Geosci. Model Dev. Discuss., 8, 2653–2689, 2015
www.geosci-model-dev-discuss.net/8/2653/2015/
doi:10.5194/gmdd-8-2653-2015

Geoscientific Model Development Discussions

Open Access

Matching soil grid unit resolutions with polygon unit scales for DNDC modelling of regional SOC pool

H. D. Zhang[1,3], D. S. Yu[1,3], Y. L. Ni[1,3], L. M. Zhang[1,2], and X. Z. Shi[1,3]

[1]State Key Laboratory of Soil and Sustainable Agriculture, Institute of Soil Science, Chinese Academy of Sciences, Nanjing, 210008, China
[2]College of Resource and Environment, Fujian Agriculture and Forestry University, Fuzhou, 350002, China
[3]Graduated University of Chinese Academy of Sciences, Beijing 100393, China

Received: 3 February 2015 – Accepted: 27 February 2015 – Published: 9 March 2015

Correspondence to: D. S. Yu (dshyu@issas.ac.cn)

Published by Copernicus Publications on behalf of the European Geosciences Union.

Matching soil grid unit resolutions with polygon unit scales for DNDC modelling

H. D. Zhang et al.

Abstract

Matching soil grid unit resolution with polygon unit map scale is important to minimize uncertainty of regional soil organic carbon (SOC) pool simulation as their strong influences on the uncertainty. A series of soil grid units at varying cell sizes were derived from soil polygon units at the six map scales of 1 : 50 000 (C5), 1 : 200 000 (D2), 1 : 500 000 (P5), 1 : 1 000 000 (N1), 1 : 4 000 000 (N4) and 1 : 14 000 000 (N14), respectively, in the Tai lake region of China. Both format soil units were used for regional SOC pool simulation with DeNitrification–DeComposition (DNDC) process-based model, which runs span the time period 1982 to 2000 at the six map scales, respectively. Four indices, soil type number (STN) and area (AREA), average SOC density (ASOCD) and total SOC stocks (SOCS) of surface paddy soils simulated with the DNDC, were attributed from all these soil polygon and grid units, respectively. Subjecting to the four index values (IV) from the parent polygon units, the variation of an index value (VIV, %) from the grid units was used to assess its dataset accuracy and redundancy, which reflects uncertainty in the simulation of SOC. Optimal soil grid unit resolutions were generated and suggested for the DNDC simulation of regional SOC pool, matching with soil polygon units map scales, respectively. With the optimal raster resolution the soil grid units dataset can hold the same accuracy as its parent polygon units dataset without any redundancy, when VIV < 1 % of all the four indices was assumed as criteria to the assessment. An quadratic curve regression model $y = -8.0 \times 10^{-6} x^2 + 0.228x + 0.211$ ($R^2 = 0.9994$, $p < 0.05$) was revealed, which describes the relationship between optimal soil grid unit resolution (y, km) and soil polygon unit map scale ($1 : x$). The knowledge may serve for grid partitioning of regions focused on the investigation and simulation of SOC pool dynamics at certain map scale.

GMDD
8, 2653–2689, 2015

Matching soil grid unit resolutions with polygon unit scales for DNDC modelling

H. D. Zhang et al.

1 Introduction

Soil organic carbon (SOC) is the largest terrestrial carbon pool (Schlesinger, 1997), with stocks about four times the biotic (trees, etc.) pool and about three times the atmospheric pool (Lal, 2004). Relatively modest changes in SOC storage can result in a significant alteration in the atmospheric CO_2 concentration (Davidson and Janssens, 2006). Therefore, an accurate SOC pool estimation has become an important requirement for assessing the global carbon balance and for global climate change.

Agricultural soils are a highly sensitive part of the global carbon cycle (Shi et al., 2010; Wang et al., 2011), carbon sequestration by agricultural soils presents an immediate viable option for increasing soil carbon pool and reducing atmospheric CO_2 and mitigating global warming (Sun et al., 2010). For complexities of human activities and tillage practices affecting agricultural soil, SOC dynamic changes are increasingly to be simulated over broad space and time scales by process-based models (Giltrap et al., 2010; Xu et al., 2012a), such as DeNitrification–DeComposition (DNDC) (Li et al., 2003).

The DNDC model developed by Li et al. (1992a, b) can simulate C and N biogeochemical cycles occurring in agricultural systems, driven by both the environmental factors (e.g. soil organic matter, texture, pH, bulk density, hydraulic properties, daily temperatures and precipitation, etc.) and management practices (e.g. crops, tillage, fertilization, manure application, grazing, etc.). It has been validated through long-term applications internationally at the plot scale, including many sites of North America, Europe, Asia, etc. (Pathak et al., 2005; Li et al., 2006; Tonitto et al., 2007), and is one of the most widely accepted biogeochemical models in the world (Li, 2007; Tang et al., 2006; Li et al., 2010).

The DNDC model has also been utilized to upscale estimates of SOC from plot to region scale. At the region scale the DNDC modelling conducted initially has used counties as basic simulation units, where minimum and maximum soil parameter values for each county were derived from soil maps to simulate an upper and a lower estimate of

GMDD
8, 2653–2689, 2015

several C and N pools (Cai et al., 2003; Li et al., 2004). However, county scale model simulations are subject to great uncertainties as soil properties are averaged for each county, largely ignoring the nonlinear impacts of soil heterogeneity therein (Rüth and Lennartz, 2008; L. M. Zhang et al., 2014).

Recently for DNDC up-scaled utilization, a region is partitioned into many simulation units, within which all soil properties are assumed to be as homogeneous as they are at the plot scale (Li et al., 2005; Zhang et al., 2012). The homogeneity assumption is a possible major source of error when extending DNDC modelling from the plot to the region scale (Li et al., 2002, 2004). As the area of the basic simulation unit increases so does soil property variability or heterogeneity, calling into question the accuracy of its capture (Smith and Dobbie, 2001; Bouwman et al., 2002).

Soil polygons derived from soil vector maps are used as basic simulation units, that is one way to reduce effects of soil heterogeneity on DNDC modelling as they can as possible (Xu et al., 2012b; Yu et al., 2013; Zhang et al., 2012). Even so, the soil heterogeneity within a soil polygon unit still exists, and depends on the soil vector map scale, smaller map scale resulting higher heterogeneity (Yu et al., 2013). To different broad regions, multi-scales of the polygon unit simulated with DNDC ranged widely from 1 : 50 000 to 1 : 14 000 000, taken effects extremely on accuracy and uncertainly of the modelling (Xu et al., 2011, 2012b; Yu et al., 2013; L. M. Zhang et al., 2014).

Another way to reduce effects of the soil heterogeneity on the DNDC modelling is that soil grid cells are used as the basic simulation units (Huang et al., 2004; Y. Q. Yu et al., 2007; Shi et al., 2010; Yu et al., 2011). Cell size or resolution of the soil grid units is one of rulers to scale the soil heterogeneity therein, lower resolution or larger cell size resulting higher soil heterogeneity likewise. The cell size or resolution take effects extremely also on the accuracy and uncertainly of the soil grid unit simulation with DNDC (Yu et al., 2011).

The soil grid units are more often applied to simulation of SOC pool (Qiu et al., 2005; Tang et al., 2006; Yu et al., 2011; Liu et al., 2011), as they are more easily manipulated for spatial model simulation, geo-statistics and spatial analysis than the soil polygon

GMDD
8, 2653–2689, 2015

units (Huang et al., 2004; Li et al., 2005). They are often derived by data conversion from the soil polygon units, but the grid resolution choice varies by researcher even if the soil polygon units are at same map scale and in same region (Y. Q. Yu et al., 2007; Shi et al., 2010; Yu et al., 2012). For example, the soil polygon units compiled in the Soil Database of China (Yu et al., 2007a) at the map scale of 1 : 1 000 000 have been converted to the gird units at the resolutions of 1 km × 1 km (Yu et al., 2007b) and 10 km × 10 km (Y. Q. Yu et al., 2007, 2012) to simulate and estimate agricultural SOC pools in China, respectively. The soil grid units at the resolution of 2 km × 2 km (Shen et al., 2003) and 50 km × 50 km (Wan et al., 2011) converted from the original soil polygon units at the map scale of 1 : 4 000 000, were used for the grid simulation of SOC dynamics in different regions, respectively. The original soil polygon units at the map scale of 1 : 50 000 were converted to grid units at the resolution of 100 m × 100 m (Shi et al., 2010) and 30 m × 30 m (Su et al., 2012) for the grid simulation of SOC dynamics in agro-ecosystem, respectively.

Our concerning is whether these soil grid units at different cell sizes are equivalent in accuracy or granularity to their parent soil polygon units at a corresponding map scale for DNDC modelling. In other words, whether these soil grid unit datasets regulate coarser data or contain redundant data of soil properties, contrasting to their parent soil polygon unit dataset at a certain map scale. The coarser or redundant dataset affects the simulation unit inner homogeneity of soil properties, and farther affects the common outcome too, being that modelling error will be lower if all features within the simulation unit are more homogeneous (Cai et al., 2003; Yu et al., 2011, 2013).

