1 *Article*

2 A Conversion Method to Determine Regional
3 Vegetation Cover Factor from Standard Plots based
4 on Large Sample Theory and TM Images: A Case
5 Study in the Eastern Farming-pasture Ecotone of
6 Northern China

7 **Degen Lin[1,2,3,4,8], Yuan Gao[1,3,4], Yaoyao Wu[1,3,4], Peijun Shi[1,3,4,6,7]*, Huiming Yang[1,3,4,5] and Jing'ai**
8 **Wang[1,3,4]**

9 [1] State Key Laboratory of Earth Surface Processes and Resource Ecology, Beijing Normal University, Beijing
10   100875, China; ldg@zjnu.edu.cn
11 [2] College of Geography and Environmental Sciences, Zhejiang Normal University, Zhejiang 321004, China;
12 [3] Key Laboratory of Regional Geography, Beijing Normal University, Beijing 100875, China;
13 [4] Faculty of Geographical Science, Beijing Normal University, Beijing 100875, China;
14 [5] College of Life Science & Technology, Hebei Normal University of Science & Technology, Qinhuangdao
15   066000, China;
16 [6] Key Laboratory of Environmental Change and Natural Disaster, MOE, Beijing Normal University, Beijing
17   100875, China;
18 [7] Academy of Disaster Reduction and Emergency Management, Beijing Normal University, Beijing 100875,
19   China
20 [8] Remote Sensing Application & Environmental Disaster Research Center, Zhejiang Normal University,
21   Zhejiang 321004, China;
22 * Correspondence: spj@bnu.edu.cn; Tel.: +86-010-58808179

23 **Abstract:** The key to simulating soil erosion is to calculate the vegetation cover (C) factor. Methods
24 that apply remote sensing to calculate C factor at regional scale cannot directly use the C factor
25 formula. That is because the C factor formula is obtained by experiment, and needs the coverage
26 ratio data of croplands, woodlands and grasslands at standard plot scale. In this paper, we present
27 a C factor conversion method from a standard plot to a km-sized grid based on large sample theory
28 and multi-scale remote sensing. Results show that: 1) Compared with the existing C factor formula,
29 our method is based on the coverage ratio of croplands, woodlands and grasslands on a km-sized
30 grid, takes the C factor formula obtained from the standard plot experiment and applies it to
31 regional scale. This method improves the applicability of the C factor formula, and can satisfy the
32 need to simulate soil erosion in large areas. 2) The vegetation coverage obtained by remote sensing
33 interpretation is significantly consistent (paired samples t-test, $t = -0.03$, $df = 0.12$, 2-tail significance
34 $p < 0.05$) and significantly correlated with the measured vegetation coverage. 3) The C factor of the
35 study area is smaller in the middle, southern and northern regions, and larger in the eastern and
36 western regions. The main reason for that is the distribution of woodlands, the Hunshandake and
37 Horqin sandy lands and the valleys affected by human activities. 4) The method presented in this
38 paper is more meticulous than the C factor method based on the vegetation index, improves the
39 applicability of the C factor formula, and can be used to simulate soil erosion on large scale and
40 provide strong support for regional soil and water conservation planning.

41 **Keywords:** farming-pasture ecotone; TM image, remote sensing, vegetation cover factor; scale
42 conversion; land use; high resolution image

43

44 1. Introduction

45   The soil erosion models have provided strong support for regional soil and water conservation
46   planning. With the development of Geographic Information System (GIS) and Remote sensing (RS),
47   the water erosion equation has been widely used in regional soil erosion simulations [1-6]. Current
48   water erosion models include the universal soil loss equation (USLE)[7], the revised universal soil
49   loss equation (RUSLE)[8,9], the China soil loss equation [10] and so on. These models are being widely
50   used to estimate soil loss in agriculture and environmental management. In the models, soil loss
51   A=R×K×L×S×C×P, where R, K, L, S, C and P, respectively, are rainfall erosivity (R), soil erodibility (K),
52   slope length (L), slope steepness (S), vegetation cover and management (C), and support practice
53   factor (P). The R, K, L and S factors are controlled by the natural environment, therefore will not be
54   changed by short-term soil and water conservation measures and activities. However, the C factor
55   has the greatest change range among the factors of water erosion models, and the changes can differ
56   by 2–3 orders of magnitude[11]. According to Benkobi, *et al.* [12] and Biesemans, *et al.* [13], the
57   vegetation cover factor together with slope steepness and length factors are most sensitive for soil
58   loss, and have the most significant effect on the overall effectiveness of the USLE / RUSLE models
59   [14]. Therefore, calculating and improving the accuracy of regional C factors has become the key to
60   improving regional soil erosion simulations.

