

A review of recent developments in HASM

Tian-Xiang Yue¹ · Li-Li Zhang¹ · Na Zhao¹ · Ming-Wei Zhao¹ · Chuan-Fa Chen² ·
Zheng-Ping Du¹ · Dun-Jiang Song³ · Ze-Meng Fan¹ · Wen-Jiao Shi⁴ ·
Shi-Hai Wang⁵ · Chang-Qing Yan⁶ · Qi-Quan Li⁷ · Xiao-Fang Sun⁸ ·
Hai Yang⁹ · John Wilson^{1,10} · Bing Xu¹¹

Received: 1 December 2014 / Accepted: 1 May 2015 / Published online: 26 May 2015
© Springer-Verlag Berlin Heidelberg 2015

Abstract Ground observation is able to obtain highly accurate data with high temporal resolution at observation points, but these observation points are too sparsely to satisfy the application requirements at regional scale. Satellite remote sensing can frequently supply spatially continuous information on earth surface, which is impossible from ground-based investigations, but remote sensing description is not able to directly obtain process parameters. In fact, in terms of fundamental theorem of surfaces, a surface is uniquely defined by the first fundamental coefficients, about the details of the surface observed when we stay on the surface, and the second fundamental coefficients, the change of the surface observed from outside the surface. A method for high accuracy surface modeling (HASM) has been developed initiatively to find solutions for error problem and slow-speed problem of earth surface modeling since 1986. HASM takes global approximate information (e.g., remote

sensing images or model simulation results) as its driving field and local accurate information (e.g., ground observation data and/or sampling data) as its optimum control constraints. Its output satisfies the iteration stopping criterion which is determined by application requirement for accuracy. This paper reviews problems to be solved in every development stage and applications of HASM.

Keywords Surface modeling method · HASM · Error · Computational speed · Memory requirement · Accuracy

Introduction

Ground observation can obtain high accuracy data at observation points, but observations at fixed positions are confined within some limited dispersal points and not able

✉ Na Zhao
zhaon@lreis.ac.cn

Tian-Xiang Yue
yue@lreis.ac.cn

¹ State Key Laboratory of Resources and Environment Information System, Institute of Geographical Science and Natural Resources Research, Chinese Academy of Sciences, Beijing, China

² Geomatics College, Shandong University of Science and Technology, Qingdao, China

³ Institute of Policy and Management, Chinese Academy of Sciences, Beijing, China

⁴ Institute of Geographical Science and Natural Resources Research, Chinese Academy of Sciences, Beijing, China

⁵ China Aerospace Science and Technology Consultant Co., Ltd, Beijing, China

⁶ Department of Information Engineering, Shandong University of Science and Technology, Qingdao, China

⁷ College of Resources and Environment, Sichuan Agricultural University, Ya'an, China

⁸ School of Geography and Tourism, Qufu Normal University, Qufu, China

⁹ Institute of Water Resources and Hydropower Research, Hohai University, Nanjing, China

¹⁰ Spatial Sciences Institute, University of Southern California, Los Angeles, USA

¹¹ School of Environment, Tsinghua University, Beijing, China

to directly calculate relative parameters at regional scale. Satellite remote sensing can frequently supply surface information of geographical processes and ecological processes, but remote sensing description is not able to directly obtain process parameters. Remote-sensing data can generate information about earth surface that is impossible from ground-based studies. The timing and extent of land-cover change and the relationship between climate and phenology highlight unique information that is available only from satellite and airborne sensors (Chambers et al. 2007). However, maps derived from satellites observations are patchy and can not be used reliably as an independent source of information for earth surface monitoring because of the well-known limitations of satellite retrievals, such as missing data for cloud-covered pixels (Emili et al. 2011).

In fact, earth surface systems are controlled by a combination of global factors and local factors, which can not be understood without accounting for both the local and global components. The system dynamic can not be recovered from the global or local controls alone (Phillips 2002). In terms of fundamental theorem of surfaces, a surface is uniquely defined by the first fundamental coefficients and the second fundamental coefficients. The first fundamental coefficients express the information observed when we stay on the surface, about the details of the surface. The second fundamental coefficients express the change of the surface observed from outside the surface (Somasundaram 2005).

Many studies are based on either satellite observations (e.g., Gupta et al. 2014) or ground observations (e.g., Alkhasawneh et al. 2014; Liu et al. 2013). It has been proven that simulation results would have bigger errors when only satellite measurements or ground measurements are considered. For instances, no matter what kind of support vector machine (SVR) models and what parameter optimization method were used, the relative errors of many points exceeded 20 % when predicting mining subsidence in the water area using original observation data of the subsidence area without water directly in Eastern China (Li et al. 2014). A comprehensive survey in Guyana indicated that estimates only based on remotely sensed data may be inaccurate (Butt et al. 2015).

Combination of satellite measurements with ground measurements can considerably improve simulation accuracy. For instance, utilizing a satellite image as secondary information decreased errors associated with yield monitor data and also allowed better prediction in areas where no reliable yield measurements were available (Dobermann and Ping 2004). Gross primary production (GPP) and net ecosystem exchange (NEE) were simulated by assimilating meteorological data derived from the measurements from existing weather stations, forest volume data derived from a previous investigation, satellite data, flux tower data, and other ancillary data, which rendered the simulation more stable and accurate (Chiesi et al. 2011). Errors of water vapor estimation were reduced up to certain extent using both satellite observations and ground observations over the Walnut Creek region in USA (Srivastava et al. 2014).