In fact the accuracy and the redundancy are two important issues to soil simulation units' dataset conversion from polygon to grid format, which are often neglected in modelling at regional scale. The accuracy of the grid unit dataset determine reliability and uncertainty of SOC grid simulation (Batjes, 2000; Ni, 2001), the redundancy of the dataset results in mistaken understanding of data accuracy and redundant workload and cost of the simulation (Yu et al., 2011, 2013). Some researches focus on data accuracy but neglect the data redundancy (Yu et al., 2007b; Shi et al., 2010), while

GMDD
8, 2653–2689, 2015

others neglect the data accuracy (Batjes, 2000; Y. Q. Yu et al., 2007) when conduct data conversion, they always search for an individual solution in every case.

Given the variety of datasets and number of simulations, in combination with data accuracy and redundancy as well as computational costs (Schmidt et al., 2008), important questions are raised. How sensitive is DNDC modelling to different simulation units at varied vector map scales or raster grid resolutions? Which raster resolution is optimal to DNDC grid simulation at a fixed soil map scale for error and cost controls? Matching the soil grid unit resolution with polygon unit map scale is one of essential issues to DNDC modelling.

In the present study, paddy soil polygon simulation units at six vector map scales from 1 : 50 000 to 1 : 14 000 000 were converted to grid simulation units at varied raster resolution, respectively, in the Tai Lake region of China. Soil organic carbon pools were simulated by polygon simulations and grid simulations with the DNDC model at the varied vector map scales and raster resolutions, respectively.

The objectives of the study were to (1) reveal the impact of vector map scale and raster resolution of soil simulation units on the DNDC modelling, (2) determine an optimal raster resolution of grid simulation units at a fixed soil vector map scales, based on an assessment of the simulation units' data accuracy and redundancy metrics, and (3) construct relationship between soil vector map scale of polygon units and optimal raster resolution of grid units for DNDC modelling at regional scale. The results will serve as a reference for soil simulation unit conversion from polygon to grid format, in the support of soil carbon cycle modelling at regional scale.

2 Materials and methods

2.1 Study area

The Tai Lake region (118°50′–121°54′ E, 29°56′–32°16′ N) (Fig. 1) is located in the middle and lower reaches of the Yangtze River in China, covers a watershed area of

36 500 km^2, including parts of Jiangsu and Zhejiang provinces and the entire Shanghai City administrative area. The terrain of the region is generally flat plains, broken by a high density of rivers. A northern subtropical monsoon climate prevailed in the area with mean annual temperature and precipitation of 16 °C and 1100–1400 mm, respectively (Xu et al., 1980). The soil types in the region are mainly Paddy, Fluvo-aquic and Red soils, which covers 90 % of total area. Paddy soils, the largest single proportion of any soil type in the Tai Lake region, occupy 23 200 km^2, approximately 66 % of total area (Yu et al., 2014). Derived from loess, alluvium and lacustrine deposit, Paddy soils in the Tai Lake region are recognized as the most typical of their type in China (Yu et al., 2013), with a long history of rice cultivation spanning over several centuries. A summer rice (planted in June and harvested in October) and winter wheat (planted in November and harvested in May) doublecrop rotation has been intensively cultivated in this region (L. M. Zhang et al., 2012, 2014). Six subgroups, Bleached, Gleyed, Percogenic, Degleyed, Submergenic and Hydromorphic are included in the Paddy soils. They are cross referenced in US Soil Taxonomy (ST) as Typic Epiaquepts (Bleached, Percogenic, Hydromorphic) and Typic Endoaquepts (Gleyed, Degleyed, Submergenic) (Shi et al., 2006; Soil Survey Staff, 1994).

2.2 Development of polygon and grid simulation unit datasets at different map scales

First of all, paddy polygon unit datasets for DNDC simulation at six soil vector map scales, e.g. 1 : 50 000 (C5), 1 : 200 000 (D2), 1 : 500 000 (P5), 1 : 1 000 000 (N1), 1 : 4 000 000 (N4) and 1 : 14 000 000 (N14), were developed in the Tai Lake region, respectively. They were generated respectively by vector overlay from paddy polygons at the six map scale datasets and polygons depicting county boundaries at a scale of 1 : 50 000 using the Union function supported by the ESRI ARCGIS 9.0 software (ESRI, Redlands, CA). All simulation units of paddy polygons at one certain map scale within one county have same feature input value for DNDC modelling such as crops, agricultural management and climate, except soil feature, such as soil types, soil organic matter content, clay content, bulk density, rock fragments content, soil layer thickness, pH, hydraulic properties, etc (Yu et al., 2013).

The paddy polygon unit datasets at the six map scales were developed by a Gis Linkage technique based on Soil Type (Yu et al., 2005, 2007a, b), namely PKB (Pedological Knowledge Based) method (Zhao et al., 2006), from soil vector maps at their corresponding map scales, respectively. The soil vector maps were compiled using a standard soil mapping system formulated as part of the Second National Soil Survey of China conducted in the 1980s (Office for the Second National Soil Survey of China, 1994). To the six map soils, soil species is the basic mapping unit for C5 and D2, soil family is for P5 and N1, while soil subgroup is for N4 and N14 (Yu et al., 2014). The soil properties attributed to all paddy polygons were derived from soil profiles, which were surveyed, compiled and authorized in the Second Soil Survey of China in 1980s (Shi et al., 2006). The number of representative soil profiles whose measured data were applied to attribute paddy polygons at C5, D2 and P5 scales totaled 1107, 136 and 127, respectively. The datasets were all taken from three books: Soils of County, Soils of District and Soils of Province, respectively. The paddy polygons at national map scale (N1, N4 and N14) were origined from 49 soil profiles described from the book "Soils of China" (Shi et al., 2006; Yu et al., 2014).

Secondly, paddy grid unit datasets for DNDC simulation were developed from above paddy polygon unit datasets at the six map scales. Each vector paddy polygon unit dataset was converted to a series of paddy grid unit datasets of differing grid cell sizes. The gird cell size ranged from a default size to a maximum, with the size increment set to approximately 10 % of the default. The default was determined by the soil vector map scale and the lowest mapping unit size (2 mm × 2 mm), which can be described and exhibited in hard copy of the map (Yu et al., 2014). For conversions of the six paddy polygon unit datasets (C5, D2, P5, N1, N4 and N14), the default grid cell sizes are 100, 400 m, 1, 2, 8 and 28 km, respectively. In addition, the paddy polygon unit dataset at N14 scale was also converted to gird unit datasets at cell sizes ranging from the default

GMDD
8, 2653–2689, 2015

to a minimum size, with the approximate decrement of 10 % of the default cell size. The minimum and maximum grid cell sizes were that at which the difference of the paddy soil SOC pool simulated by DNDC with the grid unit dataset exceeds the simulation from its parent polygon unit dataset by 30 %. All the data conversions were conducted using the Polygons to Raster Conversion Tools (PRCT), a component of the ArcGIS 9.0 software, with the grid cell value assignment type option of Maximum-Area.

Finally, all simulation units rendered as vector (polygon unit) and raster (gird unit) datasets describing the soil properties, daily weather, cropping systems, and agricultural management practices of rice paddy fields, are required to initialize and run the DNDC model at regional scale (Yu et al., 2011, 2013). Each simulation unit has own data records specifically that were used as input for the DNDC modelling of SOC dynamics (L. Zhang et al. 2009; L. M. Zhang et al., 2009, 2012, 2014).

2.3 DNDC modelling and validation

The DNDC (DeNitrification–DeComposition) model is a process-base model of carbon (C) and nitrogen (N) biogeochemistry in agroecosystems (Li et al., 1992a, b), it can simulate soil C and N biogeochemical cycles in paddy rice ecosystems, depending on a series of anaerobic processes being supplemented in the model (Li et al., 2002, 2004; Li, 2007).

For DNDC modelling of SOC dynamics, farming management scenarios were compiled based on five assumptions from L. Zhang et al. (2009) and L. M. Zhang et al. (2009, 2012, 2014), did not vary with the soil simulation unit within counties. The DNDC modelling runs span the time period 1982 to 2000, duration of 19 years. A total of 65 340 paddy polygon unit simulations were executed, as well as half million paddy grid unit simulations roughly. At present study DNDC in 9.1 versions was run.

To validate and assess performance of DNDC modelling, observed values of SOC content acquired in 2000 from 1033 soil sampling sites within paddy polygon units at C5 map scale, were used to against modelling values (L. M. Zhang et al., 2014). The observed SOC content of top layer (0–15 cm) varied from 1.9 to 36 g kg^{-1}, and the simulated SOC ranged from 5.1 to 34 $g\,kg^{-1}$ in 2000, where 99.6 % of simulated polygon units in C5 were within the ranges produced by the observed values. Four statistical criteria, the correlation coefficient (r), the relative error (E), the mean absolute error (MAE) and the root mean square error (RMSE), were employed to evaluate the model performance. The r of 0.5 at significant level $p < 0.01$, the E of 6.4 %, MAE of 4.0 $g\,kg^{-1}$ and RMSE of 5.0 $g\,kg^{-1}$, all indicated that the modeled results were encouragingly consistent with the observations, the DNDC model were acceptable for SOC modelling of paddy soils in the Tai Lake region (L. M. Zhang et al., 2014). For a more complete discussion of DNDC model validation and error assessment associated for the region can see L. M. Zhang et al. (2012, 2014).

