The 18th Biennial Conference of International Society for Ecological Modelling

A Software Tool for Earth Surface Modeling of Environmental variables

Chen-Liang Wang a, Tian-Xiang Yue a,b*

aInstitute of Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences, 11A, Datun Road, Anwai, 100101 Beijing, China;
bEcological Complexity and Modeling Laboratory, University of California, Riverside, CA 92521-0124, USA

Abstract

A software tool for earth surface modeling of climate change through high-accuracy and high-speed methods are presented here. Environmental variables including temperature, rainfall and soil property constitute critical driving factors of ecosystem modeling. However, there are a couple of challenges behind modeling environmental factors including heterogeneity of multiple data sets. To handle compatibility between heterogeneous data sets, the modeling process need to represent environmental data in grid cells with earth surface modeling. However, there are error problems in classical methods of surface modeling. Furthermore, there is still limited experience in developing multi-scale surface modeling tool of climate change for ecosystem modeling. High-accuracy and high-speed methods for surface modeling (HASM) were used to solve error problems, improve the accuracy of spatial distribution simulation of environmental variables and handling multi-scale information. However, HASM which is more theoretically perfect and more complicated than other classical methods of surface modeling. This tool decompose the HASM modeling process into several components including HASM initial field, HASM optimistic controller and HASM iterative solver. HASM can opt remote sensing images or existing short precision grid data for initial field and can also simulate filed when lacking of initial fields. HASM optimistic controller need to get knowledge from environmental database to generate controlling condition rulers ensuring modeling more reasonable and more precision. HASM iterative solver is responsible for solution of HASM by means of integration of various solve methods which we have developed. Rather than linking environmental data with these modeling components, a supermatic surface modeling approach ensuring effective information exchange components is also designed for HASM. This software is easily applied to the development of ecosystem modeling integrated with surface modeling. This study from this integration is also helpful for future researches that aim to integrate surface modeling and ecosystem modeling.

© 2011 Published by Elsevier B.V. Selection and/or peer-review under responsibility of School of Environment, Beijing Normal University. Open access under CC BY-NC-ND license.

* Corresponding author. Tel.: +0086-10-64889041; fax: +0086-10-64889041.
E-mail address: txyue@ucr.edu
1. Introduction

Nowadays, environmental and ecological modeling processes have been interdisciplinary modeling processes. Various natural indicators, social factors and different technologies are linked together to investigate the driving mechanism [1-6]. However, there are a couple of challenges behind modeling that is heterogeneity of multiple data sets. To handle compatibility between heterogeneous data sets, environmental and other data can be represented in united unstructured grid cells with surface modeling. Several methods of surface modeling including IDW, kriging, Spline and others were developed and applied to estimate spatial distribution or convert the point data to continuous surface in the various natural domains [7-12]. But there are some critical problems in surface modeling. Error problem has been studied being a vital factor affecting the accuracy of modeling especially in the DEM creation [13,14]. The accuracy of results can be seriously affected by several conditions such as sample and grid size, different methods. [15,14,16-18] And user may obtain unqualified data in a certain field ignoring the accuracy or reliability of modeling [19,16,20]. To obtain necessarily high accuracy the computation cost must be moderately increased and difficult to deal with [18]. HASM (high accuracy and high speed surface modeling) focus on these two major issues of surface modeling. It is a relatively more accurate and quick that classical methods solving the error problem theoretically. Furthermore, it has been applied to various ecological and environmental contexts and achieving the desired effect of accuracy and speed [21-25]. However, there is no easy interface or framework of HASM. And HASM is based on theory of surface and solves partial differential equations (PDEs). Obviously, modelers dislike such complicated models because of their gaps between the theory of mathematics and modeling implementation [2]. Reuse and extension of existing work are constricted. This paper presents a software framework of HASM which enables user to perform HASM quickly and automatically completes other data operations.

2. HASM description

HASM is a newly developed surface modeling method. In terms of the fundamental theorem of surfaces, a surface is uniquely defined by the first and second fundamental coefficients which constitute Gauss-Codazzi equations [26]. HASM is based on the second order PDEs of Gauss-Codazzi equations for surface modeling. By means of the solution of elliptic partial differential equations and optimum formulation of sampled values, HASM can obtain more relatively accurate results than classical methods [21-23].