A high accuracy and speed method (HASM) for surface modeling has been developed initiatively to efficiently assimilate remote sensing data with ground-based observation data since 1986 so that solutions could be found for error and slow-speed problems which have long troubled earth surface modeling. This paper focuses on HASM, especially how problems, appearing in every development stage, were solved by introducing appropriate theories and methods.

Development of high accuracy surface modeling

The development process of HASM can be divided into four stages (Table 1). *In the first stage from 1986 to 2001*, studies were based on curve-theorem, dealing with a surface as a combination of its profiles. It was learnt that slope and curvature are significant variables of surface analysis (Evans 1980). In fact, a curve is uniquely determined by its slope and its curvature in terms of curve-theorem in the plane (Spivak 1979). Following this consideration, a model for modeling cirques was constructed in terms of curve-theorem (Yue and Ai 1990), which was then developed for change detection of earth surface (Yue et al. 2002).

In the second stage from 2001 to 2007, studies were based on fundamental theorem of surfaces, paying attention

Table 1 The most important features of HASM per stage

Development stage	Theoretic foundation	Research focus
From 1986 to 2001	Curve-theorem	Dealing with a surface as a combination of its profiles
From 2001 to 2007	Fundamental theorem of surfaces	Error problems
From 2008 to 2011	Fundamental theorem of surfaces	Low computational-speed and large memory-requirement problems
Since 2012	Fundamental theorem of surfaces	Improvement of HASM and its parallelization

to error problem (Yue 2011). It is proven that the equation of Earth’s surface can be formulated as (Kerimov 2009), $z = f(x, y)$, where z is an attribute value of the earth’s surface at location (x, y) . For the surface $z = f(x, y)$, an iterative formulation of HASM was developed in terms of the fundamental theorem for surfaces (Yue et al. 2007), which was transformed into a symmetric positive-definite and large sparse linear system.

In the third stage from 2008 to 2011, studies were based on fundamental theorem of surfaces, paying attention to low computational-speed and large memory-requirement problems. HASM has a huge computation cost because it must use an equation set for simulating each lattice of a surface. To speed up the computation of HASM, we developed a multi-grid method of HASM (HASM-MG) (Yue and Song 2008; Yue et al. 2013a), an adaptive method of HASM (HASM-AM) (Yue et al. 2010a), an adjustment computation of HASM (HASM-AC) (Yue and Wang 2010), and a preconditioned conjugate gradient algorithm of HASM (HASM-PCG) (Yue et al. 2010b). Multi-grid method is the fastest numerical method for solving partial differential equations, which is based on two principles that are error smoothing and coarse grid correction. The principle of the adaptive method is that grid cells where the error is large will be marked for refinement, while grid cells with a satisfied accuracy are left unchanged. The adjustment computation permits all observations, regardless of their number or type, to be entered into the adjustment and used simultaneously in the computations by means of least squares. A conjugate gradient algorithm was originally viewed as an acceleration technique for the effective solution of large linear systems by a succession of well-convergent approximations; the preconditioned conjugate gradient algorithm can be developed by introducing a preconditioner to ensure faster convergence of the conjugate gradient method.

In the fourth stage since 2012, Gauss–Codazzi equation set was introduced into HASM (Yue et al. 2013b; Zhao and Yue 2014a). We found that accuracy of HASM was not so satisfied in a few cases of multitudinous numerical tests and real-world tests. In other words, HASM performance was not stable enough for all applications. The reason was that second-order central difference stencil, that we employed for HASM solution, had no element on the diagonal corresponding to $f_{i,j}$. This led to algebraic systems with loss of diagonal dominance.

A combination of a forward difference stencil and a backward difference stencil has produced a symmetric stencil, which has a non-zero coefficient in the diagonal and can thus restore the diagonal dominance of the corresponding matrix in the algebraic systems. This refined

symmetric stencil is of second-order on a uniform grid and can give a solution to the instability when solving the algebraic equations of HASM.

Let $f_{i,j}^{(0)} = \bar{f}_{i,j}$ at the sampled point (x_i, y_j) in the computational domain, $(x_i, y_j) \in \Phi$, and $\Phi = \{(x_i, y_j, \bar{f}_{i,j}) | 0 \leq i \leq I + 1, 0 \leq j \leq J + 1\}$ be the set of sampling points, then the matrix formulation of HASM can be expressed as,

$$[\mathbf{A} \quad \mathbf{B} \quad \mathbf{C} \quad \lambda \mathbf{S}] \mathbf{Z}^{(n+1)} = [\mathbf{A} \quad \mathbf{B} \quad \mathbf{C} \quad \lambda \mathbf{S}] \begin{bmatrix} \mathbf{d}^{(n)} \\ \mathbf{q}^{(n)} \\ \mathbf{p}^{(n)} \\ \lambda \mathbf{k} \end{bmatrix} \tag{1}$$

where $\mathbf{Z}^{(n+1)} = (f_{1,1}^{(n+1)}, \dots, f_{1,J}^{(n+1)}, \dots, f_{I,1}^{(n+1)}, \dots, f_{I,J}^{(n+1)})$; \mathbf{A} , \mathbf{B} , and \mathbf{C} , respectively, represent coefficient matrixes of the first, second, and third equation of HASM master equation set; $\mathbf{d}^{(n)}$, $\mathbf{q}^{(n)}$, and $\mathbf{p}^{(n)}$ are, respectively, the right-hand side vectors of the HASM master equation set; the non-zero element of the sample matrix \mathbf{S} can be expressed as $S_{p,(i-1) \times I + j} = 1$ and the non-zero element of the sample vector $k_p = \bar{f}_{i,j}$; λ is the weight of the sampling points and determines the contribution of the sampling points to the simulated surface.

