Biogeosciences Discuss., doi:10.5194/bg-2016-473, 2017
Manuscript under review for journal Biogeosciences
Published: 2 January 2017

No significant changes in topsoil carbon in the grasslands of northern China between the 1980s and 2000s

Author: Shangshi Liu[1,2], Yuanhe Yang[1], Haihua Shen[1], Huifeng Hu[1], Xia Zhao,[1] He Li[1,2], Taoyu Liu[1,2], Jingyun Fang[1,3*]

[1]State Key Laboratory of Vegetation and Environmental Change, Institute of Botany, Chinese Academy of Sciences, Beijing 100093, China

[2]University of Chinese Academy of Sciences, Beijing 100049, China

[3] Department of Ecology, Key Laboratory for Earth Surface Processes of the Ministry of Education, Peking University, Beijing 100871, China

Corresponding author: Professor Jingyun Fang, Tel: +86-10-6276 5578, Fax: +86-10-6275 6560, E-mail: jyfang@urban.pku.edu.cn

Biogeosciences Discuss., doi:10.5194/bg-2016-473, 2017
Manuscript under review for journal Biogeosciences
Published: 2 January 2017

Abstract

The grasslands of northern China store a large amount of soil organic carbon (SOC), and the small changes in SOC stock could significantly affect the regional C cycle. However, recent estimates of SOC changes in this region are highly controversial. In this study, we examined the changes in the SOC density (SOCD) in the upper 30 cm of the grasslands of northern China between the 1980s and 2000s, using an improved approach that integrates field-based measurements into machine learning algorithms (artificial neural network and random forest). The random forest-generated SOCD averaged 5.55 kg C m^{-2} in the 1980s and 5.53 kg C m^{-2} in the 2000s. The change ranged between -0.17 and 0.22 kg C m^{-2} at the 95% confidence level, suggesting that the overall SOCD did not change significantly during the study period. However, the change in SOCD exhibited large regional variability. The topsoil of the Inner Mongolian grasslands experienced a significant C loss (4.86 *vs*. 4.33 kg C m^{-2}), whereas that of the Xinjiang grasslands exhibited an accumulation of C (5.55 *vs*. 6.46 kg C m^{-2}). In addition, the topsoil C in the Tibetan alpine grasslands remained relatively stable (6.12 *vs*. 6.06 kg C m^{-2}). A comparison of different grassland types indicated that SOCD exhibited significant decreases in typical steppe, whereas showed increases in mountain meadow, and were stable in the remaining grasslands (alpine meadow, alpine steppe, mountain steppe and desert steppe). Climate change could partly explain these changes in the SOCD of the different grassland types. Increases in precipitation could lead to SOC increase in temperate grasslands and SOC loss in alpine grasslands, while climate warming is likely to cause SOC loss in temperate grasslands. Overall, our study shows that northern grasslands in China remained a neutral SOC sink between the 1980s and 2000s.

Keywords: alpine grasslands, artificial neural network, carbon density, climate change, grassland ecosystems, random forests, soil organic carbon, temperate grasslands.

Biogeosciences Discuss., doi:10.5194/bg-2016-473, 2017
Manuscript under review for journal Biogeosciences
Published: 2 January 2017

1 Introduction

Soil is the largest carbon (C) pool in the terrestrial ecosystem (Batjes, 1996), and even a few percent changes in stored C can have profound impacts on terrestrial C cycling (Johnson *et al*., 2007). Previous studies have shown that climate warming will likely accelerate the microorganism decomposition of soil organic C (SOC) and induce greater carbon dioxide (CO_2) emissions into the atmosphere, thus promoting positive C-climate feedback. However, this C loss from soil might be offset if the increased input of plant C to soil exceeds the increase in decomposition, which would promote negative C-climate feedback (Davidson and Janssens, 2006). Hence, accurately unraveling SOC dynamics under climate change is crucial to determine whether positive C-climate feedback has occurred.

Grasslands are widely distributed in temperate regions of the world, and most C in grasslands stores in soil. (Fang *et al*., 2010). Due to this point, the SOC dynamics of grasslands have attracted considerable attention in recent decades (e.g., Hanegraaf *et al*., 2009; Mestdagh *et al*., 2009; Yang *et al*., 2010). However, soil produces the largest amount of uncertainty with respect to estimating the terrestrial C budget (Ciais *et al*., 2013; Todd-Brown *et al*., 2013). For example, using a repeated soil inventory, Bellamy *et al*. (2005) reported that soils in England and Wales, including grassland soils, had experienced a significant C loss from 1978 to 2003. However, Emmett *et al*. (2010) analyzed a large-scale survey of soil profiles and found that the topsoil C concentration in Great Britain did not change as much as that reported by Bellamy *et al*. (2005) during the same study period. As an important component of global grassland ecosystem, nature grasslands in China cover more than 40% of country's territory, and the C in soil account for about 96.6 % grassland ecosystem C stock in this area (Fang *et al*., 2010). Similarly, contrasting studies have also been reported with respect to the estimation of the SOC change in China's

Biogeosciences Discuss., doi:10.5194/bg-2016-473, 2017
Manuscript under review for journal Biogeosciences
Published: 2 January 2017

northern grasslands. Using the Terrestrial Ecosystem Model, Yan *et al*. (2015) suggested that Tibetan alpine grasslands sequestered C over the past 50 years at an annual rate of 10.1 Tg C. In contrast, Yang *et al*. (2008, 2010) demonstrated that the SOC stock in the grasslands of northern China did not change considerably between the 1980s and 2000s. These conflicting results suggest that intensive investigations and appropriate methodology are required to accurately assess the SOC dynamics of the grasslands in northern China.

In this study, we used repeated observations from two periods in the 1980s and 2000s and two advanced machine learning algorithms, artificial neural network (ANN) and random forest (RF) algorithms, to estimate long-term SOC changes in grasslands of northern China. The ANN and RF approaches consider the spatial relationships between study sites and the influence of climatic and edaphic parameters and could thus produce more accurate predictions compared with traditional spatial interpolation approaches, such as process-based models or Kriging-based spatial interpolations (Grimm *et al*., 2008; Li *et al*., 2013; Sreenivas *et al*., 2014; Yang *et al*., 2014). Specifically, we aimed to answer the following three questions: (1) Do the new approaches we used produce different estimates of SOC stocks and spatial heterogeneity compared with those obtained from previous studies? (2) How do the SOC changes vary spatially across the grasslands of northern China? and (3) How these changes are associated with climatic factors and vegetation types?

2 Materials and methods

2.1 Soil inventories

Soil profile data in the 1980s were obtained from the National Soil Inventory (Wu, 1991; Hou and Zeng, 1992; Feng and Wang, 1993; Zhao, 1993; Gong, 1994; Liu, 1995) and were used to assess the historical status of the SOC density (SOCD) in the grasslands of northern China. The soil inventory recorded the physical and chemical properties of the soils, such as geographical location, land cover, layer thickness, bulk density, proportion of rock fragments, and soil organic matter (SOM) concentration, which was determined using the Walkley-Black method (Nelson and Sommers, 1982). We used a constant value of 0.58 to convert SOM to SOC (Wu *et al*., 2003; Xie *et al*., 2007; Yang *et al*., 2010). A detailed description of the inventory soil profiles can be found in Yang *et al*. (2010).