2.4 Data calculation and analysis

Simulated SOC density (SOCD, $kg\,C\,m^{-2}$) of a paddy polygon or grid unit is calculated according to the following equation (Yu et al., 2014):

$$\mathrm{SOCD} = \sum_{i=1}^{n} \frac{(1-\delta_i\%) \times \rho_i \times C_i \times T_i}{100} \tag{1}$$

Where n is the number of soil pedogenic layers, $\delta_i\%$ represents the volumetric percentage of the fraction > 2 mm (rock fragments), ρ_i is the bulk density ($g\,cm^{-3}$), C_i is the simulated soil organic C content ($g\,kg^{-1}$) in 2000, and T_i represents the thickness (cm) of the layer i. The simulated SOCD of surface paddy soil is calculated to the depth of 20 cm.

Four indices of surface paddy soil, Paddy soil area (AREA, M ha), number of paddy soil type (STN), the simulated SOC stocks (SOCS, Tg) and average SOCD (ASOCD, $kg\,C\,m^{-2}$), were selected to assess data accuracy between a paddy grid unit dataset and its parent polygon unit dataset. The four index values (IVs) determined from each polygon unit dataset are recognized as a benchmark for comparison with those values from their affiliated grid unit datasets. Except for the index value (IV) of STN obtained by accounting, the IV of AREA, SOCS and ASOCD were calculated as follows respectively:

$$\mathrm{IV(AREA)} = \sum \mathrm{AREA}_j. \quad (2)$$

$$\mathrm{IV(SOCS)} = \sum \mathrm{SOCD}_j \times \mathrm{AREA}_j. \quad (3)$$

$$\mathrm{IV(ASOCD)} = \mathrm{IV(SOCS)}/\mathrm{IV(AREA)}. \quad (4)$$

Where AREA_j is the area of the paddy polygon or grid unit; SOCD_j is simulated SOC density of a paddy polygon or grid unit; j is the number of paddy polygon or grid unit (Yu et al., 2014).

Variation of an index value (VIV, %) obtained from a grid unit dataset ($\mathrm{IV\text{-}}_{\mathrm{raster}}$) and its parent polygon unit dataset ($\mathrm{IV\text{-}}_{\mathrm{vector}}$) is recognized as a ruler to scale the magnitude of the consistency between the two datasets. The two format datasets accuracy may be consistent or identical, only if absolute values of all these indices VIVs are less than 1 % (Yu et al., 2014). The VIV is calculated as follow:

$$\mathrm{VIV(\%)} = \mathrm{ABS}(100 \times (\mathrm{IV\text{-}}_{\mathrm{vector}} - \mathrm{IV\text{-}}_{\mathrm{raster}})/\mathrm{IV\text{-}}_{\mathrm{vector}}). \quad (5)$$

Where ABS is absolute function, $\mathrm{IV\text{-}}_{\mathrm{vector}}$ is an index value obtained from a polygon unit dataset; $\mathrm{IV\text{-}}_{\mathrm{raster}}$ is the index value obtained from it's an affiliated grid unit dataset.

The optimal soil grid unit size for a polygon unit dataset conversion to grid unit dataset is the maximum grid cell size of which the two datasets are scaled identically. Statistical analyses were conducted by using the Excel and Origin 10 software.

3 Results

3.1 Variations of input soil parameters among simulation unit datasets

Soil organic matter, clay content, pH and soil bulk density are all sensitive parameters as input for DNDC SOC simulations (Li et al., 2002; Levy et al., 2007; L. M. Zhang et al., 2012, 2014). The spatial distribution characteristics of these soil properties depicted by various simulation unit datasets differ from each other. The difference of the input parameter value affects uncertainty of the modelling (Valade et al., 2014; Zhu and Zhuang, 2014). A map scale or raster resolution decrease yielded a change in their estimated content (Tables 1–6), and a corresponding change in the simulated SOC (Table 7).

Weather data (precipitation, maximum and minimum air temperature) and farming management scenarios (sowing method, nitrogen fertilizer application rates, livestock, planting and harvest dates, etc.) variability among these simulation unit datasets for the purposes of this analysis can be neglected, because they were from the same weather and farming management county scale database (Yu et al., 2011, 2013) overlain with these soil polygon datasets. Change in soil type and their attributes as well as soil type area are the main source of SOC variability simulated by DNDC associated with the simulation unit scale and resolution (Yu et al., 2011, 2013).

3.2 Index values determined from simulation polygon units at different map scales

The basic mapping unit's type, numbers of paddy soil type (STN) and polygon unit (SPN) as well as soil area (AREA) determined from the six paddy polygon unit datasets at different map scales, which describe the physical characteristics of these soil datasets, differ from each other (Table 7). For instance, four of the six paddy soil subgroups, Bleached, Percogenic, Degleyed and Submergenic paddy soil, do not get described in N14 polygon unit dataset but in other five datasets. The data scarcity should be one of the substantial causes of the uncertainties in modelling on regional scales (W. Zhang et al., 2014) did. And understandably, the C5 paddy polygon unit dataset containing the maximum numbers of soil polygon units, soil families and species (Table 7), is the most detailed and accurate database in the Tai Lake region (L. M. Zhang et al., 2009, 2012; Yu et al., 2011, 2013). That the IVs of STN, AREA, SOCS and

ASOCD obtained in C5 dataset are considered to be the most believable in the region (Yu et al., 2011, 2013).

The IVs of SOCS and ASOCD for surface paddy soils simulated by DNDC with the six polygon unit datasets display pronounced difference from each other, as well (Table 7). In the main, the IV of SOCS increased with decreasing of the map scale of polygon unit dataset. The highest IV of SOCS was simulated with the N14 polygon unit dataset, due to the largest area of the Hydromorphic paddy soils with the highest SOCD ($> 8\,\mathrm{kg\,C\,m^{-2}}$) simulated mapped in the dataset. The area of Hydromorphic paddy soils mapped in the N14 dataset with the darkest polygons of SOCD simulated by DNDC (Fig. 2f) is 7 times of that in C5 dataset (Fig. 2a) roughly. Spatial distribution maps of SOCD simulated with these polygon unit datasets display differences from each other, too (Fig. 2). Being synthesized much cursorily, the SOCD maps of surface paddy soil simulated by DNDC with the N4 and N14 polygon unit dataset differ distinctly from the others. Obviously, the map scale of soil polygon unit dataset would significantly influence the results of regional SOC pool simulation (Zhao et al., 2006; Xu et al., 2011, 2012b; L. M. Zhang et al., 2014).

3.3 Optimal soil grid unit resolutions for SOC modelling at regional map scales

The three paddy polygon unit datasets C5, D2, P5 are representative of regional scale digital maps, describing soil features at the county, district and province levels, respectively (Yu et al., 2013). The VIVs of the four assessment indices (STN, AREA, SOCS and ASCOD) determined from grid unit datasets and their parent polygon unit dataset, increases with increasing grid cell size (Fig. 3a–c). VIV magnitude and trend vary with grid cell size and by dataset and index. For instance, the VIV of STN from C5 or D2 datasets varies with grid cell size best described by an exponential curve (Yu et al., 2011), while the VIV from P5 varies as a logarithmic curve (Yu et al., 2014).

To the C5 polygon unit dataset and affiliated grid unit datasets, VIVs of the four indices are all $> 1\,\%$ when the grid cell size set as > 0.5 km. And only the VIV of ASOCD is $< 1\,\%$ when the grid unit resolution ranges from 0.3 to 0.5 km. With the grid cell size

GMDD
8, 2653–2689, 2015

decreasing to 0.3 km, three of four VIVs are all < 1 % except the SOCS index. Only when the grid cell size is ≤ 0.2 km (Fig. 3a) and STN index depicted with soil species (Table 7), the VIVs of the four indices are all < 1 %. That the 0.2 km × 0.2 km resolution is optimal for C5 dataset conversion from polygon to grid unit, as it is at this cell dimension that the grid and parent polygon unit datasets are roughly equivalent in their information content, and the data redundancy is at a minimum (Fig. 3a) when simulating regional SOC pool with DNDC.

Similarly, for D2 and P5 dataset conversion, only the VIV of ASOCD is < 1 % when the raster unit resolution is > 1 and 2 km, respectively. But when the grid cell size for D2 conversion decreases to the range of 0.8–1 km, all of the index VIVs are < 1 % except the STN index of soil species, and all VIVs > 1 % for P5 conversion when grid cells size increase over 2 km and the STN index depicted with soil family (Table 7). VIVs of the four indices derived from D2 and P5 dataset conversions are all < 1 % only when their grid cell sizes are ≤ 0.7 and ≤ 1 km, respectively. It is at those cell dimensions that the grid and parent polygon unit datasets are nearly identical and the cell size is maximized, which minimizes the time and cost of simulation process (Fig. 3b and c). The optimal grid unit resolution for D2 and P5 conversion of simulating regional SOC pool with DNDC is 0.7 km × 0.7 km and 1 km × 1 km, respectively.

3.4 Optimal soil grid unit resolutions for SOC modelling at national map scales

The three paddy polygon unit datasets of N1, N4 and N14, describe soil features at the national scale (Yu et al., 2013). Generally, almost all VIVs of the four assessment indices from these grid unit datasets and their parent polygon unit datasets increase with increasing grid cell size except N14 (Fig. 3d–f).