In terms of fundamental existing theorem for surfaces, if the first and second coefficients satisfy Gauss–Codazzi equations, there exists a surface uniquely determined within a Euclidean displacement (Somasundaram 2005). The Gauss–Codazzi equation can be transformed into

$$(\varphi_{1y} - \phi_{2x} - \varphi_2 P - \phi_1 Q)^2 + (\varphi_{2x} - \phi_{1y} - \varphi_1 Q - \phi_2 P)^2 + (Q_x + P_y + \varphi_1 \varphi_2 - \phi_1 \phi_2)^2 = 0 \tag{2}$$

where E , F , and G are the first fundamental coefficients; L , M , and N represent the second fundamental coefficients;

$$\varphi_1 = \frac{L}{\sqrt{E}}; \varphi_2 = \frac{N}{\sqrt{G}}; P = \frac{\sqrt{E_y}}{\sqrt{G}}; Q = \frac{\sqrt{G_x}}{\sqrt{E}}; \phi_1 = \frac{M}{\sqrt{G}} \text{ and } \phi_2 = \frac{M}{\sqrt{E}}$$

Thus, we can design an iteration stopping criterion of the improved HASM as,

$$(\varphi_{1y} - \phi_{2x} - \varphi_2 P - \phi_1 Q)^2 + (\varphi_{2x} - \phi_{1y} - \varphi_1 Q - \phi_2 P)^2 + (Q_x + P_y + \varphi_1 \varphi_2 - \phi_1 \phi_2)^2 < EI \tag{3}$$

where EI is determined by the requirement of an application for simulation accuracy.

Although HASM performance has been considerably improved because of the introduction of the refined symmetric stencil and the Gauss–Codazzi equations, as a consequence it has caused the low-speed problem once again.

Applications of high accuracy surface modeling

HASM has been successfully applied to constructing digital elevation models (Yue et al. 2007; Chen and Yue 2010; Chen et al. 2013a, b), filling voids in the Shuttle Radar Topography Mission (SRTM) dataset (Yue et al. 2012), simulating climate change (Yue et al. 2013a, b; Zhao and Yue 2014a, b), filling voids on remotely sensed XCO₂ surfaces (Yue et al. 2015a), modeling surface soil properties (Shi et al. 2011) and soil pollution (Shi et al. 2009), and analyzing ecosystem responses to climatic change (Yue et al. 2015b). In all these applications, HASM produced more accurate results than the classical methods.

For instance, HASM was applied to simulate the elevation surface of the Dongzhi tableland in the Loess Plateau region of China. The validation results showed that HASM-AM has the highest accuracy and the fastest computation speed, compared with widely used classic methods (Yue et al. 2010a). The Dongzhi tableland has an area of 2.724 km², which makes for 27.24 million pixels at a spatial resolution of 10 m × 10 m, while the area of the Earth's surface is 510 million km², which makes for 5.1 million pixels at a spatial resolution of 10 km × 10 km. In other words, HASM has the capacity of computing more than five of the Earth's surfaces at a spatial resolution of 10 km × 10 km.

The shuttle radar topography mission (SRTM) datasets, which were derived from the Space Shuttle Endeavour in February 2000, have become a useful source of elevation data and are critical to modern imagery analysis and geospatial intelligence requirements. However, STRM data have variously sized voids, resulting in incomplete data sets. These voids account for 0.15 % of the total data set in China. They cover as much as 30 % of the surface in rugged terrain. HASM was used to fill voids in China (Yue et al. 2012). Verification in nine regions with three different geomorphologic units consisting of hills, plateaus and mountains demonstrated that HASM results always had the highest accuracy compared to all of the classic methods, whether auxiliary data were added or not, whether landform complexity was higher or lower, and whether the void areas were larger or smaller in all the nine regions.