The soil data in the 2000s were obtained from field sampling campaigns conducted from 2001-2005 across the grasslands of northern China (Yang *et al*., 2010). The SOC concentration was determined following the same methodology as the National Soil Inventory (*i.e*., Walkley-Black method). The descriptions of this sampling method and the soil profiles (with environmental variables) are detailed in Yang *et al*. (2008, 2010).

2.2 Vegetation information and climate/Normalized Difference Vegetation Index (NDVI) data

According to the vegetation map of China (Chinese Academy of Sciences, 2001), six major grassland types occur in China: desert steppe (DS), typical steppe (TS), meadow steppe (MS), mountain meadow (MM), alpine steppe (AS), and alpine meadow (AM) (Fig. 1). To understand the variations in the SOCD dynamics among the different grassland types, the vegetation at each study site was classified into temperate grasslands (*i.e*., DS, TS, MS, and MM) and alpine grasslands (*i.e*., AS and AM) to examine the relationships between the SOCD dynamics and climate change. In addition, to compare our results with those of previous regional studies, we

Biogeosciences Discuss., doi:10.5194/bg-2016-473, 2017
Manuscript under review for journal Biogeosciences
Published: 2 January 2017

divided our study area into three landscape regions: Inner Mongolia, Xinjiang (northwestern China), and the Tibetan Plateau.

To examine the relationships between the SOCD dynamics and environmental variables, we documented the following environmental variables for each site where soil profile information was collected: mean annual temperature (MAT), mean annual precipitation (MAP), seasonal climate variables (precipitation, mean temperature, maximum temperature and minimum temperature of each season). (National Meteorological Information Center; http://www.nmic.gov.cn), and annual NDVI data at a spatial resolution of 0.1° × 0.1° (GIMMS NOAA/AVHRR NDVI) (Tucker *et al*., 2005).

2.3 Model validation and prediction

First, we obtained site-level observations from Yang *et al*. (2010). Specifically, Yang *et al*. (2010) calculated the SOCD in the top 30 cm of soil using Eq. 1 to provide a consistent comparison with China's National Soil Inventory from the 1980s. Because of the lack of bulk density information for certain profiles surveyed in the 1980s, Yang *et al*. (2010) developed an empirical relationship between bulk density and the SOC concentration (Eq. 2).

$$SOCD = \sum_{i=1}^{n} T_i \times BD_i \times SOC_i \times (1 - C_i/100)/100, \qquad (1)$$

$$BD = 0.29 + 1.296\exp(-0.0167\mathrm{SOC}) \qquad (2)$$

$(r^2 = 0.63, P = 0.0001)$

Second, we predicted the SOCD using the RF and ANN approaches (Fig. 2). The RF approach is a method that consists of averaging multiple deep decision trees trained on different parts of the same training set, and the goal is to reduce variance (Hastie *et al*., 2009). The general technique of bootstrap aggregating was applied in the training algorithm of the RF model. The typical ANN used in this study included an input layer, a hidden layer and an output layer. Based on the National Soil Inventory from the 1980s and related environmental parameters, we trained the RF model and ANN to estimate the SOCD during the 1980s for those sites that were surveyed during the 2000s. Moreover, to assess the SOCD in the 2000s for sites that were surveyed in the 1980s, we used field observations recorded by Yang *et al*. (2010) and the related environmental parameters to construct the RF model and ANN. Specifically, the input variables of both approaches included latitude, longitude, altitude, grassland type, soil texture, NDVI, MAT, MAP, and precipitation, mean temperature, maximum temperature, minimum temperature of each season. The collected data sets for each period were divided into two parts. The training data set (90% of all data) was used to construct the RF model and ANN. The test data set (10% of all data) was used to validate the reliability of the two approaches. Moreover, the input variables of the ANN were normalized to a scale between -1 and 1, and the predictions of the ANN were rescaled to the actual SOCD. The predicted results were compared with field measurements to evaluate the validation and performance of the ANN and RF model. We repeated the training and prediction process 5000 times and adapted the mean value in both approaches. The parameters of both algorithms were optimized to achieve their best performance.

The final results were used to estimate the SOCD in the two periods. The SOCD changes in the entire grassland region, three different regions, and six grassland types were examined using a paired *t*-test. Furthermore, linear regression model was used to explore the relationship between the SOCD changes and the climate variables by grassland type.

2.4 Uncertainty analysis

The accuracy of SOCD temporal changes was assessed by Monte Carlo simulation in this study. Here, to quantify the uncertainties of our estimation, we performed 5000 Monte Carlo simulations for each sites. During each stimulation, the input variables were randomly generated based on the probability distribution functions of their errors. Specifically, the error of climate data (*i.e.*, MAT, MAP etc.) were calculated from the interpolation process. For the errors associated with the extrapolate models prediction of SOC density, we used values of the standard deviation of the error derived from the cross-validation, we applied this error randomly to each prediction in every sites. Moreover, for the soil profiles which bulk density were not recorded, we also use the error of relative pedotransfer function to present the uncertainty induced by lackness of bulk density record.

We also assumed any actual SOCD measurements was enough to present the true SOC stock, thus the uncertainty of SOC change was generated by those sites had not been surveyed. Finally, the interquartile range of the 5000 iterations of the simulation for each unsurveyed sites were calculated, and the summed interquartile were used to assess the relative uncertainty of the SOCD change over each region and grassland type (Ding *et al*., 2016). All analyses were performed using R software (R Development Core Team, 2014).

3 Results

3.1 Topsoil C changes and its uncertainty

We tested the performance of the RF and ANN algorithms according to the root mean square error (RMSE) and determined the coefficient of variation (r^2) through 5000 iterations. Positive linear relationships were observed between the measured and predicted SOCD during the 1980s

Biogeosciences Discuss., doi:10.5194/bg-2016-473, 2017
Manuscript under review for journal Biogeosciences
Published: 2 January 2017

and 2000s; the mean r^2 was 0.73 for RF and 0.62 for ANN, and the mean RMSE was 1.98 for RF and 2.39 for ANN (Fig. 3), demonstrating that both algorithms generated accurate and dependable predictions of SOCD during both time periods.

The SOCD during the 1980s for the sites surveyed in the 2000s was estimated as 5.21 kg C m^{-2} using the ANN approach and 5.39 kg C m^{-2} using the RF approach (Table 1). The SOCD during the 2000s for the sites surveyed in the 1980s was estimated as 5.70 kg C m^{-2} (ANN) and 5.97 kg C m^{-2} (RF). The comparison between the actual measurements and the predictions from both approaches showed that there were no significant differences in these two methods (paired *t*-test, $P > 0.05$), suggesting that the ANN and RF methods generated accurate estimates. Notably, we only presented RF estimates in further analyses because of a higher r^2 and lower RMSE.