For example, the VIVs of three index (SOCS, AREA and ASOCD) from the N14 dataset conversion varies with grid cell size in the diagram of random scatter except the STN index of soil subgroup when the grid cell size ranges from 18 to 36 km, which is around the center of its default grid cell size (28 km). The VIV random scatter diagram complicates the selection of an optimal grid unit resolution as the VIV values for the four indices are not consistent with grid cell size variation. To simulate regional SOC pool with DNDC, the optimal grid resolution for N14 dataset conversion was determined to be 17 km × 17 km, as all VIVs are < 1 % when the grid cell size is ≤ 17 km (Fig. 3f).

The results for N1 and N4 datasets conversion demonstrate that the VIVs of ASOCD and STN are < 1 % and the VIVs of SOCS and AREA are > 1 %, when the grid cell size is > 2 and > 8 km, and the STN index depicted with soil family and subgroup (Table 7), respectively. The VIVs of the four indices obtained from their grid unit datasets meet the criteria of < 1 %, only when the grid cell size ≤ 2 and ≤ 8 km, respectively. Accordingly, the grid resolution of 2 km × 2 km for N1 and 8 km × 8 km for N4 dataset conversion is optimal from paddy polygon to grid units, which as simulation units for DNDC modelling of regional SOC pool (Fig. 3d and e).

3.5 Relationship between polygon unit map scale and matched optimal grid unit resolution for the simulation of regional SOC pool

Correlation analysis indicated a statistically significant relationship between paddy polygon unit map scale (1 : x) and matched optimal grid unit resolution (y, km),which can be described as follows:

$$y = -8.0 \times 10^{-6} x^2 + 0.228x + 0.211 \; (R^2 = 0.9994,\; p < 0.05). \tag{6}$$

The quadratic curve regression deviates from a standard linear regression, which describes the relationship between soil polygon unit map scales and their default grid cell sizes. The quadratic model implies that when the map scale for the regional SOC simulation with DNDC is less than 1 : 4 000 000, the optimal grid cell size is less than the default, and the deviation increases with map scale decreasing (Fig. 4).

4 Discussion

4.1 Comparison of simulation grid unit resolutions at different map scale among referenced researches

At map scale of C5 (1 : 50 000), the original soil polygon units were converted to grid cells at size of 100 m × 100 m (Yang et al., 2009; Shi et al., 2010) and 30 m × 30 m (Su et al., 2012) as basic assessment units to simulate the SOC dynamics of agro-ecosystem. Compared to the default optimal resolution (200 m × 200 m), the soil grid units are redundant by the standards suggested here. Similarly, both the grid unit datasets at a cell size of 1 km × 1 km converted from the 1 : 1 000 000 (N1) scale soil polygon unit dataset (Yu et al., 2007b) and 2 km × 2 km converted from the 1 : 4 000 000 scale (N4) dataset (Shen et al., 2003), contain a lot of redundancy, compared to the optimal resolution 2 km × 2 km and 8 km × 8 km achieved in this study, respectively. Although grid unit datasets used by these researchers kept the same data content as their parent polygon unit dataset, the grid cell size is not real resolution matching with their map scales due to the data redundancy. Workload and cost of the regional SOC investigation and simulation tripled due to the increased number of grid cells, if the grid cell was designed as soil sampling and simulating unit.

By contrast, the grid units at the cell size of 50 km × 50 km converted from 1 : 4 000 000 (N4) (Wan et al., 2011) scale's and 10 km × 10 km from the 1 : 1 000 000 scale's (N1) (Y. Q. Yu et al., 2007) soil polygon unit dataset, were used as assessment unit for modelling of SOC dynamics in different regions, respectively. It can be anticipated that the simulated results will have higher uncertainty than its parent polygon units' simulations, because the grid unit datasets is coarser than their parent polygon unit datasets. If such the grid cell is designed as soil sampling and simulating unit, the regional SOC investigation and simulation will not be matching in accuracy to the map scale.

The harmonized world soil database (HWSD), completed by FAO/IIASA/ISRIC/ISSCAS/JRC in 2009, was produced at a cell size of about 1 km × 1 km from an original polygon unit dataset, which contains over 16 000 different soil mapping units and was derived from the Soil Map of the World (1 : 5 000 000), regional Soils and Terrain Digital Database (SOTER) (1 : 1 000 000 to 1 : 5 000 000) as well as the European Soil Map and the Soil Map of China (1 : 1 000 000, digitized and compiled by authors D. S. Yu et al.) (FAO/IIASA/ISRIC/ISSCAS/JRC, 2009). Based on the relationship developed here and assuming a scale of 1 : 5 000 000, the effective resolution of the grid unit dataset would be 10 km × 10 km roughly, rather than 1 km × 1 km. Utilization of the HWSD database at grid cell size of 1 km × 1 km for global SOC pool research would be subject to elevated data redundancy and uncertainty at the map scale (Yu et al., 2011, 2013, 2014), although it is a perfect global soil database in the world at present.

Considering Fig. 2 and Table 1 we see that the influence of the geomagnetic and magnetospheric terms is negligible. Furthermore, Eqs. (1) and (2) add no insight to the problem. We must therefore conclude that Phillips (1999) incorrectly supposed such a connection to exist.

In spite of this negative result, research will continue on this highly interesting question. For if it were to prove correct, then the consequences would be enormous to say the least.

4.2 Comparison of optimal soil raster unit resolutions between calculation and simulation of regional SOC pool

Yu et al. (2014) did similar study by using similar method and same basic data in same region as this study. A difference of method adopted in Yu et al. (2014) from this study was that SOC content (C_i, g kg^{-1}) in Eq. (1) was observed data in 1982 (Yu et al., 2014), which is one of input parameters for DNDC modelling in this study; while the C_i (g kg^{-1}) in this study was simulated data in 2000 by the DNDC modelling. It leads to slight difference of results from each other. For example, the optimal grid sizes matching to 1 : 4 000 000 (N4) and 1 : 14 000 000 (N14) map scales were 9 and 20 km when C_i (g kg^{-1}) was observed (Yu et al., 2014), 8 and 17 km when C_i (g kg^{-1}) was simulated, respectively. But the optimal grid sizes matching to other four map scales

GMDD
8, 2653–2689, 2015

(C5, D2, P5, P1), respectively, did not find any difference, no matter C_i ($\mathrm{g\,kg^{-1}}$) was observed or simulated. Accordingly, the relationships between optimal grid size (y, km) and map scale ($1:x$) has slight difference too in their regression parameters:

$$y = -8.03 \times 10^{-6}x^2 + 0.256x - 0.087 \; (R^2 = 0.9982,\ p < 0.05), \tag{7}$$

observed C_i ($\mathrm{g\,kg^{-1}}$) (Yu et al., 2014).

$$y = -8.0 \times 10^{-6}x^2 + 0.228x + 0.211 \; (R^2 = 0.9994,\ p < 0.05), \tag{8}$$

simulated C_i ($\mathrm{g\,kg^{-1}}$) (this study).

The reason for the slight difference is that more soil features data were used when the C_i ($\mathrm{g\,kg^{-1}}$) was simulated than observed, e.g. soil clay content and pH are two input parameters for DNDC modelling. More Soil features involved implies more rigorous criteria to assess data consistency between grid unit datasets and their parent polygon unit datasets, and leads to increase of optimal raster resolution further, even if the same indices and criteria were applied as Yu et al. (2014) did in their study. Fortunately, the slight difference happened only in polygon unit dataset conversions at small map scales of N4 and N14; and the relationships between optimal grid cell size (y, km) and map scale ($1:x$) revealed in the two studies, respectively, are all described in a quadratic curve regression model (Eqs. 7 and 8).

The quadratic curve regression model (Eq. 8) revealed in this study differ from a standard linear regression too, as Yu et al. (2014) did, which describes the relationship between soil polygon unit map scales and matched default grid cell sizes (Fig. 4). The quadratic model implies that when the map scale is larger than 1 : 4 000 000 (N4) the optimal grid cell size may be larger than the default. Soil grid units at the default cell size converted from polygon units at these map scales will result in data redundancy. When the map scale is less than 1 : 4 000 000 (N4) the optimal grid cell size is less than the default, and the deviation increases with map scale decreasing (Fig. 4). For soil polygon units at these map scales, their conversion to grid units at the default cell size, will result in a drop of data accuracy and an increase in simulation uncertainty. Thus, the quadratic model is more important to soil polygon unit dataset conversion at less N4 map scales than the other map scales. The quadratic model (Eq. 8) also can be substitution of Eq. (7), when C_i ($\mathrm{g\,kg^{-1}}$) was observed for the calculation of regional SOC pool, as the optimal grid unit resolution determined from the Eq. (8) may higher than that from the Eq. (7) at a certain map scale. Soil assessment unit dataset accuracy and result certainty are more critical than the dataset redundancy.

4.3 Application of the quadratic curve regression model for DNDC modelling at different map scales

Almost all map scales of soil polygon unit datasets for China being frequently used are involved in this study, which were generated from the Second National Soil Survey of China. The six soil map scales were designed for soil mapping at different administrative levels including county, district, province and the whole country (Shi et al., 2006).

