ii) Run this module, and select the input data type in the pop-up window. Here the ASCII Raster data type is selected; (iii) Select the name of the layer to convert (a single land type or all land types); (iv) Input the name of the output data. The output data will be saved in the selected path. "Integer" format is selected as the type of output grid unit data, which can be directly opened in the active window.

4.5 Summary

The DLS model is a powerful tool for simulating the dynamics of land use changes. It comprehensively considers the control of external demand and influences of various neighborhood driving factors, emphasizes internal suitability, controls random disturbance factors, has specific decision rules and constructs multiple objective functions. Therefore, it is robust for simulating land use pattern changes in terms of both the expression of mechanisms and simulation effects (Deng et al., 2008; Zhan et al., 2007).

DLS is a software tool developed based on the DLS model for the dynamic simulation of the land use pattern changes. This chapter introduces the main procedures of DLS including installation, parameter configuration, running steps and results output. It can measure the influence of driving factors that are closely associated with changes in the land use pattern, including natural conditions, socioeconomic factors and even land use management policies. It can simulate the spatiotemporal process of pattern changes in land uses and export maps of pattern changes in land uses with high spatial and temporal resolution by setting conversion rules of land use types and designing change scenarios.

References

Aguiar A P D, Câmara G, Escada M I S, 2007. Spatial statistical analysis of land-use determinants in the Brazilian Amazonia: Exploring intra-regional heterogeneity. Ecological Modelling, 209(2–4): 169–188.

Bacon-Shone J. 2003. Modelling structural zeros in compositional data. In S. Thió-Henestrosa and Martín-Fernández J A (Eds.), Proceedings of CODAWORK'03, The First Compositional Data Analysis Workshop. Spain: University of Girona. pp. 15–17.

Burgi M and Turner M G. 2002. Factors and processes shaping land cover and land cover changes along the Wisconsin River. Ecosystems, 5(2): 184–201.

Burgi M, Hersperger A M, Schneeberger N. 2004. Driving forces of landscape change—Current and new directions, Landscape Ecology, 19, 857–868.

Chen Y Q. 2000. Multi1scale spatial characterization of land use/land cover in China. Scientia Geographica Sinica, 20(3): 198–202.

Dai E, Wu S, Shi W, Cheung C K, Sharker A. 2005. Modeling change-pattern-value dynamics on land use: An integrated GIS and artificial neural networks approach. Environmental Management, 36(4): 576–591.

Deng X Z, Su H B, Zhan J Y. 2008. Integration of multiple data sources to simulate the Dynamics of Land Systems. Sensors, 8(2): 620–634.

Deng X Z, Jiang Q O, Lin Y Z, Han J Z. 2010a. Simulation of changes of soil organic carbon stock of cropland in China. Geographical Research, 29(1): 93–101 (in Chinese).

Deng X Z, Huang J K, Rozelle S, Uchida E. 2010b. Economic growth and the expansion of urban land in China. Urban Studies, 47(4): 813–843.

Ferrers N M. 1866. An Elementary Treatise on Trilinear Coordinates. London: Macmillan.

Frisch R. 1934. Statistical Confluence Analysis by Means of Complete Regression Systems. Oslo: Universitetets Okonomiske Institute.

Gao Z Q and Deng X Z. 2002. Analysis on spatial features of LUCC based on remote sensing and GIS in China. Chinese Geographical Science, 12(2): 107–113.

Gao Z Q, Liu J Y, Cao M K, Li K R, Tao B. 2005. Impacts of land-use and climate changes on ecosystem productivity and carbon cycle in the cropping-grazing transitional zone in China. Science in China Series D—Earth Sciences, 48(9): 1479–1491.

Ge Q S and Dai J H. 2005. Farming and forestry land use changes in China and their driving forces from 1900 to 1980. Science in China Series D—Earth Sciences, 48(10): 1747–1757.

Haberl H, Wackernagel M, Wrbka T. 2004. Land use and sustainability indicators. An introduction. Land Use Policy, 21(3): 193–198.

Lambin E F, Turner B L, Geist H J, Agbola S B, Angelsen A, Bruce J W, Coomes O T, Dirzo R, Fischer G, Folke C, George P S, Homewood K, Imbernon J, Leemans R, Li X, Moran E F, Mortimore M, Ramakrishnan P S, Richards J F, Skanes H, Steffen W, Stone G D, Svedin U, Veldkamp T A, Vogel C, Xu J. 2001. The causes of land-use and land-cover change: Moving beyond the myths. Global Environmental Change, 11(4): 261–269.