2.1. Theoretical Formulation

The most suitable PDEs of HASM could be expressed as [27],

\[
\begin{align*}
 f_{xx} & = \Gamma_{11} f_x + \Gamma_{12} f_y + \frac{L}{\sqrt{E + G - 1}} \\
 f_{yy} & = \Gamma_{22} f_x + \Gamma_{22} f_y + \frac{N}{\sqrt{E + G - 1}}
\end{align*}
\] (1)
Where \( E = 1 + f_x^2; F = f_x f_y; G = 1 + f_y^2; L = \frac{f_{xx}}{\sqrt{1 + f_x^2 + f_y^2}}; N = \frac{f_{yy}}{\sqrt{1 + f_x^2 + f_y^2}} \);

\[
\Gamma_{11}^1 = \frac{GE_x - 2FF_x + FE_y}{2(EG - F^2)}; \quad \Gamma_{22}^1 = \frac{2GF_y - GG_x - FG_y}{2(EG - F^2)};
\]

\[
\Gamma_{11}^2 = \frac{2EF_x - EE_y - FE_x}{2(EG - F^2)}; \quad \Gamma_{22}^2 = \frac{EG_x - 2FF_y + FG_x}{2(EG - F^2)};
\]

Firstly, HASM calculates the first and second coefficients, \( E, F, G, L \) and \( N \) according to sampled values. \( \{ \tilde{f}_{i,j} \} \) are denoted as the sampled value of \( f \) at sampling points \( \{(x_i, y_j)\} \) and \( \{ \tilde{f}_{i,j} \} \) are interpolations in terms of the sampled values \( \{ \tilde{f}_{i,j} \} \). Let \( f_{i,j}^n = \tilde{f}_{i,j} \) and \( h \) represent simulation step length, \( f_{i,j}^n (n \geq 0, 0 \leq i \leq I + 1 \text{ and } 0 \leq j \leq J + 1) \) are the \( n \)th iteration values of lattices whose centers are points of \( \{(x_i, y_j)\} \), then finite difference of the basic equations could be formulated as,

\[
\begin{align*}
\frac{f_{i+1,j}^n - 2f_{i,j}^n + f_{i-1,j}^n}{h^2} = & \left( \Gamma_{11}^1 \right)_{i,j} f_{i+1,j}^n - f_{i-1,j}^n + \left( \Gamma_{11}^1 \right)_{i,j} f_{i,j+1}^n - f_{i,j-1}^n + \frac{L_{i,j}^n}{\sqrt{E_{i,j}^n + G_{i,j}^n - 1}} \quad (2) \\
\frac{f_{i,j+1}^n - 2f_{i,j}^n + f_{i,j-1}^n}{h^2} = & \left( \Gamma_{22}^1 \right)_{i,j} f_{i+1,j}^n - f_{i-1,j}^n + \left( \Gamma_{22}^1 \right)_{i,j} f_{i,j+1}^n - f_{i,j-1}^n + \frac{N_{i,j}^n}{\sqrt{E_{i,j}^n + G_{i,j}^n - 1}}
\end{align*}
\]

Where \( n \geq 0; f_{0,j}^0 = f_{i,j}^{n+1} \quad (0 \leq j \leq J + 1); f_{i,0}^0 = f_{i,0}^{n+1} \quad (0 \leq i \leq I + 1); \)

\( f_{i,j}^0 = f_{i,j}^{n+1} \quad (0 \leq j \leq J + 1); f_{i,j+1}^0 = f_{i,j+1}^{n+1} \quad (0 \leq i \leq I + 1); \)

\( f_{0,j}^0, f_{i,0}^0, f_{i+1,j}^0 \) and \( f_{i,j+1}^0 \) are boundary conditions.

If computational domain is normalized to \([0,1] \times [0,1]\), the basic equations (2) could be expressed as,

\[
\begin{align*}
A_1 F_{i,j}^{n+1} = B_1^n \\
A_2 F_{i,j}^{n+1} = B_2^n
\end{align*}
\]

Where \( F_{i,j}^{n+1} = (f_{i,1}^{n+1}, \ldots, f_{i,j}^{n+1}, f_{j,2}^{n+1}, \ldots, f_{j,1}^{n+1}, \ldots, f_{j-1,1}^{n+1}, \ldots, f_{i-1,j}^{n+1}, f_{i-1,j}^{n+1}, \ldots, f_{1,j}^{n+1})^T \);
$I + 2$ and $J + 2$ are lattice number in direction $x$ and direction $y$ respectively; $h = \frac{1}{I + 1} = \frac{1}{J + 1}; A_1^n$ and $B_1^n$ are respectively matrix of left-hand item and vector of right-hand item of the first equation of equation sets (3); $A_2$ and $B_2^n$ are respectively matrix of left-hand item and vector of right-hand item of the second equation of equation sets (3).