The Tai Lake region is a typical area in China where paddy soil prevails. Although it is located in the Yangtze Delta plain in East China, where rice fields are integrated with a high density of river or pond, garden and urban land, the spatial pattern of rice field distribution is similar to hilly or mountain regions where rice fields coexist with crop, grass, shrub and forest and urban land (Yu et al., 2011, 2013, 2014). We may assert with some degree of confidence that the knowledge obtained in this present study can be rolled out elsewhere in East and South China where distributes 95 % of rice filed in China (Li, 1992c).

While in the North and West China, soil vector mapping unit is larger in size than that of East and South China at various map scales, because of simpler natural conditions and reduced spatial variability. We may draw a conclude from it that the optimal grid cell size determined from the quadratic model (Eq. 8) can be smaller than the real optimal size in the region (Yu et al., 2011, 2013, 2014). The optimal grid cell size applying will result in a little redundancy of grid unit dataset, but not affect its accuracy matching to their soil polygon units' map scales. Although the quadratic model was obtained from a specific case study, and it would vary with the research region, the knowledge can be used as a guideline for soil unit conversion from polygon to grid, and for optimizing field sampling strategies, to support the regional simulation of SOC pool dynamics in China.

Within China a few administrative region extents are different from those used here, which is caused by their history anthropogeography and physical geography, resulting in additional soil datasets with non-traditional map scales, such as 1 : 75 000, 1 : 100 000 and 1 : 150 000 scales of soil polygon maps for county level, 1 : 250 000 or 1 : 350 000 for district level, and 1 : 750 000 or 1 : 1 500 000 for province level (Shi et al., 2006), respectively. The soil polygon unit conversion for DNDC modelling at these map scales, the optimal grid resolutions can also be informed from the guidelines published here.

5 Conclusion

The DNDC model has been utilized to upscale estimates of SOC from the plot to region scale. For DNDC up-scaled utilization, a region is partitioned into many simulation units, e.g. soil vector polygon units or raster grid units, within which all properties are assumed to be as homogeneous as they are at plot scale. The homogeneity assumption is a possible major source of error when extending DNDC modelling from the plot to region scale. The homogeneity of simulation units is linked to soil polygon units map scale and grid units resolution, which has a strong influence on the results of SOC pool simulation.

Soil grid units are more often applied to SOC pool simulation, as they are more easily manipulated for spatial model simulation, geo-statistics and spatial analysis than soil polygon units. Most of them are derived by data conversion from soil polygon units, but the grid unit resolution choice varies by researcher even if they are derived from a certain vector polygon unit dataset. An optimal raster resolution matched with a certain map scale, for soil polygon unit conversion to grid unit, was put forward in this

GMDD
8, 2653–2689, 2015

study. The optimal raster resolution is the maximum grid cell size of which the soil grid unit dataset and the vector polygon unit dataset are scaled identically. The optimal soil grid unit resolution was found as 0.2 km × 0.2 km, 0.7 km × 0.7 km, 1 km × 1 km, 2 km × 2 km, 8 km × 8 km and 17 km × 17 km at different polygon unit map scales of 1 : 50 000 (C5), 1 : 200 000 (D2), 1 : 500 000 (P5), 1 : 1 000 000 (N1), 1 : 4 000 000 (N4) and 1 : 14 000 000 (N14), respectively. An quadratic curve regression model $y = -8.0 \times 10^{-6} x^2 + 0.228x + 0.211$ ($R^2 = 0.9994$, $p < 0.05$) was revealed in this study too, which describes the relationship between the optimal soil grid unit resolution (y, km) and soil polygon unit map scale (1 : x).

For the investigation and simulation of regional SOC pool, the quadratic curve model is more important to the soil polygon unit conversion at N4 less map scales than the other map scales. Although the quadratic curve model was revealed from a specific case study and would vary with the investigated region, the knowledge can be used as a guideline for soil assessment unit conversion from vector polygon to raster grid, optimizing field sampling strategies, and minimizing uncertainty of the investigation and simulation of regional SOC pool at different map scales further.

Author contributions. D. S. Yu and H. D. Zhang pondered the rationale of the method. X. Z. Shi collected the observed and simulated datasets. Y. L. Ni and L. M. Zhang performed the DNDC model simulation. H. D. Zhang and D. S. Yu prepared the manuscript with contributions from all coauthors.

Acknowledgements. We gratefully acknowledge support for the research from "Strategic Priority Research Program – Climate Change: Carbon Budget and Related Issues" (XDA05050507), the National Basic Research Program of China (2010CB950702), and the Natural Science Foundation of China (40921061).

References

Batjes, N. H.: Effects of mapped variation in soil conditions on estimates of soil carbon and nitrogen stocks for South American, Geoderma, 97, 135–144, 2000.

Bouwman, A. F., Boumans, L. J. M., and Batjes, N. H.: Modeling global annual N_2O and NO emissions from fertilized fields, Global Biogeochem. Cy., 16, 1080, doi:10.1029/2001GB001812, 2002.

Cai, Z. C., Sawamoto, T., Li, C. S., Kang, G. D., Boonjawat, J., Mosier, A., Wassmann, R., and Tsuruta, H.: Field validation of the DNDC model for greenhouse gas emissions in East Asian cropping systems, Global Biogeochem. Cy., 17, GB1107, doi:10.1029/2003GB002046, 2003.

Davidson, E. A. and Janssens, I. A.: Temperature sensitivity of soil carbon decomposition and feedbacks to climate change, Nature, 440, 165–73, 2006.

Eswaran, H., Berg, E. V. D., and Reich, P.: Organic carbon in soils of the world, Soil Sci. Soc. Am. J., 57, 192–194, 1993.

FAO/IIASA/ISRIC/ISSCAS/JRC: Harmonized World Soil Database (version 1.1), FAO, Rome, Italy and IIASA, Laxenburg, Austria, 38 pp., 2009.

Giltrap, D. L., Li, C. S., and Saggar, S.: DNDC: a process-based model of greenhouse gas fluxes from agricultural soils, Agr. Ecosyst. Environ., 136, 292–300, 2010.

Huang, Y., Zhang, W., and Zheng, X.: Modeling methane emission from rice paddies with various agricultural practices, J. Geophys. Res., 109, D08113, doi:10.1029/2003JD004401, 2004.

Lal, R.: Soil carbon sequestration tomitigate climate change, Geoderma, 123, 1–22, 2004.

Levy, P. E., Mobbs, D. C., Jones, S. K., Milne, R., Campbell, C., and Sutton, M. A.: Simulation of fluxes of greenhouse gases from European grasslands using the DNDC model, Agr. Ecosyst. Environ., 121, 186–192, 2007.

Li, C. S.: Quantifying greenhouse gas emissions from soils: scientific basis and modeling approach, Soil Sci. Plant. Nutr., 53, 344–352, 2007.

Li, C. S., Frolking, S., and Frolking, T. A.: A model of nitrous oxide evolution from soil driven by rainfall events: I. Model structure and sensitivity, J. Geophys. Res., 97, 9759–9776, 1992a.

Li, C. S., Frolking, S., and Frolking, T. A.: A model of nitrous oxide evolution from soil driven by rainfall events: II. Model applications, J. Geophys. Res., 97, 9777–9783, 1992b.

Li, C. S., Qiu, J. J., Frolking, S., Xiao, X. M., Salas, W., Moore III, B., Boles, S., Huang, Y., and Sass, R.: Reduced methane emissions from large-scale changes in water management in China's rice paddies during 1980–2000, Geophys. Res. Lett., 29, 1972, doi:10.1029/2002GL015370, 2002.

Li, C. S., Zhuang, Y. H., Frolking, S., Galloway, J., Harriss, R., Moore III, B., Schimel, D., and Wang, X. K.: Modeling soil organic carbon change in croplands of China, Ecol. Appl., 13, 327–336, 2003.
Li, C. S., Mosier, A., Wassmann, R., Cai, Z. C., Zheng, X. H., Huang, Y., Tsuruta, H., Boonjawat, J., and Lantin, R.: Modeling greenhouse gas emissions from rice-based production systems: sensitivity and upscaling, Global Biogeochem. Cy., 18, GB1043, doi:10.1029/2003GB002045, 2004.
Li, C. S., Frolking, S., and Xiao, X. M.: Modeling impacts of farming management alternatives on CO_2, CH_4, and N_2O emissions: a case study for water management of rice agriculture of China, Global Biogeochem. Cy., 19, GB3010, doi:10.1029/2004GB002341, 2005.
Li, C. S., Salas, W., DeAngelo, B., and Rose, S.: Assessing alternatives for mitigating net greenhouse gas emissions and increasing yields from rice production in China over the next twenty years, J. Environ. Qual., 35, 1554–1565, 2006.
Li, H., Qiu, J. J., Wang, L. G., Tang, H. J., Li, C. S., and Eric, V. R.: Modelling impacts of alternative farming management practices on greenhouse gas emissions from a winter wheat–maize rotation system in China, Agr. Ecosyst. Environ., 135, 24–33, 2010.
Li, Q. K.: Paddy Soil of China, Science Press, Beijing, China, 1992 (in Chinese).
Liu, Z. P., Shao, M. A., and Wang, Y. Q.: Effect of environmental factors on regional soil organic carbon stocks across the Loess Plateau region, China, Agr. Ecosyst. Environ., 142, 184–194, 2011.
Ni, J.: Carbon storage in terrestrial ecosystems of China: estimates at different spatial resolutions and their responses to climate change, Climatic Change, 49, 339–358, 2001.
Pathak, H., Li, C., and Wassmann, R.: Greenhouse gas emissions from Indian rice fields: calibration and upscaling using the DNDC model, Biogeosciences, 2, 113–123, doi:10.5194/bg-2-113-2005, 2005.
Qiu, J. J., Wang, L. G., Tang, H. J., Li, H., and Li, C. S.: Studies on the situation of soil organic carbon storage in croplands in northeast of China, Agric. Sci. China, 37, 1166–1171, 2005 (in Chinese with English abstract).
Rüth, B. and Lennartz, B.: Spatial variability of soil properties and rice yield along two catenas in southeast China, Pedosphere, 18, 409–420, 2008.
Schlesinger, W. H.: Evidence from chronosequence studies for a low carbon-storage potential of soils, Nature, 48, 232–234, 1990.