By summarizing the estimates based on the measured SOCD profiles combined with the predicted data, we estimated the mean SOCD from 573 sites as 5.55 kg C m^{-2} and 5.53 kg C m^{-2} for the 1980s and 2000s, respectively. A paired *t*-test indicated that the SOCD during the 2000s was not significantly different from the corresponding value in the 1980s at the site level (P = 0.81), the change amounted to -0.17~ 0.22 kg C m^{-2} at the 95% confidence level. Taking the area of each grassland type into consideration, the total change in the SOC stock in the upper 30 cm depth was estimated as 11.03 Pg C in the 1980s and 11.16 Pg C in the 2000s. The overall rate of the change in the SOCD at this depth was 3.68 g C m^{-2} yr^{-1} (Table 2).

3.2 SOCD changes in different grassland types and regions

We illustrated spatial distributions in the SOCD in the topsoil in the 2000s (Fig. 4a) and its changes between the 1980s and 2000s (Fig. 4b) using the RF approach. Although there was no significant difference between the mean SOCD in the 2000s and the corresponding value in the

1980s at the site level (Tables 1, 2 and 3), the spatial distribution and temporal change of SOCD among different regions and grassland types were markedly different (Fig. 4). The mean SOCD decreased significantly in Inner Mongolia (4.86 *vs*. 4.33 kg C m^{-2}, $P < 0.01$), at an average rate of -28.46 g C m^{-2} yr^{-1}, but increased significantly in Xinjiang (5.55 *vs*. 6.46 kg C m^{-2}, $P < 0.01$), at an average rate of 70.36g C m^{-2} yr^{-1}. The SOCD was relatively stable in the Tibetan alpine grasslands (6.12 *vs*. 6.06 kg C m^{-2}, $P = 0.70$) (Fig. 5a). When considering the different grassland types, the mean SOCD increased significantly in the MM and decreased significantly in the TS, but no significant changes occurred in the AS, MS, AM and DS (Fig. 5b).

3.3 Effects of climate and soil geochemistry on SOCD changes

The SOCD change values for the temperate grasslands decreased with the temperature of spring and summer ($P < 0.01$) and increased with the precipitation change values ($P < 0.01$) (Figure 6 and 7). However, in the alpine grasslands, the SOCD change values decreased with the MAP change values ($P < 0.01$), and there was no significant trend related to other annual or seasonal temperature change values (Fig. 6 and 7). Moreover, we found a significant negative linear correlation between the change in SOCD and the original SOCD ($P < 0.01$, Fig. 8) across northern China. This relationship was also observed in the temperate and alpine grasslands ($P < 0.01$, Fig. 8).

4 Discussion

4.1 SOC changes in China's grasslands

Several studies have explored SOC dynamics across China's grasslands and presented considerably different and even contrasting estimates because various approaches were used

(Table 4). For example, Huang *et al*. (2010) estimated the C sink rate in soils of Chinese grasslands as 4.9±1.9 Tg C yr^{-1} (from 1981 to 2000). In contract, Xie *et al*. (2007) claimed that Chinese grasslands experienced tremendous C losses in soil (-178.2 Tg C yr^{-1}). In a recent study, Yang *et al*. (2010) noted that there were insignificant changes in the SOC of grasslands in northern China, based on a comparison between historical national soil inventory data and their field soil campaign data. In the current study, we used a machine learning approach, national soil inventory data and our own field survey and concluded that while Chinese grasslands remained an neutral SOC sink, but it was spatially heterogeneous during the 1980s to the 2000s.

The uncertainties in the estimation of SOC changes are largely induced by insufficient observations and methodological issues (Yang *et al*., 2014). In this study, we used two machine learning algorithms (*i.e*., ANN and RF) to estimate the SOCD in the 2000s for sites that were only surveyed in the 1980s, and vice versa. These algorithms generated promising predictions with high coefficients of determination (r^2) and low RMSE values. For both approaches, we integrated field-based measurements and related environmental factors to train the models and thus acquired comparable SOCD values for the two time periods at each site by generating reliable SOCD predictions for unsurveyed sites at each time period. The underlying assumptions of this approach were that the grassland type and soil texture did not change substantially in two decades, and the SOCD change was driven by the balance of C inputs from vegetation litter (inferred from Temperature, Precipitation and NDVI) and SOC decomposition (inferred from Temperature and Precipitation). Compared with other widely used spatial extrapolation methods, such as Kriging-based interpolations, a machine learning method combined with multiple sources of relevant information shows better performance in theory and practice (Table S1)(Li *et al*., 2013).

Biogeosciences Discuss., doi:10.5194/bg-2016-473, 2017
Manuscript under review for journal Biogeosciences
Published: 2 January 2017

Although we used an improved approach that integrates field survey data into machine learning-based models to determine SOCD dynamics, there are some uncertainties in our study (Table 3). The major sources of the uncertainties may originate from the models and their input variables, as well as changes in land cover. First, the uncertainties and errors of the input variables of the models may generate uncertainties in our estimations. For example, the lack of bulk density data for some soil profiles may result in uncertainties in our estimates because bulk density is a key variable for calculating SOCD. However, Yang *et al*. (2010) developed a well-quantified relationship between bulk density and SOC ($r^2 = 0.63$), and using it may not result in a large bias in our analysis. Climate data (*i.e*., MAT and MAP) are also a source of uncertainty because they are based on spatial interpolations. Furthermore, uneven spatial distribution of field measurements may also introduce additional uncertainties.

Second, we used two machine learning models (ANN and RF) to predict regional SOCD distribution from field-based observations, which will certainly produce some uncertainties because model simulation is sensitive to upscaling, and the extrapolation ability of any model is limited. However, the simulation performance of both models (ANN and RF) illustrated a good agreement between the predicted and field-observed SOCD (Fig. 3). This suggests that the models we used are appropriate for the estimation of SOC in Chinese grasslands. Finally, we did not consider the effects of land cover changes on the estimation of SOCD dynamics, which may introduce some uncertainties. However, there were no substantial land cover changes across most of the study region; thus, it is likely that the estimation was not greatly biased.

4.2 Effects of climate changes on SOC changes in different grassland types

As shown in Fig. 6 and 7, climate change may play a critical role in controlling the SOCD changes, and different grassland types (alpine and temperate grasslands) revealed different responses to climate change. Increasing temperature, especially spring and summer maximum temperature, induced a C loss from the topsoil of temperate grasslands ($P < 0.001$), but this trend was not observed in the Tibetan alpine grasslands where the topsoil SOC remained relatively stable ($P = 0.70$). The decrease in the topsoil C of the temperate grasslands may be caused by increased substrate utilization efficiencies of soil microbial communities (Zogg *et al*., 1997) and decreased C input into soil due to reduced biomass when temperature increases in growing season (Yang *et al*., 2010). On the other hand, no significant change in SOC in the alpine grasslands has been shown by soil warming manipulation experiments (Zhang *et al*. 2015) and a soil transplantation experiment (Yue *et al*. 2015).