61   Most of the regional C factors have been determined with remote sensing data through the
62   following methods: 1) The direct assignment of land use/coverage [15]. This method is simple, but
63   the accuracy of the computed C factor is poor [11]; 2) The vegetation index estimated C factor method
64   [16-18]. This method can express the regional vegetation coverage more finely, but it is less
65   comprehensive, and multiple layers and shallow roots are usually ignored; 3) The spectral mixing
66   analysis (SMA) estimated C factor method [19,20]. The SMA method considers the contributions of
67   litter, gravel, etc.; it can fully reflect the C factor information independent of the measurements, and
68   the soil background does not affect it. However, the SMA method cannot be used when vegetation
69   and/or litter completely covers the surface, or when the data is affected by multiple scattering [11]; 4)
70   Experimental approaches combined with geostatistical methods [18,21]. With this method, the C
71   factor can be interpolated using GIS and remote sensing images as auxiliary variables. Wang, *et al.*
72   [18] improved mapping the C factor for the USLE by geostatistical methods with TM images. Based
73   on multiple primary variables (canopy cover, ground cover and vegetation height), Gertner, *et al.* [21]
74   mapped the C factor in regions from joint co-simulation.

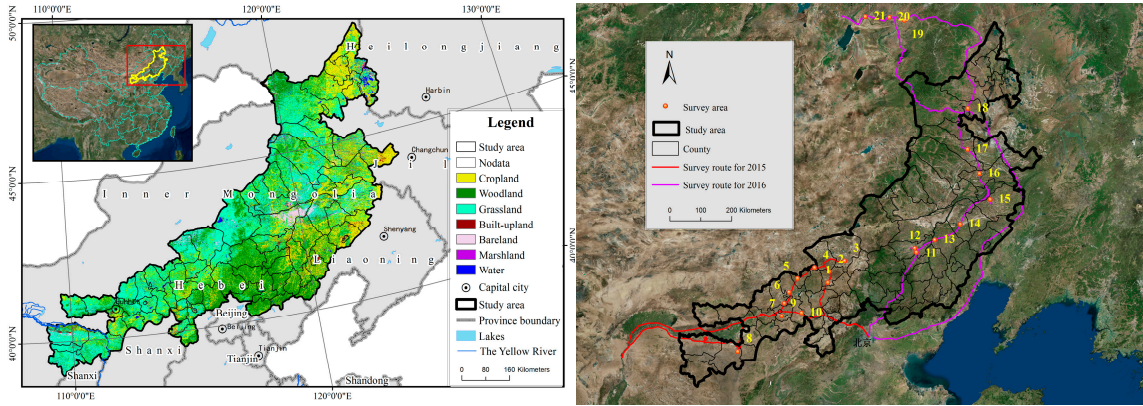
75   All of these methods that apply remote sensing have the following problems: Firstly,
76   the methods above used the C factor formula obtained from standard plot experiments, while such
77   formula cannot be directly applied to an entire region. The C factor formula of croplands, woodlands
78   and grasslands, is calculated by experiments on the erosion rate of croplands, woodlands and
79   grasslands plots with bare land in the standard plots. However, the vegetation cover of a region is a
80   complex combination of croplands, woodlands and grasslands, differing from the vegetation cover
81   modelled in the standard plot. Secondly, the RSI method requires a large number of field samples,
82   and could not be used in the fragmented landscape areas, such as farming-pasture ecotones of
83   northern China. Not only is it difficult to ensure the accuracy of the spatial interpolation, but it is also
84   time-consuming, laborious and difficult to promote. Thirdly, there are very few ways to properly
85   consider the effects of surface coverage and canopy coverage on soil erosion in the modelled region.

86   Based on the results of previous regional C factor estimations, we analyzed the key factors of soil
87   erosion and highlighted the key factors affecting the scale conversion [22]. In order to improve the

88   accuracy of regional C factor estimation, and obtain large-scale C factor map for macro-scale soil
89   erosion simulation, we built a C factor estimation method based on large sample theory and Landsat
90   Thematic Mapper (TM) images, and we show that our method solves the key problem of transitioning
91   from a standard plot to km-sized grids, and hence accurately estimates regional C factors.

92   **2. Study Area**

93   The eastern section of the farming-pasture ecotone (hereinafter referred to as the study area (see
94   Figure 1) includes 84 counties (banners) in the Inner Mongolia Autonomous Region, Liaoning
95   Province, Heilongjiang Province, Jilin Province and Hebei Province, and has an area of 4.402×10^5
96   km². In 2015–2016, our research group carried out two field trips to determine the land use and soil
97   erosion in the study area, and the length of survey route was nearly 7000 km. According to the typical
98   geographic unit and landform type, we set up 21 survey sample areas. From July 19 to July 25, 2015,
99   the western part of the study area was inspected, where 10 inspection points were created over a
100   route of 2930 km. From August 7 to August 14, 2016, the eastern part of the study area was inspected,
101   where 11 inspection points were created over a route of more than 4,000 km.