GMDD
8, 2653–2689, 2015

Matching soil grid unit resolutions with polygon unit scales for DNDC modelling

H. D. Zhang et al.

Schlesinger, W. H.: Biogeochemistry: an Analysis of Global Change, Academic Press, San Diego, 1997.
Schmidt, K., Behrens, T., and Scholten, T.: Instance selection and classification tree analysis for large spatial datasets in digital soil mapping, Geoderma, 146, 138–146, 2008.
Shen, Y., Huang, Y., Zong, L. G., Zhang, W., Xu, M., and Liu, L. W.: Simulation and prediction of soil organic carbon dynamics in Jiangsu province based on model and GIS techniques, Scientia Agricultura Sinica, 36, 1312–1317, 2003 (in Chinese with English abstract).
Shi, X. Z., Yu, D. S., Warner, E. D., Sun, W. X., Petersen, G. W., Gong, Z. T., and Lin, H.: Cross-reference system for translating between genetic soil classification of China and soil taxonomy, Soil Sci. Soc. Am. J., 70, 78–83, 2006.
Shi, X. Z., Yang, R. W., Weindorf, D. C., Wang, H. J., Yu, D. S., Huang, Y., Pan, X. Z., Sun, W. X., and Zhang, L. M.: Simulation of organic carbon dynamics at regional scale for paddy soils in China, Climatic Change, 102, 579–593, 2010.
Smith, K. A. and Dobbie, K. E.: The impact of sampling frequency and sampling times on chamber-based measurements of N_2O emissions from fertilized soils, Glob. Change Biol., 7, 933–945, doi:10.1046/j.1354-1013.2001.00450.x, 2001.
Soil Survey Staff (Eds.): Keys to Soil Taxonomy, 6th edn., U. S. Gov. Print. Office, Washington, DC, 437 pp., 1994.
Su, S. L., Zhang, Z. G., Xiao, R., Jiang, Z. L., Chen, T., Zhang, L. M., and Wu, J. P.: Geospatial assessment of agroecosystem health: development of an integrated index based on catastrophe theory, Stoch. Env. Res. Risk A., 26, 321–334, 2012.
Sun, W. J., Huang, Y., Zhang, W., and Yu, Y. Q.: Carbon sequestration and its potential in agricultural soils of China, Global Biogeochem. Cy., 24, GB3001, doi:10.1029/2009GB003484, 2010.
Tang, H., Qiu, J., Eric, V. R., and Li, C.: Estimations of soil organic carbon storage in cropland of China based on DNDC model, Geoderma, 134, 200–206, 2006.
Valade, A., Ciais, P., Vuichard, N., Viovy, N., Caubel, A., Huth, N., Marin, F., and Martiné, J.-F.: Modeling sugarcane yield with a process-based model from site to continental scale: uncertainties arising from model structure and parameter values, Geosci. Model Dev., 7, 1225–1245, doi:10.5194/gmd-7-1225-2014, 2014.
Wan, Y. F., Lin, E. D., Xiong, W., Li, Y., and Guo, L. P.: Modeling the impact of climate change on soil organic carbon stocks in upland soils in the 21st century in China, Agr. Ecosyst. Environ., 141, 23–31, 2011.

GMDD
8, 2653–2689, 2015

Matching soil grid unit resolutions with polygon unit scales for DNDC modelling

H. D. Zhang et al.

Wang, S. H., Shi, X. Z., Zhao, Y. C., Weindorf, D. C., Yu, D. S., Xu, S. X., Tan, M. Z., and Sun, W. X.: Regional simulation of soil organic carbon dynamics for dry farmland in east China by Coupling a 1 : 500 000 soil database with the Century model, Pedosphere, 21, 277–287, 2011.
Xu, Q., Lu, Y. C., Liu, Y. C., and Zhu, H. G.: Paddy Soil of Tai Lake Region in China, Science Press, Shanghai, China, 1980 (in Chinese).
Xu, S. X., Shi, X. Z., Zhao, Y. C., Yu, D. S., Wang, S. H., Zhang, L. M., Li, C. S., and Tan, M. Z.: Modeling carbon dynamics in Paddy soil in Jiangsu Province of China with soil databases differing in spatial resolution, Pedosphere, 21, 696–705, 2011.
Xu, S. X., Shi, X. Z., Zhao, Y. C., Yu, D. S., Wang, S. H., Tan, M. Z., Sun, W. X., and Li, C. S.: Spatially explicit simulation of soil organic carbon dynamics in China's paddy soils, Catena, 92, 113–121, 2012a.
Xu, S. X., Zhao, Y. C., Shi, X. Z., Yu, D. S., Li, C. S., Wang, S. H., Tan, M. Z., and Sun, W. X.: Map scale effects of soil databases on modeling organic carbon dynamics for paddy soils of China, Catena, 104, 67–76, 2012b.
Yang, R. W., Xu, Q., Shi, X. Z., and Huang, Y.: The research of paddy soil organic carbon simulation at regional scale, Anhui Agricultural Science Bulletin, 15, 126–128, 2009 (in Chinese with English abstract).
Yu, D. S., Shi, X. Z., Sun, W. X., Wang, H. J., Liu, Q. H., and Zhao, Y. C.: Estimation of China soil organic carbon storage and density based on 1 : 1,000,000 soil database, Chinese J. Appl. Ecol., 16, 2279–2283, 2005 (in Chinese with English abstract).
Yu, D. S., Shi, X. Z., Wang, H. J., Sun, W. X., Warner, E. D., and Liu, Q. H.: National scale analysis of soil organic carbon stocks in China based on Chinese soil taxonomy, Pedosphere, 85, 680–689, 2007a.
Yu, D. S, Shi, X. Z., Wang, H. J., Sun, W. X., Chen, J. M., Liu, Q. H., and Zhao, Y. C.: Regional patterns of soil organic carbon stocks in China, J. Environ. Manage., 85, 680–689, 2007b.
Yu, D. S., Yang, H., Shi, X. Z., Warner, E. D., Zhang, L. M., and Zhao, Q. G.: Effects of soil spatial resolution on quantifying CH_4 and N_2O emissions from rice fields in the Tai Lake region of China by DNDC model, Global Biogeochem. Cy., 25, GBC2004, doi:10.1029/2010GB003825, 2011.
Yu, D. S., Zhang, L. M., Shi, X. Z., Warner, E. D., Zhang, Z. Q., and Zhao, Q. G.: Soil assessment unit scale affects quantifying CH_4 emissions from rice fields, Soil Sci. Soc. Am. J., 77, 664–672, 2013.

Yu, D. S., Ni, Y. L., Shi, X. Z., Wang, N., Warner, E. D., Liu, Y., and Zhang, L. M.: Optimal soil raster unit resolutions in estimation of soil organic carbon pool at different map scales, Soil Sci. Soc. Am. J., 78, 1079–1086, doi:10.2136/sssaj2013.07.0262, 2014.
Yu, Y. Q., Huang, Y., Zhang, W., and Sun, W. J.: Modeling farmland soil organic carbon dynamics in eastern China: spatio-temporal pattern, Geography Geo-Information Sci., 23, 97–100, 2007 (in Chinese with English abstract).
Yu, Y. Q., Huang, Y., and Zhang, W.: Modeling soil organic carbon change in croplands of China, 1980–2009, Global Planet. Change, 82–83, 115–128, 2012.
Zhang, L., Yu, D., Shi, X., Weindorf, D., Zhao, L., Ding, W., Wang, H., Pan, J., and Li, C.: Quantifying methane emissions from rice fields in the Taihu Lake region, China by coupling a detailed soil database with biogeochemical model, Biogeosciences, 6, 739–749, doi:10.5194/bg-6-739-2009, 2009.
Zhang, L. M., Yu, D. S., Shi, X. Z., Weindorf, D. C., Zhao, L. M., Ding, W. X., Wang, H. J., Pan, J. J., and Li, C. S.: Simulation of global warming potential (GWP) from rice fields in the Tai-Lake region, China by coupling 1 : 50 000 soil database with DNDC model, Atmos. Environ., 43, 2737–2746, 2009.
Zhang, L. M., Yu, D. S., Shi, X. Z., Xu, S. X., Wang, S. H., Xing, S. H., and Zhao, Y. C.: Simulation soil organic carbon change in China's Tai-Lake paddy soils, Soil Till. Res., 121, 1–9, 2012.
Zhang, L. M., Yu, D. S., Shi, X. Z., Xu, S. X., Xing, S. H., and Zhao, Y. C.: Effects of soil data and dimulation unit resolution on quantifying changes of soil organic carbon at regional scale with a biogeochemical process model, PLoS ONE, 9, e88622, doi:10.1371/journal.pone.0088622, 2014.
Zhang, W., Zhang, Q., Huang, Y., Li, T. T., Bian, J. Y., and Han, P. F.: Uncertainties in estimating regional methane emissions from rice paddies due to data scarcity in the modeling approach, Geosci. Model Dev., 7, 1211–1224, doi:10.5194/gmd-7-1211-2014, 2014.
Zhao, Y. C., Shi, X. Z., Weindorf, D. C., Yu, D. S., Sun, W. X., and Wang, H. J.: Map scale effects on soil organic carbon stock estimation in north China, Soil Sci. Soc. Am. J., 70, 1377–1386, 2006.
Zhu, Q. and Zhuang, Q.: Parameterization and sensitivity analysis of a process-based terrestrial ecosystem model using adjoint method, J. Adv. Model. Earth Syst., 6, 315–331, doi:10.1002/2013MS000241, 2014.