Lambin E F, Geist H J, Lepers E. 2003. Dynamics of land-use and land-cover change in tropical regions. Annual Review of Environment and Resources, 28: 205–241.

Li L, Zhang P Y, Hou W. 2005. Land use/cover change and driving forces in Southern Liaoning Province since 1950s. Chinese Geographical Science, 15(2): 131–136.

Liu J Y, Deng X Z, Liu M L, Zhang S W. 2002. Study on the spatial patterns of land-use change and analyses of driving forces in Northeastern China during 1990—2000. Chinese Geographical Science, 12(4): 299–308.

Liu X H, Wang J F, Liu M L, Meng B. 2005. Spatial heterogeneity of the driving forces of cropland change in China. Science in China Series D—Earth Sciences, 48(12): 2231–2240.

Paustian K, Levine E, Post W M, Ryzhova I M. 1997. The use of models to integrate information and understanding of soil C at the regional scale. Geoderma, 79(1–4): 227–260.

Pontius R G, Walker R, Yao-Kumah R, Arima E, Aldrich S, Caldas M, Vergara D. 2007. Accuracy assessment for a simulation model of Amazonian deforestation. Annals of the Association of American Geographers, 97(4): 677–695.

Shi P J, Gong P, Li X B, Chen J, Qi Y, Pan Y Z. 2000. Method and Practice of Study on Land Use/Land Cover Changes. Beijing: Science Press.

Tobler W. 1970. A computer movie simulating urban growth in the detroit region. Economic Geography, 46(2): 234–240.

Turner B L Ⅱ. 1997. The sustainability principle in global agendas: Implications for understanding land-use/cover change. The Geographical Journal, 163(2): 133–140.

Turner B L Ⅱ, Lambin E F, Reenberg A. 2007. The emergence of land change science for global environmental change and sustainability. Proceedings, National Academy of Sciences of the United States of America, 104(52): 20666–20671.

Veldkamp A and Lambin E F. 2001. Predicting land-use change. Agriculture Ecosystems and Environment, 85(1–3): 1–6.

Veldkamp A and Verburg P H. 2004. Modelling land use change and environmental impact. Journal of Environmental Management, 72(1–2): 1–3.

Verburg P H. 2006. Simulating feedbacks in land use and land cover change models. Landscape Ecology, 21(8): 1171–1183.

Verburg P H, Chen Y Q, Veldkamp T A. 2000. Spatial explorations of land use change and grain production in China. Agriculture Ecosystems and Environment, 82(1–3): 333–354.

Verburg P H, Soepboer W, Veldkamp A, Limpiada R, Espaldon V, Mastura S S A. 2002. Modeling the spatial dynamics of regional land use: The CLUE-S model. Environmental Management, 30(3): 391–405.

Wang S Y, Liu J S, Ma T B. 2010. Dynamics and changes in spatial patterns of land use in Yellow River Basin, China. Land Use Policy, 27(2): 313–323.

Wang X D, Zhong X G, Liu S Z, Li M H. 2008. A non-linear technique based on fractal method for describing gully-head changes associated with land-use in an arid environment in China. Catena, 72(1): 106–112.

Wold S, Albano C, Dunn W J Ⅲ, Esbensen K, Hellberg S, Johansson E, Sjõstrõm M. 1983. Pattern recognition: Finding and using patterns in multivariate data. In: Martens H and Russwurm H Jr (Eds.), Food Research and Data Analysis. London: Applied Science Pub, pp. 147–188.

Zhan J Y, Deng X Z, Jiang Q O, Shi N N. 2007. The application of system dynamics and CLUE-S model in land use change dynamic simulation: A case study in Taips County, Inner Mongolia of China. Proceedings of the 2007 Conference on Systems Science, Management Science and System Dynamics: Sustainable Development and Complex Systems, 1–10: 2781–2789.

Zhang X Q, Ge Q S, Zheng J Y. 2003. Analysis of driving forces on the change of historical farmland use in Shandong Province during the Qing Dynasty (1644—1911). Chinese Geographical Science, 13(3): 207–215.