102

103   Figure 1 Study area and two field trips routes

104   **3. Materials and Methods**

105   *3.1 Basic Idea and Research Framework*

106   More than 1000 coverage data-points were obtained from Landsat TM images of Global Land
107   Survey in 2010 (GLS2010) and ground measurements, and according to large sample theory and
108   information entropy theory, their distributions and proportions are similar. Large sample theory, also
109   called asymptotic theory, is used to approximate the distribution of an estimator when the sample
110   size *n* is large. This theory is extremely useful if the exact sampling distribution of the estimator is
111   complicated or unknown[23]. In addition, based on information-entropy theory [24], when the
112   sample size is sufficiently large, the samples can be assumed approximate to a normal distribution.
113   According to the above two theories, the subsets of remote sensing data-points and field
114   measurements data-points have a similar distribution. That is, the same km-sized grid, 2,000 land
115   cover data-points obtained from field measurements have a similar distribution and similar
116   proportion of croplands, woodlands, and grasslands as 1,000 land cover data-points determined from
117   the Landsat TM images. The accuracy of the field measured vegetation cover method is better than
118   that of the remote sensing method, therefore the field measurements can be used as the reference for
119   verifying the remote sensing measurements. The research framework is shown in Figure 2.

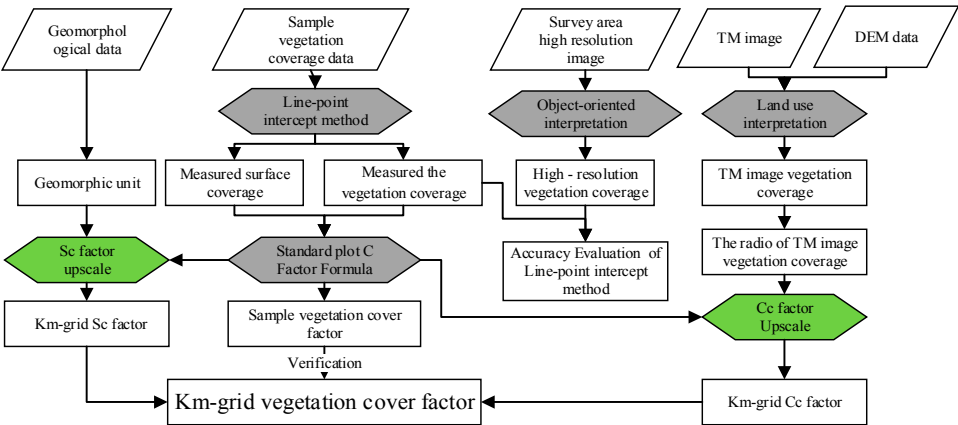



Figure 2 The C factor conversion method from a "standard plot (single vegetation type)" to a "km-sized grid" (multiple vegetation types)

The C factor conversion from a "standard plot (single vegetation type)" to a "kilometer grid" (multiple vegetation types) contains six steps: 1) The vegetation cover of the sampling method was evaluated based on the high-resolution images of the survey sample area. 2) We interpreted the Landsat TM images of the study area and derived the land use data based on the CART decision tree classification method. 3) Then set up 21 survey sample areas, each with an area of 1 km², and 2,000 canopy cover and surface cover data established. 4) Based on the resolution of land use data, 1,125 canopy coverage data were obtained in each km-sized grid of the study area, and the canopy coverage (Cc) of the croplands, woodlands, and grasslands were calculated. 5) The surface cover (Sc) factor was calculated based on the surface coverage surveyed in survey sample area, and was applied to the entire study area according to the landform type. 6) According to the Cc and Sc factors, we calculated in the C factors of study area. Finally, we verified the regional C factor by the survey sample areas.

3.2 Materials

The study mainly used three types of data: measured data (survey sample data), remote sensing data (Landsat TM image data, MODIS data and high-resolution images) and basic geographical data. We used the Landsat TM image data of the Global Land Survey in 2010 (GLS2010, https://glovis.usgs.gov/, 30m×30m). The dates of MODIS data (https://urs.earthdata.nasa.gov/profile, 1km×1km) are July 12, 2015 and July 27, 2016, near the survey dates. A total of about 42,000 coverage data-points were surveyed throughout the study area. Basic information of the survey sample areas is shown in Table 1.

Table 1 The basic information of the survey sample area

Sample ID	longitude	latitude	County	Measuring point	Geomorphic unit *
1	E115°6'55''	N41°40'57''	Taibus Banner	1983	Wavy High Plains
2	E115°14'3''	N42°0'40''	Taibus Banner	2012	Wavy High Plains
3	E115°54'25''	N42°16'53''	Zhenglan Banner	2002	Rising hills
4	E114°39'9''	N42°11'33''	Hua De County	1936	Layered high plain
5	E113°54'48''	N41°55'2''	Hua De County	1953	Wavy High Plains
6	E113°30'56''	N41°32'14''	Shangdu County	1996	Wavy High Plains
7	E113°16'26''	N41°13'44''	Qahar Right Wing Rear Banner	1989	Layered high plain
8	E111°16'16''	N39°55'19''	Jungar Banner	1985	Rising hills
9	E113°7'22''	N40°51'6''	Qahar Right Wing Front Banner	2008	Layered high plain
10	E113°55'3''	N40°53'12''	Xinghe County	1995	Rising hills
11	E118°50'4''	N42°11'30''	Hongshan District	2006	Rising hills
12	E118°48'29''	N42°19'21''	Hongshan District	2000	Rising hills