Precipitation is another critical factor that affected the SOCD change (Table 3). With an increase in precipitation, the SOCD increased in temperate grasslands ($P < 0.0001$) and decreased in alpine grasslands ($P < 0.001$). The underlying mechanisms that change in precipitation affects the SOC change in grassland ecosystems are complex (Bellamy *et al*., 2005). On the one hand, increased precipitation leads to an increase in plant growth and biomass and thus increased C input into soils in dry climates; on the other hand, it increases soil respiration and therefore enhances C emissions from soils (Thomey *et al*., 2011; Zhang *et al*., 2015). In very cold and humid regions, such an interaction between the SOC and changes in precipitation may produce an opposite result. Because we did not have regional soil respiration data in this study, we only examined how changes in precipitation affect plant growth in the temperate and alpine grasslands, using NDVI as an indicator of vegetation growth (Piao *et al*., 2005). There was a significantly positive relationship between the precipitation and NDVI changes in the temperate grasslands (Fig. 7a), while no such trend was found in the alpine grasslands (Fig. 7b). This suggests that increased precipitation does not significantly influence the SOC in alpine grasslands but could accelerate vegetation growth and therefore increase C input into the soil in temperate grasslands.

Biogeosciences Discuss., doi:10.5194/bg-2016-473, 2017
Manuscript under review for journal Biogeosciences
Published: 2 January 2017

5 Conclusion

In summary, we integrated field survey data into two machine learning algorithms (ANN and RF) to estimate the SOC stocks and dynamics in grasslands of northern China. We found that the overall SOC stocks did not change significantly between the 1980s and the 2000s, which indicates that the soils of these northern grasslands are a neutral C sink. Despite this overall stable SOC stock, there were large spatial differences in the SOCD changes between different regions and grassland types, suggesting differences in the climate changes and in the interactions between ecosystem C processes and climate systems. Our analysis indicated that the SOC exhibited a positive feedback to climate warming in the temperate grasslands, while this was not found for the alpine grasslands. In addition, the response of the SOCD dynamics to the changes in precipitation depended on grassland type.

Acknowledgements

This study was funded by the National Basic Research Program of China on Global Change (2014CB954001), the National Natural Science Foundation of China (31330012, 31321061, 31470525), and the Strategic Priority Research Program of the Chinese Academy of Sciences (XDA05050000).

Biogeosciences Discuss., doi:10.5194/bg-2016-473, 2017
Manuscript under review for journal Biogeosciences
Published: 2 January 2017

References

Batjes, N.H. (1996) Total carbon and nitrogen in the soils of the world. *European Journal of Soil Science*, 47, 151-163.

Bellamy, P.H., Loveland, P.J., Bradley, R.I., Lark, R.M. and Kirk, G.J. (2005) Carbon losses from all soils across England and Wales 1978-2003. *Nature*, 437, 245-8.

Chinese Academy of Sciences. (2001) *Vegetation atlas of China*. Science Press, Beijing.

Ciais, P., Sabine, C., Bala, G., Bopp, L., Brovkin, V., Canadell, J., Chhabra, A., Defries, R., Galloway, J. and Heimann, M. (2013) Carbon and other biogeochemical cycles. 465-570.

Dai, E., Zhai, R., Ge, Q. and Wu, X. (2014) Detecting the storage and change on topsoil organic carbon in grasslands of Inner Mongolia from 1980s to 2010s. *Journal of Geographical Sciences*, 24, 1035-1046.

Davidson, E.A. and Janssens, I.A. (2006) Temperature sensitivity of soil carbon decomposition and feedbacks to climate change. *Nature*, 440, 165-73.

Ding, J., Li, F., Yang, G., Chen, L., Zhang, B., Liu, L., Fang, K., Qin, S., Chen, Y., Peng, Y., Ji, C., He, H., Smith, P. and Yang, Y. (2016) The permafrost carbon inventory on the Tibetan plateau: a new evaluation using deep sediment cores. *Global Change Biology, 22*(8).

Emmett, B.A., Reynolds, B., Chamberlain, P.M., Rowe, E., Spurgeon, D., Brittain, S.A., Frogbrook, Z., Hughes, S., Lawlor, A.J. and Poskitt, J. (2010) Countryside Survey: Soils Report from 2007. *NERC/Centre for Ecology and Hydrology*,

Fang, J., Yang, Y., Ma, W., Mohammat, A. and Shen, H. (2010) Ecosystem carbon stocks and their changes in China's grasslands. *Science China Life Sciences*, 53, 757-65.

Feng, R. and Wang, F. (1993) *Soil Species of Gansu*. Gansu Science and Technology Press, Lanzhou.

Gong, D. (1994) *Soil Species of Xizang Autonomous Region*. Science Press, Beijing.

Grimm, R., Behrens, T., Märker, M. and Elsenbeer, H. (2008) Soil organic carbon concentrations and stocks on Barro Colorado Island — Digital soil mapping using Random Forests analysis. *Geoderma*, 146, 102-113.

Hanegraaf, M.C., Hoffland, E., Kuikman, P.J. and Brussaard, L. (2009) Trends in soil organic matter contents in Dutch grasslands and maize fields on sandy soils. *European Journal of Soil Science*, 60, 213-222.

Hastie, T.J., Tibshirani, R.J. and Friedman, J.H. (2009) *The Elements Of Statistical Learning 2nd edition*. Springer.

Hou, Z. and Zeng, Q. (1992) *Soil Species of Xinjiang Uygur Autonomous Region*. Xinjiang Medical Science and Technology Press, Urumqi.

Huang, Y., Sun, W., Zhang, W. and Yu, Y. (2010) Changes in soil organic carbon of terrestrial ecosystems in China: a mini-review. *Science China Life Sciences*, 53, 766-775.

Johnson, D.W., Todd, D.E., Trettin, C.F. and Sedinger, J.S. (2007) Soil Carbon and Nitrogen Changes in Forests of Walker Branch Watershed, 1972 to 2004. *Soil Science Society of America Journal*, 71, 1639-1646.

Li, Q., Yue, T., Wang, C., Zhang, W., Yu, Y., Li, B., Yang, J. and Bai, G. (2013) Spatially distributed modeling of soil organic matter across China: An application of artificial neural network approach. *Catena*, 104, 210-218.

Liu, S. (1995) *Soil Species of Qinghai*. China Agricultural Press, Beijing.