2.2. Optimum Formulation

If $Z = \begin{bmatrix} A_1 \\ A_2 \end{bmatrix}$ and $Q^n = \begin{bmatrix} B_1^n \\ B_2^n \end{bmatrix}$, the following equality-constrained least squares problem can be developed to make the interpolated values equal to or approximate to the sampled values at the sampling points,

$$\begin{align*}
\min_{F} & \|ZF^{n+1} - Q^n\|_2 \\
\text{s.t.} & \quad C \times F^{n+1} = D
\end{align*}$$

(4)

Where $C(k, (i - 1) \cdot J + j) = 1$ and $D(k) = \tilde{f}_{i,j}$, which means that the sampled value is $\tilde{f}_{i,j}$ at the $k$th sampling point $(x_i, y_j)$.

For sufficiently large $\lambda$, the algorithm can be transferred into unconstrained least squares approximation,

$$\min_{F} \| \frac{Z}{\lambda C}F^{n+1} - \frac{Q^n}{\lambda D} \|_2$$

(5)

Or

$$[\frac{Z^T}{\lambda C^T}]\begin{bmatrix} \frac{Z}{\lambda C}F^{n+1} \\ \frac{Q^n}{\lambda D} \end{bmatrix} = [\frac{Z^T}{\lambda C^T}]\begin{bmatrix} \frac{Q^n}{\lambda D} \end{bmatrix}$$

(6)

If we denote $S_h = \begin{bmatrix} Z^T \\ \lambda C^T \end{bmatrix}$ and $B_h^n = \begin{bmatrix} Z^T \\ \lambda C^T \end{bmatrix}V^n$, under consideration of simulation step length (or grid cell size) $h$, then, formulation can be expressed as,

$$S_h F^{n+1} = B_h^n$$

(7)

We can get the iteration expression,

$$F^{n+1} = (A_1^T A_1 + A_2^T A_2 + \lambda^2 C^T C)^{-1} (A_1^T B_1^n + A_2^T B_2^n + \lambda^2 C^T D)$$

(8)
If an iteration method is used to solve equation set, the iteration cycle for solving the equation set is named inner iteration and the procedure of updating the right-hand item $B^n_k$ named as outer iteration.

3. Framework implementation

HASM is more theoretically perfect and more complicated than other classical methods of surface modeling. This tool decompose the HASM modeling process into several components including HASM initial field, HASM optimistic controller and HASM iterative solver. HASM can opt remote sensing images or existing short precision grid data for initial field and can also simulate field when lacking of initial fields. HASM optimistic controller need to get knowledge from environmental database to generate controlling condition rulers ensuring modeling more reasonable and more precision. HASM iterative solver is responsible for solution of HASM by means of integration of various solve methods which we have developed. The framework of the software is written in C# and ESRI ArcObjecls for GUI (graphical user interface) and spatial data process. The core HASM computation routines are written in C++. The programs of this project are based on object-oriented design. All spatial data are organized as ArcGIS and GDAL compatible format.

This tool comprises of three main modules, the meaning of each module is explained in Table 1, and the execution flowchart as shown in the Fig. 1.

---

**Fig. 1 the execution flowchart**
Table 1 description of the modules

| Module                  | Description                                                                 |
|-------------------------|-----------------------------------------------------------------------------|
| Sample values           | Load multi-year sample values                                               |
| Optimistic controller   | Apply equality or inequality constraints to make the interpolated values   |
|                         | equal to or approximate to the sampled values at the sampling points       |
| Iterative solver        | Discretization of PDEs. According to the different accuracy and speed       |
|                         | requirements, different iterative solvers are performed to generate the     |
|                         | surface.                                                                    |

The no covariates modeling include the following steps,

1. The first step includes batch loading raw data, spatialization of raw data and fields mapping of data tables in multiple Spatio-Temporal scales. Raw data are often organized as tables or ASCII-files, this program can load them quickly and convert them into spatial format. Through field defining files, this program can also map the fields of data into the operating fields for flexibility and extension of reading tables of different structures.