13	E119°39'56"	N42°27'48"	Aohan Banner	1990	Rising hills
14	E120°49'5"	N42°47'45"	Naiman banner	1999	Sandy land
15	E122°13'59"	N43°19'28"	Horqin Left Wing Rear Banner	1985	Sandy land
16	E122°2'13"	N44°9'47"	Horqin Left Wing Middle Banner	2000	Sandy land
17	E121°45'36"	N44°56'28"	Horqin Right Wing Middle Banner	1988	Sandy land
18	E122°9'22"	N46°9'38"	Horqin Right Wing Front Banner	2015	Sandy land
19	E120°9'48"	N49°9'29"	Evenki Autonomous Banner	2000	Valley plain
20	E119°30'58"	N49°19'55"	Chen Barag Banner	2008	Layered high plain
21	E118°25'21"	N49°28'49"	Chen Barag Banner	2008	Wavy High Plains

144   Note: * According to the Inner Mongolia landform type map to determine the landform type.

145   We used some Gaofen-2 satellite images (GF-2) that covered the investigation area to discuss the

146   relationship between survey estimated vegetation coverage and intercept vegetation coverage in the

147   km-sized grid. GF-2 satellite was designed and developed by China Academy of Space Technology

148   (CAST). It employs the CAST-CS-L3000A bus and two Panchromatic image/Multi spectral image

149   (PAN/MS) cameras: one is the MS image with four bands in the visible and near-infrared (VNIR)

150   range with a spatial resolution of 3.2 m; and the other is PAN image in the visible range with a spatial

151   resolution of 0.8 m[25]. We chose four GF-2 images that were obtained within one month either side

152   of the survey date in 2015. The GF-2 date of sample NO. 4 is August 4, 2015 and the sample NO. 5 is

153   August 9, 2015, and the sample NO. 9 is in August 19, 2015, and the sample NO. 10 is in August 9,

154   2015. With the eCognition software, we adopted the object-oriented high-precision remote sensing

155   interpretation method to obtain high-resolution land use/cover data of the survey sample area.

156   *3.3   Methods*

157   3.3.1   Canopy coverage factor upscaling

158   The canopy coverage factor can be calculated from vegetation cover, which can be estimated

159   using point intercept, line-point intercept, grid-point intercept, and ocular estimates [26,27]. Scholars

160   have compared the above-mentioned methods to calculate the vegetation cover, and found that when

161   the samples are less than 20, the point-based methods are more precise than ocular estimates

162   [26,28,29] and line-point intercepts [26]. When the sample size is large enough, the estimate of line-

163   point intercept is the same as that of the point intercept and grid-point intercept, and can correctly

164   reflect vegetation cover [26,27]. In this paper, 21 survey sample areas were set up in the study area.

165   In each survey sample area we collected a total of 2,000 samples, which is far greater than the

166   requirement of 20 samples [27].

167   The size of the sampling unit is also important for ground measurements. Duncan, *et al.* [30]

168   analyzed the influence of sampling unit size on the remote sensing regression model based on the

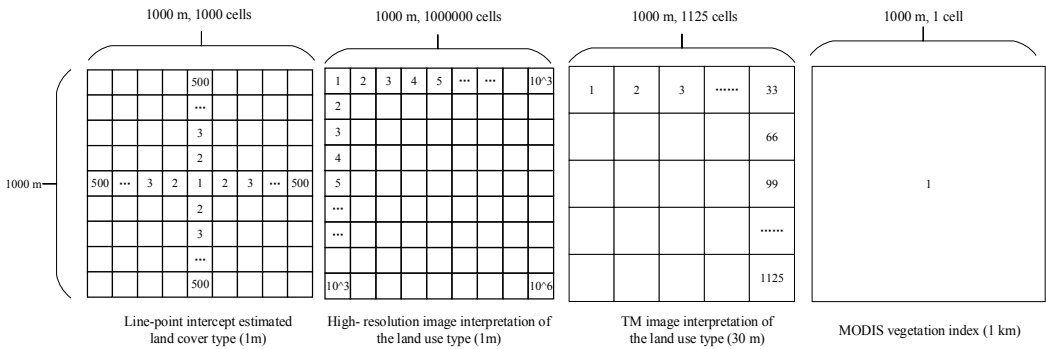
169   vegetation index and vegetation cover, and discussed the most suitable sampling unit size. In this

170   paper, we chose the km-sized grid as the statistical unit. Then, via TM image interpretation, we

171   produced more than 1,000 land use/land cover data-points in each km-sized grid. Next, the MODIS

172   vegetation cover was directly measured from the vegetation index in the km-sized grid. The different

173   methods used to estimate vegetation cover are shown in Figure 3.

174

175　　　Figure 3 The different methods used to estimate vegetation cover in the survey sample area: a: The
176　　　line-point intercept method; b: from high-resolution satellite images; c: from TM images; d: from the
177　　　MODIS vegetation index.

178　　　The key to estimating the C factor is to calculate the vegetation cover. There is no standard
179　method to monitor vegetation cover, and current methods can be divided into surface measurement
180　methods and remote sensing methods. The surface measurement method is limited by the workload
181　and the measurement area size, and is unsuitable as an independent measurement method at large
182　scale. The methods based on remote sensing to calculate the vegetation cover relies on a surface test
183　for the regression calibration. It has a certain accuracy, but it is subject to the restrictions of promotion
184　and application, especially in fragmented landscapes like farming-pasture ecotones. In terms of
185　current technology, the accuracy of the surface measurement method is higher than that of the remote
186　sensing method, and thus can be used as the basis for remote sensing measurement and data
187　verification [31].