Mestdagh, I., Sleutel, S., Lootens, P., Cleemput, O.V., Beheydt, D., Boeckx, P., Neve, S.D., Hofman, G., Camp, N.V. and Walle, I.V. (2009) Soil organic carbon stock changes in

Biogeosciences Discuss., doi:10.5194/bg-2016-473, 2017
Manuscript under review for journal Biogeosciences
Published: 2 January 2017

Flemish grassland soils from 1990 to 2000. *Journal of Plant Nutrition and Soil Science*, 172, 24–31.

Nelson, D.W., Sommers, L.E., Sparks, D.L., Page, A.L., Helmke, P.A., Loeppert, R.H., Soltanpour, P.N., Tabatabai, M.A., Johnston, C.T. and Sumner, M.E. (1982) Total carbon, organic carbon, organic matter. In: *Methods of Soil Analysis* (ed. Page AL), American Society of Agronomy, Madison.

Piao, S., Fang, J., Liu, H. and Zhu, B. (2005) NDVI-indicated decline in desertification in China in the past two decades. *Geophysical Research Letters*, 32, 347-354.

Sreenivas, K., Sujatha, G., Sudhir, K., Kiran, D.V., Fyzee, M.A., Ravisankar, T. and Dadhwal, V.K. (2014) Spatial Assessment of Soil Organic Carbon Density Through Random Forests Based Imputation. *Journal of the Indian Society of Remote Sensing*, 42, 1-11.

R Development Core Team. (2014) R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna.

Thomey, M.L., Collins, S.L., Vargas, R., Johnson, J.E., Brown, R.F., Natvig, D.O. and Friggens, M.T. (2011) Effect of precipitation variability on net primary production and soil respiration in a Chihuahuan Desert grassland. *Global Change Biology*, 17, 1505-1515.

Tucker, C.J., Pinzon, J.E., Brown, M.E., Slayback, D.A., Pak, E.W., Mahoney, R., Vermote, E.F. and Saleous, N.E. (2005) An extended AVHRR 8-km NDVI dataset compatible with MODIS and SPOT vegetation NDVI data. *International Journal of Remote Sensing*, 26, 4485-4498.

Wu, H., Guo, Z. and Peng, C. (2003) Distribution and storage of soil organic carbon in China. *Global Biogeochemical Cycles*, 17, 67-80.

Biogeosciences Discuss., doi:10.5194/bg-2016-473, 2017
Manuscript under review for journal Biogeosciences
Published: 2 January 2017

Wu, Y. (1991) *Soil Species of Ningxia Hui Autonomous Region*. Ningxia People's Press, Yinchuan.

Xie, Z., Zhu, J., Liu, G., Cadisch, G., Hasegawa, T., Chen, C., Sun, H., Tang, H. and Zeng, Q. (2007) Soil organic carbon stocks in China and changes from 1980s to 2000s. *Global Change Biology*, 13, 1989-2007.

Yan, L., Zhou, G., Wang, Y., Hu, T. and Sui, X. (2015) The spatial and temporal dynamics of carbon budget in the alpine grasslands on the Qinghai-Tibetan Plateau using the Terrestrial Ecosystem Model. *Journal of Cleaner Production*, 108, 189-196.

Yang, Y., Fang, J., Smith, P., Tang, Y., Chen, A., Ji. C., Hu, H., Rao, S.,Tan, K. and He, J.(2009). Changes in topsoil carbon stock in the Tibetan grasslands between the 1980s and 2004. *Global Change Biology*, *15*(11), 2723-2729.

Yang, Y., Fang, J., Tang, Y., Ji, C., Zheng, C., He, J. and Zhu, B. (2008) Storage, patterns and controls of soil organic carbon in the Tibetan grasslands. *Global Change Biology*, 14, 1592-1599.

Yang, Y., Fang, J., Ma, W., Smith, P., Mohammat, A., Wang, S. and Wang, W. (2010) Soil carbon stock and its changes in northern China's grasslands from 1980s to 2000s. *Global Change Biology*, 16, 3036–3047.

Yang, Y., Li, P., Ding, J., Zhao, X., Ma, W., Ji, C. and Fang, J. (2014) Increased topsoil carbon stock across China's forests. *Global Change Biology*, 20, 2687-2696.

Yue, H., Wang, M., Wang, S., Gilbert, J.A., Sun, X., Wu, L., Lin, Q., Hu, Y., Li, X., He, Z., Zhou, J. and Yang, Y. (2015) The microbe-mediated mechanisms affecting topsoil carbon stock in Tibetan grasslands. *ISME Journal*, 9, 2012-2020.

Biogeosciences Discuss., doi:10.5194/bg-2016-473, 2017
Manuscript under review for journal Biogeosciences
Published: 2 January 2017

Zhang, X., Shen, Z. and Fu, G. (2015) A meta-analysis of the effects of experimental warming on soil carbon and nitrogen dynamics on the Tibetan Plateau. *Applied Soil Ecology*, 87, 32-38.

Zhang, Y., Tang, Y., Jiang, J. and Yang, Y. (2007) Characterizing the dynamics of soil organic carbon in grasslands on the Qinghai-Tibetan Plateau. *Science in China*, 50, 113-120.

Zhao, J. (1993) *Soil Species of Neimeng Autonomous Region*. China Agricultural Press, Beijing.

Zogg, G.P., Zak, D.R., Ringelberg, D.B., White, D.C., MacDonald, N.W. and Pregitzer, K.S. (1997) Compositional and Functional Shifts in Microbial Communities Due to Soil Warming. *Soil Science Society of America Journal*, 61, 475-481.

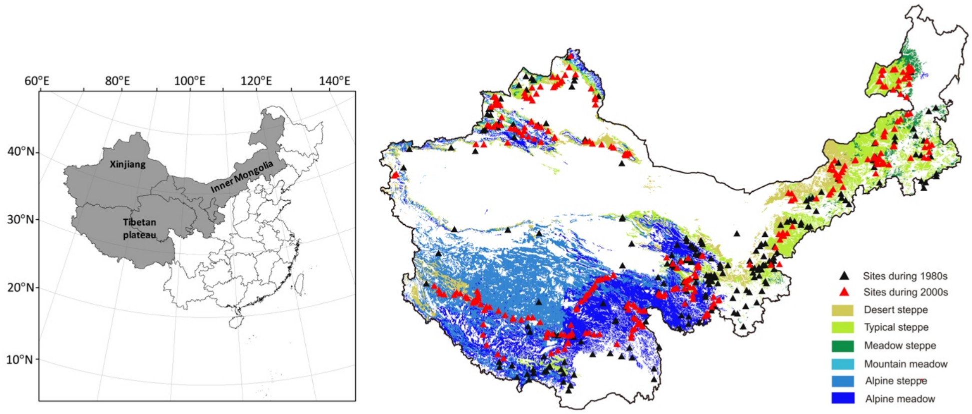

Figure 1: Location of the study area and sampling sites during the 1980s and 2000s shown the background of China's vegetation map (Chinese Academy of Sciences, 2001).