HASM has also been used to simulate climate trends in China since 1961 (Yue et al. 2013a, b). We have found that mean annual temperature (MAT) during the period 1961–2010 exhibited spatial stationarity, while mean annual precipitation (MAP) showed spatial non-stationarity. A statistical transfer function (STF) of MAT was formulated using minimized residuals output from HASM with an ordinary least squares (OLS) linear equation that used latitude and elevation as independent variables, abbreviated as HASM-OLS. The STF of MAP under a Box-Cox

(BC) transformation was derived as a combination of minimized residuals output by HASM with a geographically weight regression (GWR) using latitude, longitude, elevation, and an impact coefficient of aspect and sky view factor as independent variables, abbreviated as HASM-GWR-BC. Cross validation of HASM-OLS and HASM-GWR-BC indicates that mean absolute errors of MAT and MAP are −0.15 °C and 1.52 mm, respectively, which were much than the errors generated with the classical methods. In terms of HASM-OLS and HASM-GWR-BC, MAT shows an increasing trend since the 1960s in China, with an accelerating increasing trend since 1980. Our simulation showed that MAT has increased by 1.44 °C since the 1960s. The warming trends increase from the south to the north in China, with the exception of the Qinghai-Xizang plateau. Specifically, the 2100 °C·d contour line of annual accumulated temperature (AAT) of ≥10 °C has shifted northwestward 255 km in the Heilongjiang province of northeastern China since the 1960s. The MAP in the Qinghai-Xizang plateau and in arid regions also shows a continuously increasing trend. On average, China became wetter from the 1960s to the 1990s, but drier from the 1990s to the 2000s. The Qinghai-Xizang Plateau and northern China have experienced more climatic extremes than southern China since the 1960s.

The Coupled Model Intercomparison Project Phase 5 (CMIP5) datasets have been used for the Fifth Assessment Report of the Intergovernmental Panel on Climate Change (AR5). These simulations of general circulation models (GCMs) have offered unprecedented opportunities to analyze the attributes of climatic projections for the twenty-first century. However, due to the coarse spatial resolution of GCMs (200–500 km), it is difficult to assess climate change impacts at regional and local levels (Grotch and MacCracken 1991; von Storch et al. 1993; Raisanen 2007; Prudhomme and Davies 2009; Harding et al. 2013). Temperature and precipitation data from CMIP5 were down-scaled by means of HASM-OLS and HASM-GWR-BC (Yue 2011). Comparing the original CMIP5 historical simulation and the observed data during the baseline period, T1 (1961–2010) indicated that the mean absolute error of MAT from the original CMIP5 data was 2.04 °C (Yue 2015). Especially the mean absolute errors were, respectively, 4.64, 3.39, 2.58, and 2.64 °C in Tibet plateau, arid area, loess plateau, and Sichuan basin and Yunnan-Guizhou plateau. After the downscaling process, the mean absolute error was decreased by 67.16 % in the whole of China and by 84.48, 75.22, and 68.6 % in Tibet plateau, arid area, and loess plateau, respectively. The mean absolute error of MAP from the original CMIP5 data was 350.52 mm. The higher errors happened in Tibet plateau, loess plateau, and Sichuan basin and Yunnan-Guizhou

Table 2 Comparison of errors from different methods

Methods	MAE (kg·m ⁻²)	MRE (%)
SBA (based on satellite observation information)	1.9	49
Kriging (based on local inventory information)	2.0	50
HASM-SBA (based on both local information and global information)	0.9	23

plateau, of which the mean absolute errors are, respectively, 770.51, 656.57, and 386.03 mm. The downscaling process made the mean absolute errors decrease by 77.43 % in the whole of China and by 97.09, 91.34, and 88.53 % in Tibet plateau, arid area, and loess plateau, respectively.

The scenarios of RCP2.6, RCP4.5, and RCP8.5 of MAT and MAP in the period from 2006 to 2010 on spatial resolutions of 1° × 1° were downscaled to the ones on spatial resolution of 1 km × 1 km by a combination of the method for high accuracy surface modeling (HASM) and the statistical transfer functions (Yue 2015). The results showed that spatial distribution of MAT errors had a strong similarity under the three scenarios. Mean absolute errors of all the three scenarios were about 2.2 °C in the whole land of China in terms of the original CMIP5. The regions, where the mean absolute error was bigger than 2.2 °C, included Tibet plateau, arid area, loess plateau, and Sichuan basin and Yunnan-Guizhou plateau. The mean absolute errors of scenarios of RCP2.6, RCP4.5, and RCP8.5 were about 5 °C in Tibet plateau, 4 °C in arid area, 3 °C in loess plateau, and 3 °C in Sichuan basin and Yunnan-Guizhou plateau, respectively. After the downscaling process, the mean absolute errors of the three scenarios were reduced to about 0.6 °C in the whole land of China, 0.7 °C in Tibet plateau, 0.9 °C in arid area, 1.0 °C in loess plateau, and 0.9 °C in Sichuan basin and Yunnan-Guizhou plateau.

A simulation system for XCO2 surfaces is being developed on the basis of HASM (simply termed as HASM system), which takes satellite remote sensing data as its driving field and ground observation data as its optimum control constraints. The XCO2 surfaces from satellite remote sensing are difficult to be directly used as the driving fields because of a large number of voids caused by clouds, aerosols, and the high surface albedo. It is necessary to conduct void filling for constructing the HASM system. The verification results showed that HASM always has the highest accuracy compared with the classical methods of (Inverse distance weighted) IDW and Ordinary Kriging (OK), whether voids are inner ones or boundary ones, data sets are from Japan’s Greenhouse Gases Observing Satellite(GOSAT) or from The scanning imaging absorption spectrometer for atmospheric cartography(SCIAMACHY) on board the European Space Agency’s ENVISAT-1 satellite, and the void areas are larger or smaller (Yue et al. 2015b).