188　3.3.2　TM Image Land Use Interpretation

189　•　Land use classification system

190　　　The coverage and proportion of various land types on a km-sized grid can be calculated
191　according to the high-resolution land use data of the study area, such as the global 30 m land use data
192　produced by Chen, *et al.* [32] or Gong, *et al.* [33]. However, both the above-mentioned classification
193　data confuse grassland and bare land in the study area, and contains no secondary classification of
194　grassland. Therefore, we adopted a secondary land use classification system for the study area based
195　on the land use/cover classification system of Liu, *et al.* [34] to classify the TM images. The
196　interpretation effects of some classification methods (such as neural network classification, object-
197　oriented classification) can be very good for single scene image, but when it comes to a large area, the
198　workload and classification efficiency must be considered. The CART algorithm is based on the
199　decision tree classification method[35], and combines the advantages of the DEM, NDVI, as well as
200　supervised and unsupervised classification, resulting in a much higher interpretation efficiency than
201　other interpretation methods[35] for large area. So, we adopted the decision tree classification method
202　of the CART algorithm to carry out land use classification (we interpreted more than 40 scene Landsat
203　TM images from GLS2010).

204　•　Decision tree classification based on the CART algorithm

205　　　Prior to classification, the Landsat TM and Landsat Enhanced Thematic Mapper (ETM) data in
206　Global Land Survey 2010 (GLS2010) underwent cloud interpretation and geometric correction and
207　were combined into multi-band images composed of blue, green, red, near infrared, short-wave
208　infrared, medium-wave infrared, long-wave radiation, together with NDVI, ISODATA, DEM and

209 other bands. The NDVI data was generated from TM images and the DEM data from global ASTGTM
210 data. The ISODATA data is from unsupervised classification of the TM images, and the minimum
211 class number of unsupervised classification is 10 and the maximum is 25.

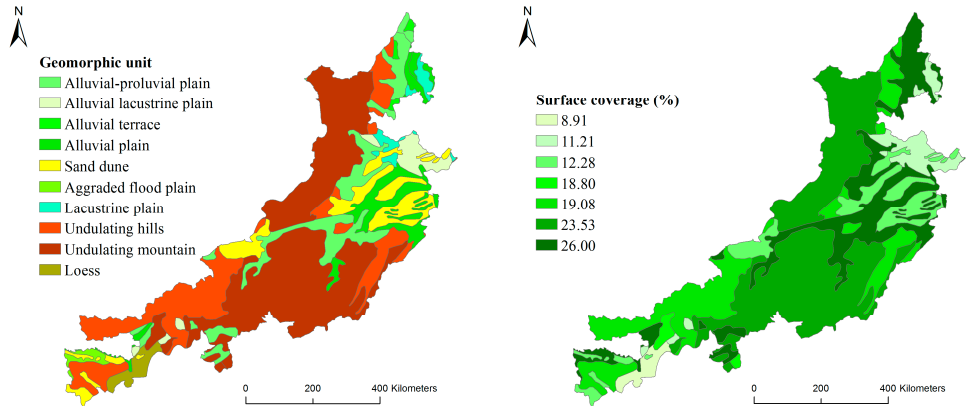
212   The main steps of decision tree classification based on the CART algorithm are: 1) Select the
213 training areas. According to the secondary classification system, a certain number of training samples
214 were selected in the multi-band image and used to obtain expert knowledge rules. The training area
215 selection order was: i) water (river canals), ii) Built-up lands, industrial and mining lands, residential
216 lands (urban land, rural residential areas, other construction land), and iii) croplands (irrigated land,
217 dry land); 2) Establish a decision tree based on the training area. We used the extension tool RuleGen
218 [35] to automatically generate decision tree rules, and used the ENVI Execute Existing Decision Tree
219 tool to establish a decision tree for land use interpretation.

220 • Classification accuracy evaluation

221   In this paper, the study area is relatively large, it is difficult to carry out scientific random field
222 verification, so this paper uses a large number of random distribution of single pixel verification point
223 method based on Google image[36,37]. We generated a total of 2,000 points randomly throughout the
224 study area (50 points in each image). In order to get the real surface cover data of the randomly points,
225 and considering the high accuracy of Google images, we first used Google Earth to distinguish the
226 real surface cover data[36,37], and second, we moved the random verification point which is in edge
227 of land cover type, to the center of land cover type, and avoid mixing pixels and ensure the accuracy
228 of the true surface data. Finally, we used the true surface data of points randomly to verify the
229 accuracy of the land use in interpretation.