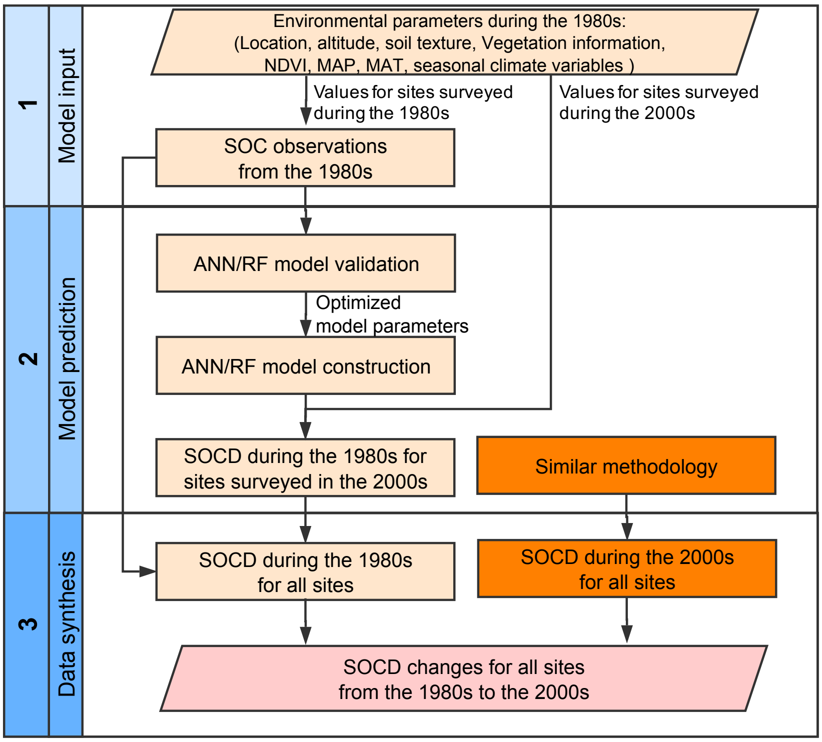

Figure 2: Methodology used for observing SOCD changes during the 1980s and 2000s. SOCD: soil organic carbon density, NDVI: Normalized Difference Vegetation Index, MAT: mean annual temperature, MAP: mean annual precipitation, ANN: artificial neural network, and RF: random forest.

Biogeosciences Discuss., doi:10.5194/bg-2016-473, 2017
Manuscript under review for journal Biogeosciences
Published: 2 January 2017

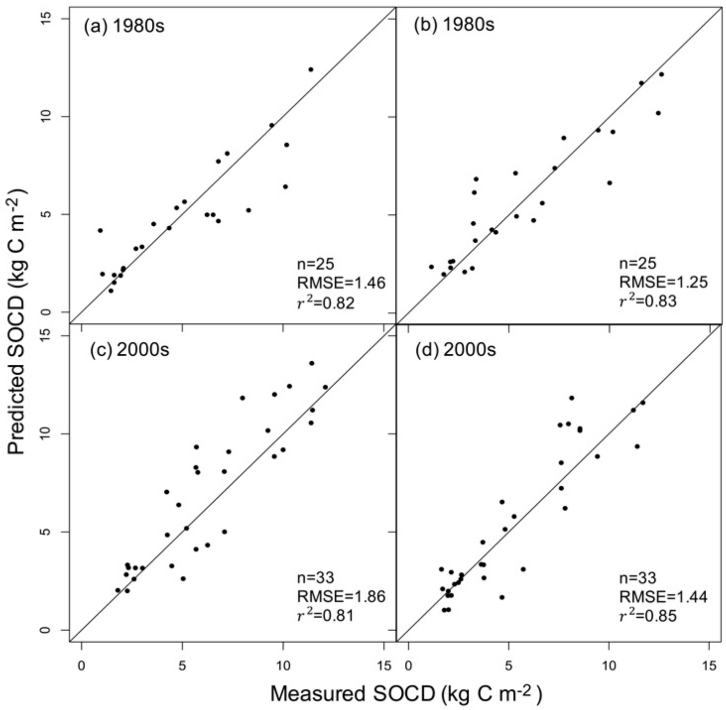

Figure 3: Comparison between measured and predicted SOCD simulated using the ANN (a and c) and RF approaches (b and d) during the test process. The diagonal is the 1:1 line. SOCD: soil organic carbon density, ANN: artificial neural network, RF: random forest, r^2: coefficient of determination, RMSE: root mean square error.

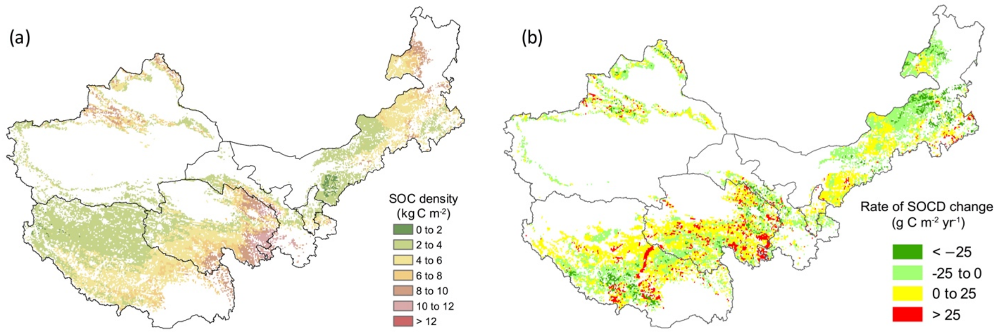

Figure 4: Spatial distributions of the SOCD in the top 30 cm for the 2000s (a) and its temporal changes (b) in the grasslands across northern China from the 1980s to the 2000s. SOCD: soil organic carbon density.

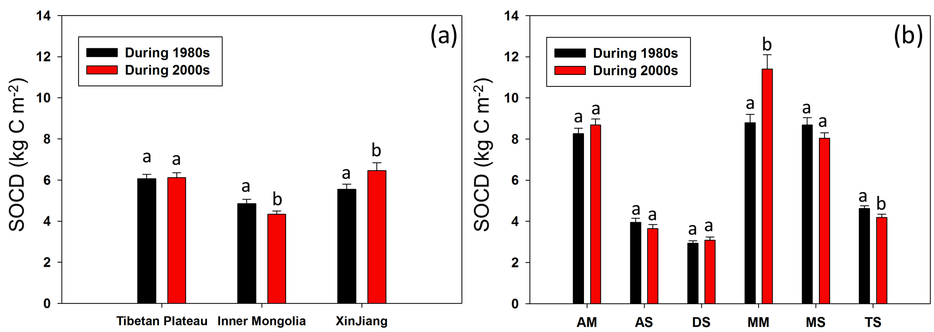

Figure 5: Comparison of SOCD in the top 30 cm during the 1980s and 2000s across different regions and grassland types. The error bar represents the standard error. SOCD: soil organic carbon density, AM: alpine meadow, AS: alpine steppe, DS: desert steppe, MM: mountain meadow, MS: mountain steppe, TS: typical steppe.