Table 3 The most important abbreviations

Abbreviation	The complete words
AAT	Annual accumulated temperature
AR5	The Fifth Assessment Report of the Intergovernmental Panel on Climate Change
BC	Box-Cox
CMIP5	The Coupled Model Intercomparison Project Phase 5
GCM	General Circulation Model
GOSAT	Japan’s Greenhouse Gases Observing Satellite
GPP	Gross primary production
GPU	Graphics Processing Unit
GWR	Geographically weight regression
HASM	High accuracy surface modeling
HASM-AC	Adjustment computation of HASM
HASM-AM	Adaptive method of HASM
HASM-MG	Multi-grid method of HASM
HASM-PCG	Preconditioned conjugate gradient algorithm of HASM
HASM-SBA	Forest inventory data (the local information) are combined with SBA by means of HASM
IDW	Inverse distance weighted
LPJ-DGVM	Lund-Potsdam-Jena dynamic global vegetation model
MAE	Mean absolute error
MAP	Mean annual precipitation
MAT	Mean annual temperature
MPI	Message Passing Interface
MRE	Mean relative error
NEE	Net ecosystem exchange
OK	Ordinary Kriging
OLS	Ordinary least squares
RCP	Representative concentration pathway
RK	Regression-kriging using a generalized linear model
SBA	Satellite-data-based approach
SCIAMACHY	The scanning imaging absorption spectrometer for atmospheric cartography
SK	Stratified kriging
STF	Statistical transfer function
SVR	Support vector machine
XCO2	Column-averaged dry air mole fraction of CO2

Turning to pedology, 5374 soil profiles collected during China’s second national soil survey (1979–1994) were taken and used as optimum control constraints in HASM to

evaluate regression relations between soil properties and their environmental factors as the driving fields, and to simulate surface soil properties at a spatial resolution of $1 \text{ km} \times 1 \text{ km}$ on the national level. The validations indicated that the HASM results were at least 17 % more accurate compared with OK method (Li et al. 2010a, b). HASM was once again validated in the middle part of Jiangxi province of China, for which 150 samples were collected in different land-cover types consisting of woodlands, croplands, and grasslands. The performance of HASM in simulating soil properties, such as pH, AN, C, N, K, Al, Ca, Mg, and Zn, was evaluated by comparing the results with three of the widely used methods of OK, stratified kriging (SK), and regression-kriging using a generalized linear model (RK). The evaluation demonstrated that HASM maps of soil properties presented more detailed and accurate maps of spatial patterns (Shi et al. 2011).

HASM has also been used to simulate the spatial distribution of grassland biomass in China by taking sampled biomass data in grasslands as optimal control constraints

and remotely sensed imagery as a driving field. The spatial distribution of forest biomass has also been simulated using the Lund-Potsdam-Jena dynamic global vegetation model (LPJ-DGVM) as the driving field and sampled forest biomass as optimum control constraints (Sun 2012). The simulation results over the last 100 years show that the vegetation biomass of China increased slowly from 1901 to 1953, and then much faster since 1953. Total biomass under the planned development scenario is predicted to be 0.09 Gt higher than that under a “business as usual” scenario in 2030.

We validated the results produced, respectively, by a satellite-data-based approach (SBA) (Piao et al. 2009) and an inventory data-based approach (Fang et al. 2001) with the results from fusing the two kinds of data by means of HASM (Yue 2015). China’s national forest inventory database from 2004 to 2008 includes 160,000 permanent sample plots and 90,000 temporary sample plots scattered over the land surface of China. The cross validation indicated that mean absolute errors (MAEs) of the carbon stock surfaces generated by SBA and OK are, respectively, 1.9