2. The next step is to generate optimistic formula according to the sample values. User can opt for equality or inequality constraints and upper and lower boundaries control [28] to make the simulated accuracy as high as possible.

3. According to the requirement of simulation accuracy, user can choose first-order or third-order truncation for finite difference. These two modes focus on different domains of model. The first-order truncation is suitable for particularly high speed simulation and relatively high accuracy than classical surface modeling methods. The third-order truncation is apt for extremely high precision model. And advanced user can also choose different iterative solving method to obtain the simulated surface.

The covariate modeling is generally similar to the no covariates modeling, the only difference is software firstly computes the residuals and predicts; generate residuals surface and add it to predict to obtain the final simulated surface.

4. A case study: temperature surfaces in China

Climate data especially temperatures are bases of the study of meteorology, agriculture, forestry and ecology. Temperature surfaces are critical environmental indexes of various geospatial and meteorology models. HASM can produce high accuracy temperature surfaces to help modelers to get more reliable outcomes. Here we take the simulation of temperature surfaces on a national level. We use a regressing transfer function of temperature and various surface modeling methods to create average annual temperature surfaces [28].

Table 2 error between temperatures of different surface modeling methods

| methods | MAE(Mean-Absolute-Error) | MRE(Mean-Relative-Error) | RMSE(Root-Mean-Square-Error) |
|---------|--------------------------|--------------------------|-----------------------------|
| Idw     | 0.56245976               | 0.23592812               | 1.0467782                   |
| Spline  | 0.59179349               | 0.20904898               | 1.0534512                   |
| kriging | 0.74907338               | 0.25565086               | 1.287444                    |
| HASM    | 0.4879699                | 0.18408516               | 0.93618777                  |
5. Conclusion

This paper presents a flexible software tool of high accuracy surface modeling. This software is also a framework to help modelers to perform multiple surface modeling in some extents.

Surface modeling methods have been implemented in many GIS. HASM has been successfully applied to simulating various ecological surfaces on different Spatio-Temporal scales. The framework can facilitate user easily to process data. This software is easily applied to the development of ecosystem modeling integrated with surface modeling. This study from this integration is also helpful for future researches that aim to integrate surface modeling and ecosystem modeling.

References

[1] Aspinall, R., Pearson, D., Integrated geographical assessment of environmental condition in water catchments: Linking landscape ecology, environmental modelling and GIS. Journal of Environmental Management 2000. 59(4) 299-319.
[2] El Yacoubi, S., El Jai, A., Jacewicz, P., Pausas, J.G., LUCAS: an original tool for landscape modelling. Environmental Modelling & Software 2003. 18(5) 429-37.
[3] Gao, Q., Xu, L.D., Liang, N., Dynamic modelling with an integrated ecological knowledge-based system. Knowledge-Based Systems 2001. 14(5-6) 281-7.

[4] He, C.S., Integration of geographic information systems and simulation model for watershed management. Environmental Modelling & Software 2003. 18(8-9) 809-13.

[5] Kragt, M.E., Newham, L.T.H., Bennett, J., Jakeman, A.J., An integrated approach to linking economic valuation and catchment modelling. Environmental Modelling & Software 2011. 26(1) 92-102.

[6] Wanless, S., Bacon, P.J., Harris, M.P., Webb, A.D., Greenstreet, S.P.R., Webb, A., Modelling environmental and energetic effects on feeding performance and distribution of shags (Phalacrocorax aristotelis): integrating telemetry, geographical information systems, and modelling techniques. Ices Journal of Marine Science 1997. 54(4) 524-44.

[7] Ali, K., Cheng, Q.M., Li, W.C., Chen, Y.Q., Multi-element association analysis of stream sediment geochemistry data for predicting gold deposits in south-central Yunnan Province, China. Geochemistry-Exploration Environment Analysis 2006. 6 341-8.

[8] Hjimans, R.J., Cameron, S.E., Parra, J.L., Jones, P.G., Jarvis, A., Very high resolution interpolated climate surfaces for global land areas. International Journal of Climatology 2005. 25(15) 1965-78.