GMDD
8, 2653–2689, 2015

Matching soil grid unit resolutions with polygon unit scales for DNDC modelling

H. D. Zhang et al.

Table 1. Statistics of soil parameters input from different resolution units at the map scale of 1 : 50 000 in the Tai Lake region of China*.

Simulation units	Clay (%) Mean	Clay (%) CV	pH Mean	pH CV	SOC ($g\,kg^{-1}$) Mean	SOC ($g\,kg^{-1}$) CV	Bulk density ($g\,cm^{-3}$) Mean	Bulk density ($g\,cm^{-3}$) CV
Polygon	29.00	37.07	6.65	9.77	16.81	33.26	1.18	10.17
0.1 km	29.01	37.06	6.67	9.75	16.79	33.34	1.18	10.17
0.2 km	28.85	37.54	6.67	9.90	16.56	33.54	1.18	10.17
0.3 km	28.35	38.77	6.67	10.04	16.30	33.45	1.19	10.08
0.4 km	27.71	40.24	6.67	10.19	16.00	33.38	1.20	10.00
0.5 km	27.20	41.62	6.67	10.34	15.74	33.43	1.20	10.00
1 km	25.61	44.01	6.71	10.88	14.94	33.66	1.22	9.84
2 km	24.55	44.77	6.74	11.13	14.46	34.14	1.22	9.84
3 km	24.45	44.09	6.77	10.78	14.33	34.52	1.22	9.84

* Mean: mean value; CV: coefficient of variation.

GMDD
8, 2653–2689, 2015
Matching soil grid unit resolutions with polygon unit scales for DNDC modelling
H. D. Zhang et al.

Table 2. Statistics of soil parameters input from different resolution units at the map scale of 1 : 200 000 in the Tai Lake region of China*.

Simulation units	Clay (%) Mean	Clay (%) CV	pH Mean	pH CV	SOC ($g\,kg^{-1}$) Mean	SOC ($g\,kg^{-1}$) CV	Bulk density ($g\,cm^{-3}$) Mean	Bulk density ($g\,cm^{-3}$) CV
Polygon	26.77	40.79	6.96	10.06	29.42	33.92	1.16	8.62
0.4 km	26.64	40.99	6.96	10.20	29.50	33.86	1.16	8.62
0.5 km	26.72	41.65	6.93	10.25	29.37	33.95	1.16	8.62
0.6 km	26.75	42.21	6.90	10.43	29.14	33.87	1.16	9.48
0.7 km	26.71	43.13	6.88	10.47	29.03	34.14	1.16	9.48
0.8 km	26.67	43.94	6.86	10.64	28.93	34.39	1.16	9.48
1 km	27.20	39.74	6.99	9.87	28.89	33.16	1.16	8.62
2 km	27.13	38.85	7.00	10.00	28.46	31.59	1.16	8.62
4 km	26.80	41.75	6.95	10.22	28.45	29.77	1.16	9.48

* Mean: mean value; CV: coefficient of variation.

GMDD
8, 2653–2689, 2015

Table 3. Statistics of soil parameters input from different resolution units at the map scale of 1 : 500 000 in the Tai Lake region of China*.

Simulation units	Clay (%)		pH		SOC ($g\,kg^{-1}$)		Bulk density ($g\,cm^{-3}$)	
	Mean	CV	Mean	CV	Mean	CV	Mean	CV
Polygon	25.26	44.02	6.97	10.47	17.50	32.85	1.16	8.62
1 km	25.61	47.36	6.79	11.63	17.06	34.12	1.17	10.26
2 km	25.63	49.63	6.67	11.99	16.83	34.46	1.18	11.02
3 km	25.84	50.23	6.67	12.29	16.70	35.02	1.19	10.92
4 km	26.31	50.29	6.63	12.37	16.67	36.77	1.19	10.08
5 km	26.56	49.10	6.63	12.22	17.06	34.53	1.20	10.83
6 km	26.86	50.22	6.59	12.14	17.06	34.87	1.20	10.83

* Mean: mean value; CV: coefficient of variation.

GMDD
8, 2653–2689, 2015

Table 4. Statistics of soil parameters input from different resolution units at the map scale of 1 : 1 000 000 in the Tai Lake region of China*.

Simulation units	Clay (%) Mean	CV	pH Mean	CV	SOC ($g\,kg^{-1}$) Mean	CV	Bulk density ($g\,cm^{-3}$) Mean	CV
Polygon	27.01	35.99	6.56	10.98	29.01	37.61	1.15	7.83
2 km	26.36	37.78	6.53	11.03	28.71	38.84	1.15	8.70
3 km	25.95	37.88	6.56	10.82	28.28	39.53	1.15	8.70
4 km	27.54	38.31	6.61	12.41	24.45	41.76	1.12	8.93
5 km	26.78	37.49	6.61	10.29	28.50	37.61	1.15	7.83
6 km	26.69	36.98	6.63	10.26	28.78	36.55	1.16	7.76
7 km	27.32	36.20	6.66	9.91	28.29	35.49	1.16	7.76

* Mean: mean value; CV: coefficient of variation.

GMDD
8, 2653–2689, 2015

Table 5. Statistics of soil parameters input from different resolution units at the map scale of 1 : 4 000 000 in the Tai Lake region of China*.

Simulation units	Clay (%) Mean	Clay (%) CV	pH Mean	pH CV	SOC ($g\,kg^{-1}$) Mean	SOC ($g\,kg^{-1}$) CV	Bulk density ($g\,cm^{-3}$) Mean	Bulk density ($g\,cm^{-3}$) CV
Polygon	25.95	26.51	6.43	6.84	15.69	33.69	1.12	6.25
8 km	26.31	25.54	6.41	7.18	15.50	33.86	1.13	6.19
9 km	25.99	26.63	6.39	6.73	15.42	35.95	1.13	6.19
10 km	27.00	23.78	6.41	6.71	15.85	33.96	1.13	6.19
12 km	27.23	23.32	6.47	5.10	16.01	33.83	1.12	5.36
14 km	27.52	23.33	6.50	4.77	15.75	32.81	1.12	4.46
16 km	27.89	20.04	6.51	4.45	15.58	28.84	1.12	4.46
18 km	28.97	16.98	6.53	4.29	16.14	26.55	1.13	4.42

* Mean: mean value; CV: coefficient of variation.

GMDD
8, 2653–2689, 2015

Table 6. Statistics of Input Soil Parameters for Different Resolution Unit at map scale of 1 : 14 000 000 in the Tai Lake Region of China*.

Simulation units	Clay (%) Mean	CV	pH Mean	CV	SOC ($g\,kg^{-1}$) Mean	CV	Bulk density ($g\,cm^{-3}$) Mean	CV
Polygon	33.46	18.29	6.51	7.37	33.38	31.34	1.14	2.63
17 km	32.77	18.31	6.61	6.51	32.34	31.08	1.14	2.63
18 km	32.91	18.14	6.60	6.67	32.11	30.36	1.14	2.63
19 km	33.39	18.75	6.58	6.53	33.39	32.32	1.14	2.63
20 km	32.12	16.69	6.58	6.99	31.23	29.84	1.14	2.63
21 km	33.10	18.31	6.59	6.83	32.27	31.14	1.14	2.63
22 km	33.35	17.96	6.54	7.03	33.09	29.07	1.14	2.63
23 km	33.09	18.68	6.59	6.68	32.33	32.11	1.14	2.63
24 km	33.07	17.96	6.54	6.88	32.79	30.16	1.14	2.63
25 km	33.12	17.30	6.52	6.75	33.06	29.19	1.14	2.63
26 km	33.17	18.66	6.57	6.85	32.87	31.15	1.14	2.63
27 km	32.60	18.10	6.61	6.96	31.29	31.80	1.14	2.63
28 km	33.13	17.27	6.54	6.88	32.13	30.10	1.14	2.63

* Mean: mean value; CV: coefficient of variation.

GMDD
8, 2653–2689, 2015

Table 7. Index values determined from DNDC simulations with the paddy polygon units at different map scales in the Tai Lake Region of China*.