230 3.3.3 Surface coverage factor upscaling

231   At regional scales, it is difficult to obtain surface cover information such as surface litter, crop
232 stubble and gravel required by soil erosion models. To overcome this limitation, we first set up 21
233 survey sample areas in the typical landform unit, measured the ratio of surface litter, crop stubble
234 and gravel in the sample area, and finally interpolated according to land surface types to obtain the
235 surface coverage factor in the study area. Secondly, the surface of the study area was classified
236 according to the geomorphic map (1:4 million) of China and its adjacent areas. Thirdly, according to
237 the natural geographic features of the study area, we found the surface of the study area can be
238 classified by the geomorphic unit that can get the surface cover information to calculate the vegetation
239 cover. Lastly, the surface coverage (Sc) of the study area was calculated based on the field coverage
240 of the survey sample area and the structure geomorphic unit of the study area, as shown in Figure 4.

241

242   Figure 4 The coverage of surface matter in the study area

243   3.3.4   Regional Vegetation Cover Factor

244   •   Standard plot C factor algorithm

245   Many scholars have experimentally determined the standard plot C factor for different regions
246   [4,11,38,39]. Different formulas are given based on different vegetation covers, for example: Jin, *et al.*
247   [40] built the C factor formula of grassland that had a 1.9% coverage in the standard plot of the
248   Huangfu River Basin. Jiang, *et al.* [41] built the C factor formula of grasslands and woodlands in a
249   standard plot with coverages of more than 5% in Ansai County of Shaanxi Province. Liu [42] built
250   the C factor formula of croplands, woodlands and grasslands in a standard plot in Beijing.

251   The surface cover is mainly composed of litter, crop stubble and gravel, which are more effective
252   in reducing soil erosion than plant canopy cover. Based on this, the C-factor algorithm of Liu [42]
253   considers both the surface cover and canopy cover of croplands, woodlands and grasslands, which
254   makes the C factor result more objective. Therefore, we chose Liu's [42] standard plot vegetation cover
255   factor algorithm to calculate the C factor, as shown in eqn. 1.

256
$$C = C_S \cdot C_C \qquad \qquad (1)$$

257   The canopy cover factor of cropland and grassland is calculated by eqn. 2, and the canopy cover
258   factor of woodland is calculated by eqn. 3. Liu [42] found that the Sc factors of croplands and
259   grasslands can be combined into a single formula by eqn. 4. The Sc factor of woodlands is calculated
260   by eqn. 5.

261
$$C_c = 1 - (0.01V_c + 0.0859)\exp(-0.0033h) \qquad \qquad (2)$$

262
$$C_c = 0.988\exp(-0.11V_c) \qquad \qquad (3)$$

263
$$C_s = 1.029 * \exp(-0.0235V_R) \qquad \qquad (4)$$

264
$$C_S = \exp(-0.0206V_R) \qquad \qquad (5)$$

265   Where C_c and C_s are the canopy cover factor and the surface cover factor, respectively, V_c and V_R
266   are the canopy cover (%) and surface cover (%), respectively, and h is the canopy height (cm). As it is
267   difficult to obtain surface information such as litter, crop stubble and gravel on the regional scale, in
268   this paper the Sc factor was calculated by using the surface coverage factor upscale method.

269   •   Regional C factor algorithm

270   The regional vegetation cover factor needs to be calculated based on the fractional cropland,
271   woodland and grassland coverage. Based on the C factor algorithm of the standard plot, the regional
272   C factor algorithm is proposed, as shown in eqn. 6.

273
$$C = (V_{crop}C_{crop} + V_{grass}C_{grass} + V_{forest}C_{forest}) / (V_{crop} + V_{grass} + V_{forest}) \qquad \qquad (6)$$

274   Where C is the C factor of the km-sized grid, V_{crop} is the cropland coverage (%), V_{grass} is the grassland
275   coverage (%), and V_{forest} is the forest coverage (%). C_{crop} is the cropland vegetation cover factor, C_{grass} is
276   the grassland vegetation cover factor, and C_{forest} is the woodland vegetation cover factor.

277   **4. Results**

278   *4.1   C factor of the survey sample area*

279   Based on land use data and surface cover data, the C factor of each survey sample area was
280   calculated according to the C factor algorithm (Table 2).