[9] Hillenbrand, C.J., Friedman, G.M., A regional investigation of depth to groundwater in New York State utilizing GIS technology. Northeastern Geology and Environmental Sciences 2005. 27(1) 54-9.

[10] Maas, S., Scheifler, R., Benslama, M., Crini, N., Lucot, E., Brahimia, Z., Benyacoub, S., Giraudoux, P., Spatial distribution of heavy metal concentrations in urban, suburban and agricultural soils in a Mediterranean city of Algeria. Environmental Pollution 2010. 158(6) 2294-301.

[11] Price, D.T., McKenney, D.W., Nalder, I.A., Hutchinson, M.F., Kesteven, J.L., A comparison of two statistical methods for spatial interpolation of Canadian monthly mean climate data. Agricultural and Forest Meteorology 2000. 101(2-3) 81-94.

[12] Werner, M.G.F., Impact of grid size in GIS based flood extent mapping using a 1D flow model. Physics and Chemistry of the Earth Part B-Hydrology Oceans and Atmosphere 2001. 26(7-8) 517-22.

[13] Aguilar, F.J., Aguilar, M.A., Aguera, F., Sanchez, J., The accuracy of grid digital elevation models linearly constructed from scattered sample data. International Journal of Geographical Information Science 2006. 20(2) 169-92.

[14] Chun, W., Guoan, T., Dunxin, J., Yini, J., Accuracy of difference associated with grid digital elevation model. Proceedings of the SPIE - The International Society for Optical Engineering 2007. 67532I (67513 pp.).

[15] Alsamamra, H., Ruiz-Arias, J.A., Pozo-Vazquez, D., Tovar-Pescador, J., A comparative study of ordinary and residual kriging techniques for mapping global solar radiation over southern Spain. Agricultural and Forest Meteorology 2009. 149(8) 1343-57.

[16] Garbrecht, J., Martz, L.W., Digital elevation model grid size, landscape representation, and hydrologic simulations - Comment. Water Resources Research 1996. 32(5) 1461-2.

[17] Goscienski, D., Application of selected statistical coefficients to accuracy analysis on surface interpolation models. 7th International Conference Environmental Engineering. Vols 1-3 2008. 1314-9.

[18] Moreno, J.I.L., Latron, J., Lehmann, A., Effects of sample and grid size on the accuracy and stability of regression-based snow interpolation methods. Hydrological Processes 2010. 24(14) 1914-28.

[19] Fisher, P., Improved Modeling of Elevation Error with Geostatistics. Geoinformatica 1998. 2(3) 215-33.

[20] Grigov, A., Krivoruchko, K., Geostatistical mapping with continuous moving neighborhood. Mathematical Geology 2004. 36(2) 267-81.

[21] Chen, C.F., Yue, T.X., A method of DEM construction and related error analysis. Computers & Geosciences 2010. 36(6) 717-25.

[22] Shi, W.J., Liu, J.Y., Song, Y.J., Du, Z.P., Chen, C.F., Yue, T.X., Surface modelling of soil pH. Geoderma 2009. 150(1-2) 113-9.

[23] Yue, T.X., Du, Z.P., Song, D.J., Gong, Y., A new method of surface modeling and its application to DEM construction. Geomorphology 2007. 91(1-2) 161-72.

[24] Yue, T.X., Fan, Z.M., Chen, C.F., Sun, X.F., Li, B.L., Surface modelling of global terrestrial ecosystems under three climate change scenarios. Ecological Modelling 2011. 222(14) 2342-61.
[25] Yue, T.X., Song, D.J., Du, Z.P., Wang, W., *High-accuracy surface modelling and its application to DEM generation*. *International Journal of Remote Sensing* 2010. 31(8) 2205-26.

[26] O'Neill, B. *Elementary differential geometry*, Amsterdam: Elsevier Academic Press; 2006

[27] Yue, T.X., Song, Y.J., *The YUE-HASM method*. *Proceedings of the 8th International Symposium on Spatial Accuracy Assessment in Natural Resources and Environmental Sciences, Vol II - Accuracy in Geomatics* 2008. 148-53.

[28] Yue, T.X. *Surface Modeling: High Accuracy and High Speed Methods*: Taylor & Francis; 2011