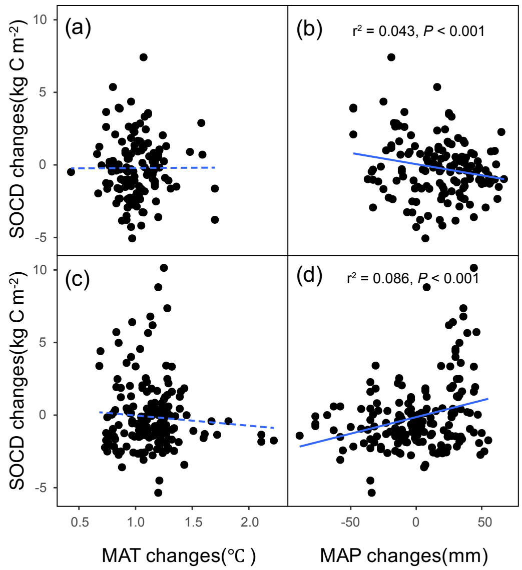

Figure 6: Relationships between the SOCD change values and the MAT / MAP change values across the alpine (a and b) and temperate (c and d) grasslands from the 1980s to the 2000s. SOCD: soil organic carbon density, MAT: mean annual temperature, and MAP: mean annual precipitation.

Biogeosciences Discuss., doi:10.5194/bg-2016-473, 2017
Manuscript under review for journal Biogeosciences
Published: 2 January 2017

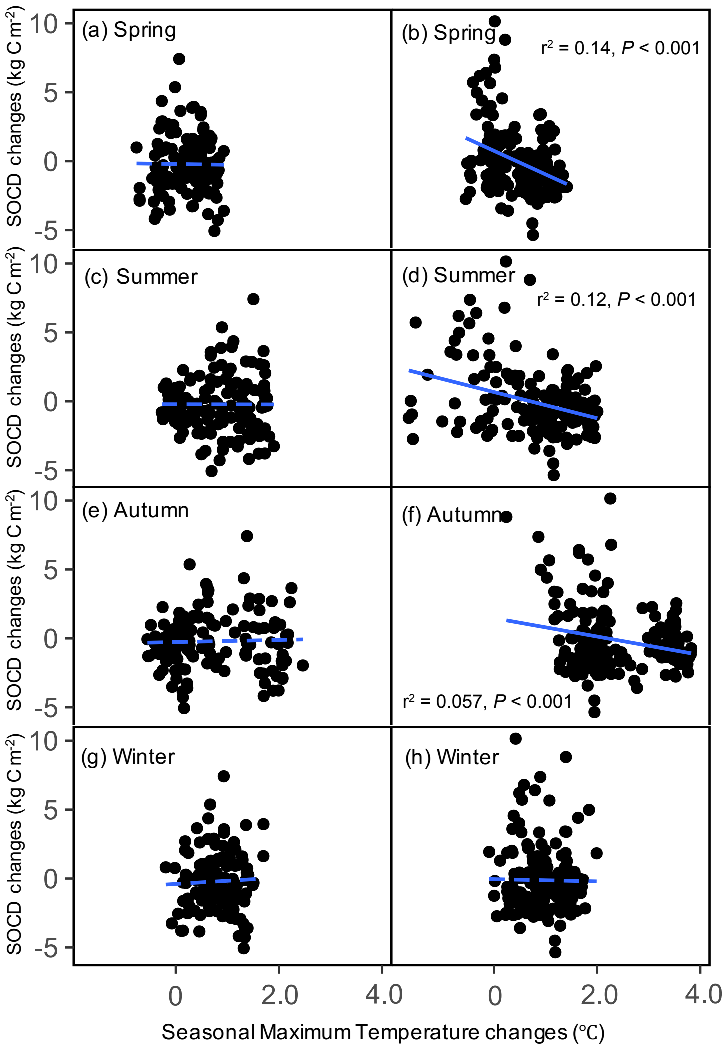

Figure7: Relationships between the SOCD change values and the seasonal maximum temperature change values across the alpine (a, c, e and g) and temperate (b, d, f and h) grasslands from the 1980s to the 2000s. SOCD: soil organic carbon density.

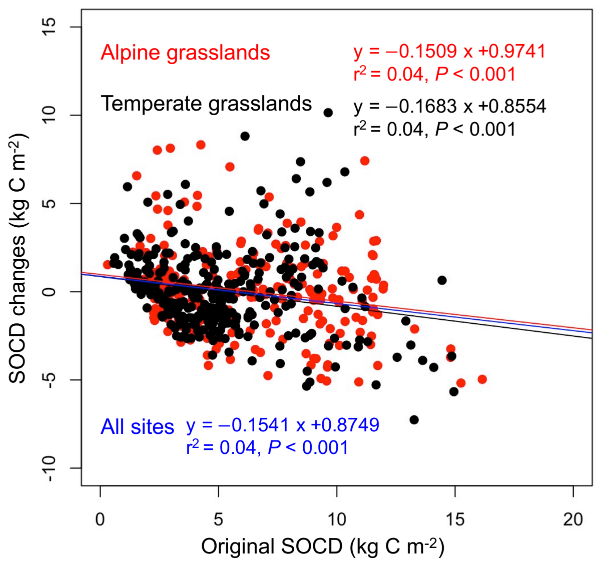

Figure 8: Relationships between the SOCD changes and the original value in the top 30 cm in grasslands across northern China from the 1980s to the 2000s. The lines are the regression lines for the whole grassland region and the two grassland types. SOCD: soil organic carbon density.

Biogeosciences Discuss., doi:10.5194/bg-2016-473, 2017
Manuscript under review for journal Biogeosciences
Published: 2 January 2017

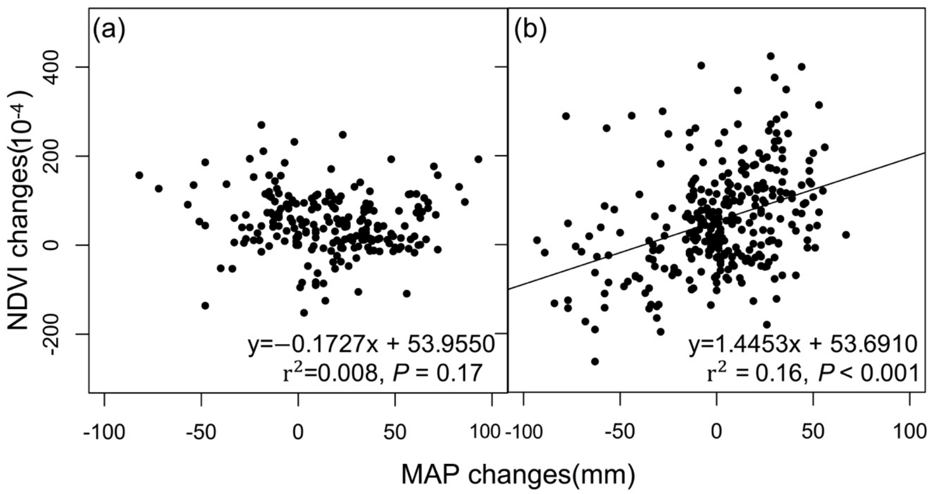

Figure 9: Relationships between NDVI and MAP changes in Chinese grasslands from the 1980s to the 2000s. (a) alpine grasslands, and (b) temperate grasslands. NDVI: Normalized Difference Vegetation Index, MAP: mean annual precipitation.