281   Table 2 The C factor of the survey sample area

Sample ID	Cropland (%)	Grassland(%)	Litter cover (%)	Fecal cover (%)	Gravel cover (%)	Wood land(%)	Bareland (%)	Road (%)	Built-up land(%)	Water (%)	Sc ratio (%)	C
1A	0.00	49.55	4.41	2.71	4.51	0.10	42.53	0.70	0.00	0.00	11.63	0.34
1B	32.76	24.14	2.13	4.46	0.81	0.00	35.50	1.01	0.00	0.00	7.40	0.55
2A	1.30	16.62	0.50	1.00	4.70	1.80	28.23	9.71	37.54	0.00	6.21	0.69
2B	27.44	34.16	4.94	1.28	6.02	0.39	30.11	6.42	0.00	0.00	12.24	0.48
3A	0.00	55.60	4.60	1.00	3.80	0.00	35.20	3.20	0.40	0.00	9.40	0.31
3B	0.00	59.38	6.39	3.89	9.88	0.00	27.74	0.00	2.59	0.00	20.16	0.22
4A	56.57	8.79	0.10	0.00	7.47	0.00	30.30	4.24	0.00	0.00	7.58	0.37
4B	20.51	28.96	1.16	0.00	20.08	1.27	42.39	5.71	0.00	0.00	21.25	0.42
5A	8.95	20.95	19.05	0.11	8.53	0.84	48.42	1.68	0.00	0.00	27.68	0.40
5B	0.00	5.48	17.05	0.80	5.68	2.09	71.19	3.39	0.00	0.00	23.53	0.50
6A	0.50	26.03	14.47	0.00	4.92	0.70	35.78	10.45	12.06	0.00	19.40	0.44
6B	0.00	36.66	12.59	0.20	7.99	4.90	43.06	2.60	0.00	0.00	20.78	0.36
7A	0.00	33.64	5.56	0.71	21.31	1.01	54.75	4.34	0.00	0.00	27.58	0.32
7B	0.00	30.03	0.00	1.00	25.13	0.00	54.75	0.90	13.31	0.00	26.13	0.35
8A	6.29	18.58	27.01	0.00	0.00	7.82	30.05	9.64	0.61	0.00	27.01	0.37
8B	7.30	25.20	14.20	0.00	1.50	3.80	18.70	2.50	28.30	0.00	15.70	0.50
9A	2.78	33.53	9.42	0.10	0.89	2.48	46.23	1.79	1.39	2.28	10.42	0.50
9B	2.80	48.60	8.80	0.30	1.50	5.00	32.80	1.70	0.00	0.00	10.60	0.38
10A	16.38	38.69	7.34	0.50	4.32	9.05	24.02	4.02	0.00	0.00	12.16	0.44
10B	12.80	35.60	1.40	0.00	8.50	8.50	22.90	2.30	7.50	9.00	9.90	0.48
11A	0.00	43.60	7.00	0.10	0.10	5.00	34.70	4.20	5.40	0.00	7.20	0.44
11B	1.19	29.03	11.93	0.10	0.20	26.64	21.77	9.34	0.00	0.00	12.23	0.28
12A	0.00	25.20	7.30	0.20	0.00	28.60	24.50	2.80	11.50	0.00	7.50	0.29
12B	15.40	24.30	6.60	0.00	2.10	35.70	11.40	3.00	3.60	0.00	8.70	0.32
13A	77.34	8.26	0.20	0.00	0.91	1.31	7.15	2.82	2.92	0.00	1.11	0.24
13B	75.83	13.24	0.00	0.00	0.00	3.91	4.81	1.91	0.30	0.00	0.00	0.30
14A	0.00	7.34	3.22	0.00	0.00	54.33	35.11	0.00	0.00	0.00	3.22	0.10
14B	0.00	42.99	6.87	0.00	0.00	31.44	18.71	0.00	0.00	0.00	6.87	0.26
15A	0.00	35.36	14.08	0.20	0.00	19.05	31.31	0.00	0.00	0.00	14.29	0.31
15B	0.00	39.78	13.03	0.20	0.00	24.85	22.14	0.00	0.00	0.00	13.23	0.27
16A	0.00	19.10	24.40	0.00	0.00	31.90	23.30	1.30	0.00	0.00	24.40	0.17
16B	0.00	38.30	3.10	0.60	0.30	0.40	52.40	0.00	0.00	5.20	4.00	0.52
17A	0.00	5.38	29.51	0.00	0.10	35.60	28.90	0.61	0.00	0.00	29.61	0.07
17B	12.57	27.05	2.79	0.00	0.10	33.23	21.96	2.40	0.00	0.00	2.89	0.38
18A	0.00	38.08	16.13	0.00	9.52	22.14	20.94	2.71	0.00	0.00	25.65	0.21
18B	19.08	18.58	3.93	0.00	0.39	22.62	31.27	4.52	0.00	0.00	4.33	0.46
19A	0.00	47.50	45.60	0.10	0.00	0.00	6.80	0.00	0.00	0.00	45.70	0.16

19B	0.00	39.10	29.60	2.30	0.80	1.70	23.00	0.80	0.00	0.00	32.70	0.26
20A	0.00	49.30	50.70	0.00	0.00	0.00	0.00	0.00	0.00	0.00	50.70	0.14
20B	0.00	47.22	43.06	0.30	0.00	4.76	1.59	0.00	0.00	2.88	43.35	0.18
21A	0.00	13.66	38.19	1.89	0.10	1.60	37.89	1.89	4.89	0.00	40.18	0.32
21B	0.00	28.16	41.79	0.70	0.10	2.29	25.37	1.69	0.00	0.00	42.59	0.25

Note: The surface coverage ratio is the sum of the litter, gravel and fecal cover.

It can be seen that the C factor of the sample area that has a higher proportion of woodland was smallest. The C factor of the survey sample areas that has a higher proportion of grassland and cropland cover was medium, and the C factor of the survey sample areas that has a higher proportion of unused land (sand and bare land) was highest.

4.2 Fractional Vegetation Cover (FVC) in the study area

Based on the 30 m resolution land use map in 2010 from GLS2010, we calculated the coverage proportion of cropland, woodland and grassland as well as the vegetation cover of the study area, as shown in Figure 5. According to the vegetation cover of the study area in 2010 based on GLS2010, the vegetation cover (including the ratio of cropland, woodland and grassland) was extracted in each survey sample area and verified by the measured vegetation coverage, as shown in Figure 6. The correlation vegetation coverage between the remote sensing and measured is shown in Table 3. The paired-samples Student's t test was used to detect the statistical significance level of the remote sensing interpretation vegetation cover. The results are shown in Table 4.