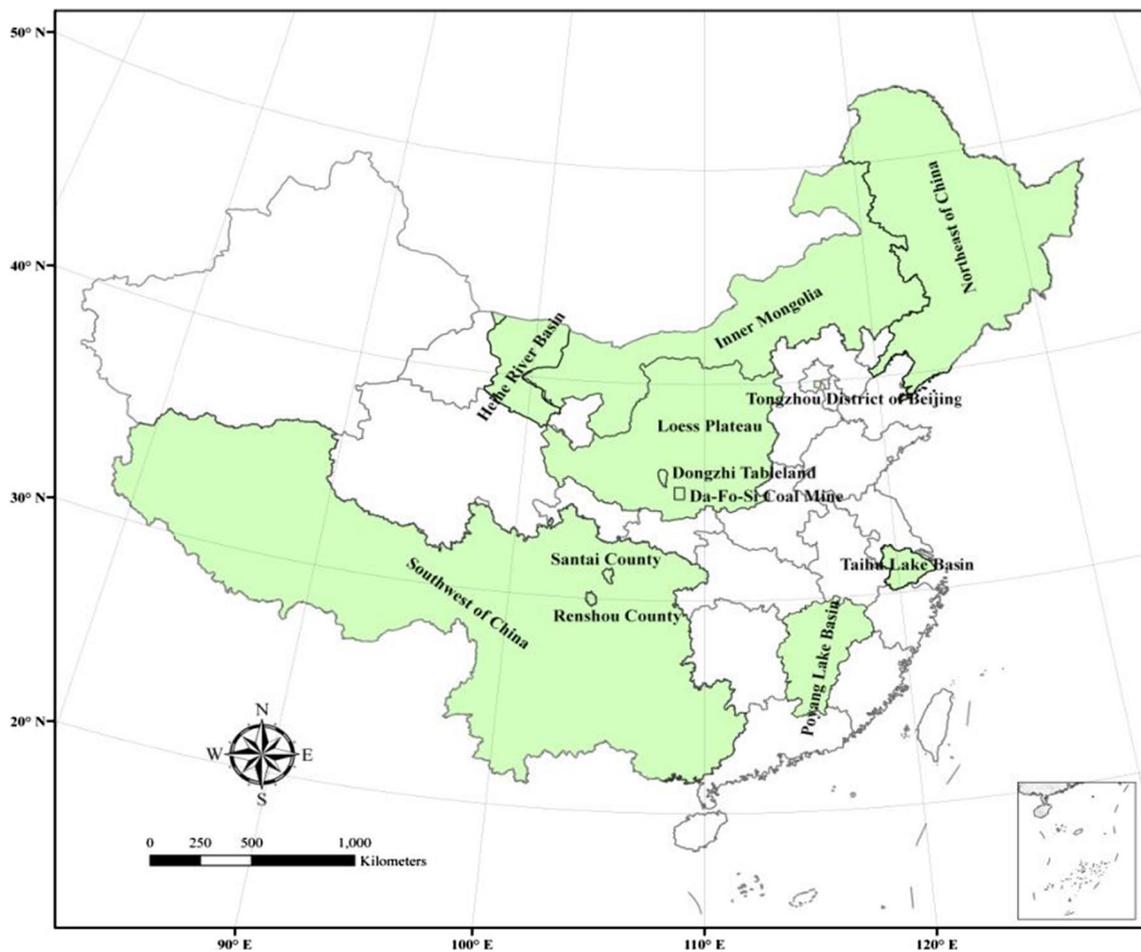


Fig. 1 Study areas of HASM in China

and 2.0 kg m^{-2} . When forest inventory data (the local information) are combined with SBA by means of HASM, which is denoted below as HASM-SBA, MAE is decreased to 0.9 kg m^{-2} . The mean relative errors (MREs) of both the global and local- information-based methods have been reduced by at least 26 % because the local information and global information were fused by means of HASM (Tables 2, 3).

Conclusions

Input of HASM is conducted by taking global approximate information (e.g., remote sensing images or simulation results from global models) as driving fields and taking local accurate information (e.g., ground observation data and/or sampling data) as optimum control constraints. Its output is the results satisfying the iteration stopping criterion which is determined by application requirement for accuracy.

In addition to the whole land of China (Yue et al. 2005, 2006, 2013b), HASM has been applied into different regions on multi-scales (Fig. 1), such as Poyang Lake Basin (Shi et al. 2009, 2011; Yue et al. 2015a), Heihe River Basin (Wang et al. 2015), Dong-Zhi tableland (Yue et al. 2010b; Chen et al. 2012), Taihu Lake Basin (Yue et al. 2013a), Loess Plateau (Fan et al. 2013), Tongzhou District of Beijing (Yue and Wang 2010), Da-Fo-Si coal mine in Shanxi province (Yue et al. 2007). In all these applications, HASM produced the highest accurate results when compared with the classical methods. In the meanwhile, HASM produced the highest accurate results when compared with either global information-based methods or local information-based approaches. However, slow computational-speed and large memory requirement remain the limitation of applications with huge computational work, although we have developed the adaptive method and a multi-grid method of HASM.

To meet the huge computational requirement of big data, one way is to use a faster single processor computer, but continually pursuing the fastest computer can be very expensive and does not scale well as problem size increases. The second way is to develop a fast numerical algorithm. It has been demonstrated that PCG is the most efficient algorithm of HASM that can be transformed into a linear system with a symmetric positive-definite matrix. A third possibility is to break down the computational problem into a number of smaller problems, e.g., using adjustment method of HASM; the smaller computational problems can be solved simultaneously on less expensive computers utilizing parallel computing methods.

Graphics processing units (GPUs) have become a powerful many-core processor. The massively parallel

architecture offers high performance in many computing applications (Galiano et al. 2012). Numerical algorithms can be significantly accelerated if the algorithms map well to the characteristics of the GPU (Helfenstein and Koko 2012). GPU-based parallel algorithm of PCG can considerably improve the efficiency and robustness of HASM (Yan and Yue 2012a, b). GPU implementation of PCG algorithm for HASM is up to 12 times faster comparing with HASM-PCG (Yan et al. 2015). However, it is not faster enough for many applications, especially at high temporal and spatial resolutions on the global level.

Message passing interface (MPI) is the most popular choice in parallel computing environments on clusters of workstations, of which version 1 was released in 1994. MPI represents the standard adopted by most of the industries and researchers (Swann 2002). When a parallel algorithm is implemented in a cluster of workstations using MPI, computational-speed can be greatly improved (Midorikawa et al. 2005). Parallel computing with MPI on clusters of workstations is an effective way to significantly solve the slow computational-speed problem of HASM, especially for huge computation of big data.

Acknowledgments This work is supported by Independent Research Project funded by State Key Laboratory of Resources and Environment Information System, by National Natural Science Foundation of China (91325204), by National High-tech R&D Program of the Ministry of Science and Technology of the People's Republic of China (2013AA122003), by National Key Technologies R&D Program of the Ministry of Science and Technology of the People's Republic of China (2013BAC03B05) and by National Basic Research Priorities Program (2010CB950904) of Ministry of Science and Technology of the People's Republic of China.