Simulation units	SPN	Index values from vector simulation unit (IV-$_{vector}$)						
		SOCS	AREA	ASOCD	STN			
		(Tg)	(M ha)	(kg C m^{-2})	S1	S2	S3	S4
C5 (1 : 50 000)	52 304	144.78	2.32	6.24	622	137	6	1
D2 (1 : 200 000)	7263	168.78	2.60	6.48	127	78	6	1
P5 (1 : 500 000)	4766	172.04	2.53	6.71		68	6	1
N1 (1 : 1 000 000)	967	161.21	2.59	6.24		48	6	1
N4 (1 : 4 000 000)	32	167.55	2.74	6.12			6	1
N14 (1 : 14 000 000)	8	207.73	2.80	7.42			2	1

*SOCS: SOC stocks of surface paddy soil; AREA: paddy soil area; ASOCD: average SOC density of surface paddy soil; STN: paddy soil type number; SPN: paddy soil unit number; S1: soil species; S2: soil family; S3: soil subgroup; S4: soil great group (Paddy soil).

GMDD
8, 2653–2689, 2015

Matching soil grid unit resolutions with polygon unit scales for DNDC modelling

H. D. Zhang et al.

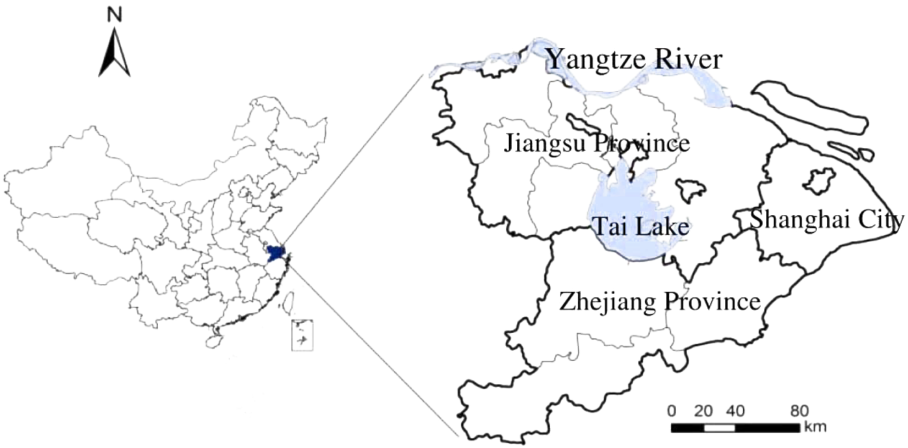

Figure 1. The location of Tai Lake region.

GMDD
8, 2653–2689, 2015

Matching soil grid unit resolutions with polygon unit scales for DNDC modelling

H. D. Zhang et al.

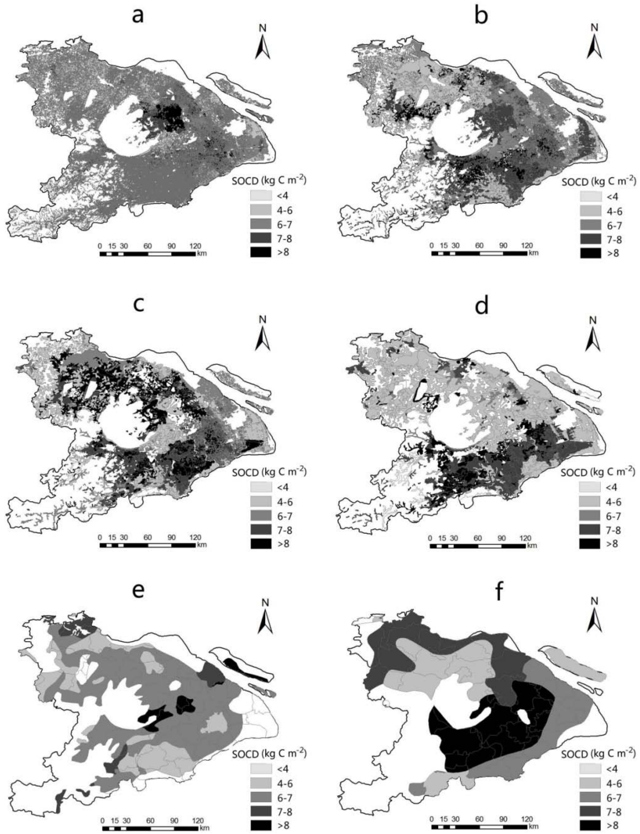

Figure 2. Map of soil organic carbon density (SOCD) simulated by DNDC from vector paddy soil units at different map scales in the Tai Lake region of China. (**a**, 1 : 50 000; **b**, 1 : 200 000; **c**, 1 : 500 000; **d**, 1 : 1 000 000; **e**, 1 : 4 000 000; **f**, 1 : 14 000 000).

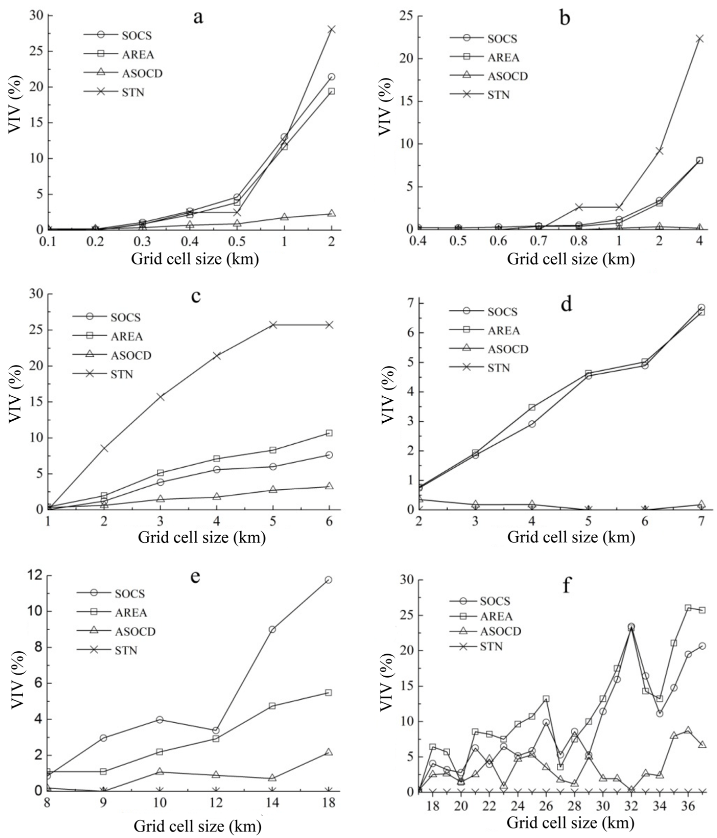

Figure 3. VIVs varied with grid unit resolutions at different soil unit map scales in the Tai Lake region of China (VIV, Variation of an index value; SOCS, soil organic carbon stocks simulated by DNDC; AREA, soil area; ASOCD, average soil organic carbon density simulated by DNDC; STN, soil type number; **a**, C5 (1 : 50 000); **b**, D2 (1 : 200 000); **c**, P5 (1 : 500 000); **d**, N1 (1 : 1 000 000); **e**, N4 (1 : 4 000 000); **f**, N14 (1 : 14 000 000)).

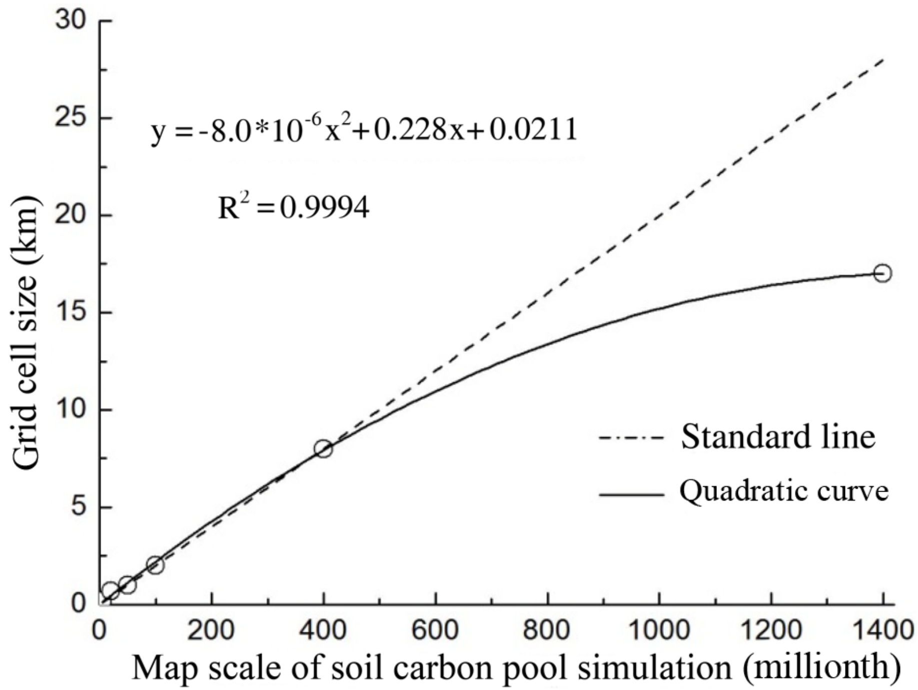

Figure 4. Relationship between paddy polygon unit map scale and matched optimal grid unit resolution for the SOC simulation with DNDC in the Tai Lake region of China.