Biogeosciences Discuss., doi:10.5194/bg-2016-473, 2017
Manuscript under review for journal Biogeosciences
Published: 2 January 2017

Table 1: Mean soil organic carbon density (SOCD) in the top 30 cm of the grasslands at the site level. The predicted SOCD was estimated using the artificial neural network (ANN) and random forest (RF) algorithms. The bold figures indicated the estimated SOCD from field measurements. SE: standard error of the mean value.

Category	No.	Mean SOCD/SE (kg C m^{-2})			
		1980s		2000s	
		ANN	RF	ANN	RF
Sites surveyed in the 1980s	246	**5.70/0.24**		5.70/0.21	5.97/0.19
Sites surveyed in the 2000s	327	5.21/0.23	5.39/0.14	**5.23/0.21**	
All sites	573	5.42/0.22	5.55/0.13	5.43/0.18	5.53/0.14

Biogeosciences Discuss., doi:10.5194/bg-2016-473, 2017
Manuscript under review for journal Biogeosciences
Published: 2 January 2017

Table 2: Changes in topsoil (0-30 cm) SOC in grasslands across northern China from the 1980s to the 2000s. Estimates of area-weighted SOC density and stocks were based on actual measured profiles and the predicted SOCD simulated using the random forest (RF) method. SOCD: soil organic carbon density, AS, alpine steppe; AM, alpine meadow; DS, desert steppe; TS, typical steppe; MS, mountain steppe; MM, mountain meadow; and SE, standard error of the mean value.

Region	Grassland type	Area /10 ^{4} km	No. Sites sampled	Area-weighted mean SOCD/SE (kg C m $^{-2}$)		Area-weighted SOC stock/SE (Pg C)		Rate of SOCD change (g C m $^{-2}$ yr $^{-1}$)
				1980s	2000s	1980s	2000s	
Inner Mongolia	DS	9.59	60	2.48/0.13	2.68/0.17	0.24/0.01	0.26/0.02	11.23
	TS	28.65	105	4.54/0.15	4.01/0.15	1.30/0.04	1.15/0.04	-29.44
	MS	5.83	40	9.28/0.42	7.68/0.31	0.54/0.02	0.45/0.02	-88.92
	Total	**44.07**	**205**	**4.72**	**4.20**	**2.08**	**1.85**	**-28.46**
Tibetan Plateau	AS	72.29	128	3.95/0.20	3.65/0.20	2.85/0.15	2.64/0.15	-16.68
	AM	53.47	123	8.26/0.27	8.69/0.29	4.42/0.14	4.65/0.15	24.03
	Total	**125.76**	**251**	**5.78**	**5.79**	**7.27**	**7.29**	**0.63**
Xinjiang	DS	6.15	51	3.48/0.16	3.56/0.26	0.21/0.01	0.22/0.02	4.16
	TS	5.23	22	5.04/0.22	5.09/0.58	0.26/0.01	0.27/0.03	2.70
	MS	7.77	16	7.17/0.61	8.93/0.46	0.56/0.05	0.69/0.04	97.93
	MM	7.36	28	8.79/0.41	11.39/0.70	0.65/0.03	0.84/0.05	144.65
	Total	**26.51**	**119**	**6.34**	**7.61**	**1.68**	**2.02**	**70.36**
Northern China	**Overall**	**196.34**	**573**	**5.61**	**5.68**	**11.03**	**11.16**	**3.68**

Biogeosciences Discuss., doi:10.5194/bg-2016-473, 2017
Manuscript under review for journal Biogeosciences
Published: 2 January 2017

Biogeosciences Discussions
Open Access

Table 3: Estimate and relative uncertainty of SOCD change of northern China grassland from 1980s to 2000s. Estimated median SOCD change with interquartile range was generated by using 5000-iteration Monte Claro resampling. SOCD: soil organic carbon density, AS, alpine steppe; AM, alpine meadow; DS, desert steppe; TS, typical steppe; MS, mountain steppe; MM, mountain meadow.

Grassland type	No. of sites	SOCD change(kg C m^{-2})		
		Median	25 Percentile	75 Percentile
Alpine grasslands	251	0.14	-1.24	1.51
AS	128	-0.48	-1.85	0.90
AM	123	0.78	-0.59	2.15
Temperate grasslands	322	0.01	-1.25	1.29
DS	111	0.12	-1.12	1.41
MM	28	2.56	1.36	3.92
MS	55	-0.65	-1.84	0.69
TS	126	-0.40	-1.68	0.86
Total	573	-0.12	-1.42	1.17

Biogeosciences Discuss., doi:10.5194/bg-2016-473, 2017
Manuscript under review for journal Biogeosciences
Published: 2 January 2017

Tables 4: Previous estimations of carbon storage in the 2000s and its balance in Chinese grassland soil at different spatial scales using different approaches over a comparable period.

Region	Method	Period	Area (10^6 ha)	Soil depth (cm)	Carbon storage in the 2000s (Pg)	Carbon balance (Tg C yr^{-1})	Ref.
Chinese grasslands	Literature synthesis	1980s-2000s	278.51	103.2	34.15	-178.2	Xie *et al*. (2007)
	Literature synthesis	1981-2000	400	NA	NA	4.9±1.6	Huang *et al*. (2010)
northern Chinese grasslands	Inventory Kriging-based estimation	1980s-2000s	196.34	30	10.47	No significant change	Yang *et al*. (2010)
	Inventory RF-based Estimation	1980s-2000s	196.34	30	11.16	No significant change	This study
Tibetan alpine grasslands	Process-based models	1961-2010	100	NA	6.28-6.60	6	Yan *et al*. (2015)
	Process-based models	1990-2002	147.74	20	9.72	Decrease	Zhang *et al*. (2007)
	Inventory satellite-based estimation	1980s-2000s	112.82	30	4.99	-0.45±40.75	Yang *et al*. (2009)
	Inventory RF-based Estimation	1980s-2000s	125.76	30	7.29	No significant change	This study
Inner Mongolian grasslands	Inventory satellite-based estimation	1980s-2010s	58.77	20	NA	No significant change	Dai *et al*. (2014)
	Inventory RF-based Estimation	1980s-2000s	44.07	30	1.85	-12.78	This study
northwest Chinese (Xinjiang) grassland	Inventory RF-based Estimation	1980s-2000s	26.51	30	2.20	18.89	This study