References

- Alkhasawneh MS, Ngah UK, Tay LT, Isa NAM (2014) Determination of importance for comprehensive topographic factors on landslide hazard mapping using artificial neural network. *Environ Earth Sci* 72(3):787–799
- Butt N, Epps K, Overman H, Iwamura T, Fragoso JMV (2015) Assessing carbon stocks using indigenous peoples' field measurements in Amazonian Guyana. *Forest Ecol Manag* 338:191–199
- Chambers JQ, Asner GP, Morton DC, Anderson LO, Saatchi SS, Espirito-Santo FDB, Souza MPC Jr (2007) Regional ecosystem structure and function: ecological insights from remote sensing of tropical forests. *Trends Ecol Evol* 22:414–423
- Chen CF, Yue TX (2010) A method of DEM construction and related error analysis. *Comput Geosci* 36:717–725
- Chen CF, Yue TX, Li YY (2012) A high speed method of SMTS. *Comput Geosci* 41:64–71
- Chen CF, Yue TX, Dai HL, Tian MY (2013a) The smoothness of HASM. *Int J Geogr Inf Sci* 27:1651–1667
- Chen CF, Li YY, Yue TX (2013b) Surface modeling of DEMs based on a sequential adjustment method. *Int J Geogr Inf Sci* 27:1272–1291

- Chiesi M, Fibbi L, Genesio L, Gioli B, Magno R, Maselli F, Moriondo M, Vaccari FP (2011) Integration of ground and satellite data to model Mediterranean forest processes. *Int J Appl Earth Obs Geoinf* 13:504–515
- Dobermann A, Ping JL (2004) Geostatistical integration of yield monitor data and remote sensing improves yield maps. *Agron J* 96:285–297
- Emili E, Popp C, Wunderle S, Zebisch M, Petitta M (2011) Mapping particulate matter in alpine regions with satellite and ground-based measurements: an exploratory study for data assimilation. *Atmos Environ* 45:4344–4353
- Evans IS (1980) An integrated system of terrain analysis and slope mapping. *Z Geomorphol Suppl Bd* 36:274–295
- Fan ZM, Li J, Yue TX (2013) Land-cover changes of biome transition zones in Loess Plateau of China. *Ecol Model* 252:129–140
- Fang JY, Chen AP, Peng CH, Zhao SQ, Ci L (2001) Changes in forest biomass carbon storage in China between 1949 and 1998. *Science* 292:2320–2322
- Galiano V, Migallon M, Migallon V, Penades J (2012) GPU-based parallel algorithms for sparse nonlinear systems. *J Parallel Distr Com* 72:1098–1105
- Grotch SL, MacCracken MC (1991) The use of general circulation models to predict regional climate change. *J Clim* 4:286–303
- Gupta M, Mohanty KK, Kumar D, Banerjee R (2014) Monitoring surface elevation changes in Jharia coalfield, India using synthetic aperture radar interferometry. *Environ Earth Sci* 71(6):2875–2883
- Harding KJ, Snyder PK, Liess S (2013) Use of dynamical downscaling to improve the simulation of Central U.S. warm season precipitation in CMIP5 models. *J Geoph Res: Atmos* 118:12522–12536
- Helfenstein R, Koko J (2012) Parallel preconditioned conjugate gradient algorithm on GPU. *J Comput Appl Math* 236:3584–3590
- Kerimov IA (2009) F-approximation of the earth's surface topography. *Izv Phys Solid Earth* 45:719–729
- Li QQ, Yue TX, Fan ZM, Du ZP (2010a) Study on method for spatial simulation of topsoil SOM at national scale in China. *J Nat Resour* 25:1385–1399 (**in Chinese with English abstract**)
- Li QQ, Yue TX, Fan ZM, Du ZP, Chen CF, Lu YM (2010b) Spatial simulation of topsoil TN at the national scale in China. *Geograph Res* 29:1981–1992 (**in Chinese with English abstract**)
- Li L, Wu K, Zhou DW (2014) Extraction algorithm of mining subsidence information on water area based on support vector machine. *Environ Earth Sci* 72(10):3991–4000
- Liu ZP, Shao MA, Wang YQ (2013) Large-scale spatial interpolation of soil pH across the Loess Plateau, China. *Environ Earth Sci* 69(8):2731–2741
- Midorikawa ET, de Oliveira HM, Laine JM (2005) PEMPIs: a new methodology for modeling and prediction of MPI programs performance. *Int J Parallel Prog* 33:499–527
- Phillips JD (2002) Global and local factors in earth surface systems. *Ecol Model* 149:257–272
- Piao SL, Fang JY, Ciais P, Peylin P, Huang Y, Sitch S, Wang T (2009) The carbon balance of terrestrial ecosystems in China. *Nature* 458:1009–1014
- Prudhomme C, Davies H (2009) Assessing uncertainties in climate change impact analyses on the river flow regimes in the UK, Part 1: baseline climate. *Clim Change* 93:177–195
- Raisanen J (2007) How reliable are climate models? *Tellus A* 59:2–29
- Shi WJ, Liu JY, Song YJ, Du ZP, Chen CF, Yue TX (2009) Surface modeling of soil pH. *Geoderma* 150:113–119
- Shi WJ, Liu JY, Du ZP, Stein A, Yue TX (2011) Surface modeling of soil properties based on land use information. *Geoderma* 162:347–357
- Somasundaram D (2005) *Differential geometry*. Alpha Science International Ltd, Harrow
- Spivak M (1979) *A comprehensive introduction to differential geometry*, vol 3. Publish or Peril, Inc., Boston
- Srivastava PK, Han DW, Rico-Ramirez MA, Bray M, Islam T, Gupta M, Dai Q (2014) Estimation of land surface temperature from atmospherically corrected LANDSAT TM image using 6S and NCEP global reanalysis product. *Environ Earth Sci* 72(12):5183–5196
- Sun XF (2012) *Spatial pattern simulation of terrestrial vegetation aboveground biomass*. Dissertation for Doctoral Degree, Graduate University of Chinese Academy of Sciences, Beijing (**in Chinese with English abstract**)
- Swann CA (2002) Maximum likelihood estimation using parallel computing: an introduction to MPI. *Comput Econ* 19:145–178
- vonStorch H, Zorita E, Cubash U (1993) Downscaling of global climate change estimates to regional scales: an application to Iberian rainfall in wintertime. *J Clim* 6:1161–1671
- Wang YF, Yue TX, Du ZP, Zhao MW (2015) Improving the accuracy of the height-diameter equation using the classified factors method. *Environ Earth Sci*. doi:10.1007/s12665-015-4168-2
- Yan CQ, Yue TX (2012a) A novel method for dynamic modeling and real-time rendering based on GPU. *J Geo-Inf Sci* 14:149–157 (**in Chinese with English abstract**)
- Yan CQ, Yue TX (2012b) GPU accelerated high accuracy surface modeling methods for DEM modeling. *Comput Eng Appl* 48:22–27 (**in Chinese with English abstract**)
- Yan CQ, Zhao G, Chen CF, Wang CL, Yue TX (2015) Speeding-up the high accuracy surface modelling method with GPU. *Environ Earth Sci* (**in press**)
- Yue TX (2011) *Surface modelling: high accuracy and high speed methods*. CRC Press, New York
- Yue TX (2015) *Principles and methods of earth surface simulation*. Science Press, Beijing (**in Chinese**)
- Yue TX, Ai NS (1990) The mathematic model of cirque morphology. *J Glaciol Geocryol* 12:227–234 (**in Chinese with English Abstract**)
- Yue TX, Song YJ (2008) The YUE-HASM method. In: Deren Li, Yong Ge and Giles M. Foody (eds) *Accuracy in geomatics*, pp 148–153
- Yue TX, Wang SH (2010) Adjustment computation of HASM: a high-accuracy and high-speed method. *Int J Geograph Inf Sci* 24:1725–1743
- Yue TX, Chen SP, Xu B, Liu QS, Li HG, Liu GH, Ye QH (2002) A curve-theorem based approach for change detection and its application to Yellow River Delta. *Int J Remote Sens* 23:2283–2292
- Yue TX, Fan ZM, Liu JY (2005) Changes of major terrestrial ecosystems in China since 1960. *Global Planet Change* 48(4):287–302
- Yue TX, Fan ZM, Liu JY, Wei BX (2006) Scenarios of major terrestrial ecosystems in China. *Ecol Model* 199(3):363–376
- Yue TX, Du ZP, Song DJ, Gong Y (2007) A new method of surface modelling and its application to DEM construction. *Geomorphology* 91:161–172
- Yue TX, Chen CF, Li BL (2010a) An adaptive method of high accuracy surface modeling and its application to simulating elevation surfaces. *Trans GIS* 14:615–630
- Yue TX, Song DJ, Du ZP, Wang W (2010b) High-accuracy surface modelling and its application to DEM generation. *Int J Remote Sens* 31:2205–2226
- Yue TX, Chen CF, Li BL (2012) A high accuracy method for filling SRTM voids and its verification. *Int J Remote Sens* 33:2815–2830
- Yue TX, Zhao N, Yang H, Song YJ, Du ZP, Fan ZM, Song DJ (2013a) The multi-grid method of high accuracy surface modelling and its validation. *Trans GIS* 17(6):943–952
- Yue TX, Zhao N, Ramsey RD, Wang CL, Fan ZM, Chen CF, Lu YM, Li BL (2013b) Climate change trend in China, with improved accuracy. *Clim Change* 120:137–151

- Yue TX, Du ZP, Lu M, Fan ZM, Wang CL, Tian YZ, Xu B (2015a) Surface modelling of ecosystem responses to climatic change. *Ecol Model* (**in press**)
- Yue TX, Zhao MW, Zhang XY (2015b) A high-accuracy method for filling voids on remotely sensed XCO₂ surfaces and its verification. *J Clean Prod* (**in Chinese**)
- Zhao N, Yue TX (2014a) A modification of HASM for interpolating precipitation in China. *Theor Appl Climatol* 116:273–285
- Zhao N, Yue TX (2014b) Sensitivity studies of a high accuracy surface modelling method. *Sci China-Earth Sci* 57:1–11