Table 3 The correlation between the remote sensing interpretation and measured vegetation coverage

	N	Correlation	Significance
Interpretation & measured Vegetation coverage	63	.760	.000

Table 4 The paired-samples Student t Test

	Paired Differences					*t*	*df*	Significance (2-tail)
	Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference				
				Lower	Upper			
Interpretation & measured coverage	−0.03	0.12	0.02	−0.06	−.002	−.03	.12	0.038

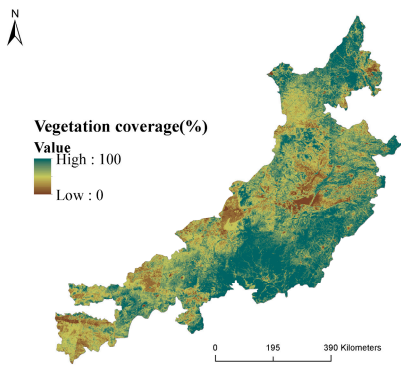



Figure 5 Vegetation coverage in the study area (2010)

$y = 0.9287x + 5.5863$
$R^2 = 0.5832$

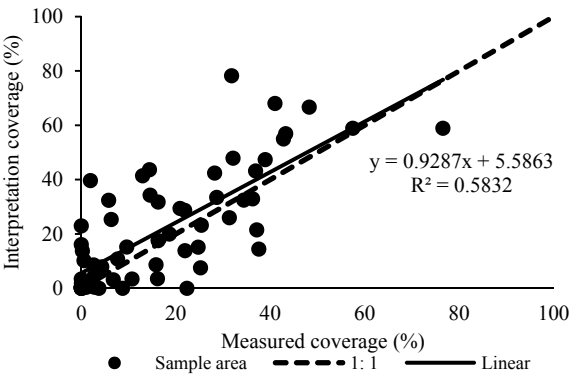



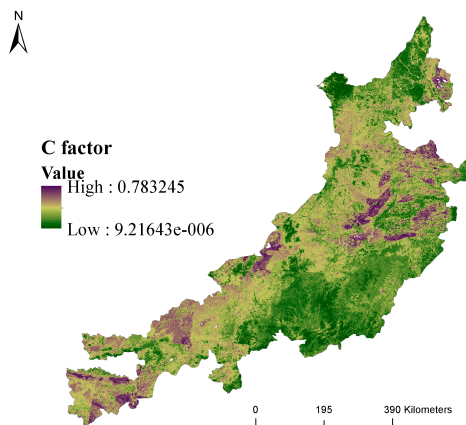
Figure 6 The verification of vegetation coverage in survey sample area

It can be seen from Figure 6, Table 3 and Table 4 that the vegetation coverage obtained by remote sensing interpretation is significantly consistent and significantly correlated with the measured vegetation coverage. The vegetation coverage obtained by remote sensing interpretation is significantly consistent with the measured vegetation coverage (paired samples *t*-test, *t* = −0.03, *df* =

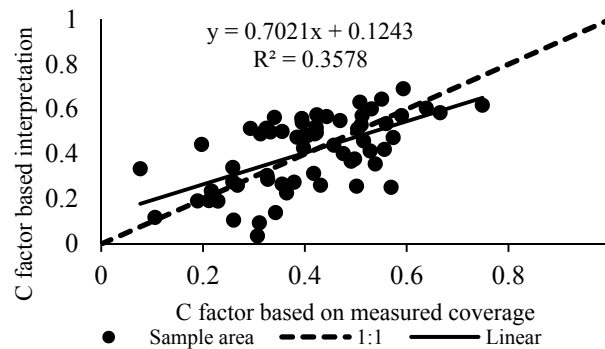
302 0.12, 2-tail significance $p < 0.05$), where R^2 is 0.58. Simultaneously, there is a statistically significant
303 linear relationship present between the two groups of data. The two groups of data are significantly
304 correlated (significance $p < 0.05$), and the correlation is 0.76. Furthermore, the two groups of data are
305 generally distributed near the 1:1 line, and the regression coefficient is 0.93. To sum up, the results
306 show that the method in this study can help accurately obtain the coverage ratio of croplands,
307 woodlands and grasslands in the km-sized grid.

308 *4.3 C factor in the Study Area*

309   Based on the C factor conversion method, and using the proportional coverage of croplands,
310 woodlands and grasslands in the km-sized grid, the C factor in the study area (2010) was calculated,
311 as shown in Figure 7.

312

313   Figure 7 The C factor in study area (2010).

314   It can be seen from Figure 7 that the C factor is smaller in the middle, southern and northern
315 regions, and larger in the eastern and western sections of the study area. In the central and northern
316 parts of the study area there are a large number of woodlands, making the C factor lower. The
317 Hunshandake sandy lands and Horqin sandy lands result in larger C factors in the eastern and
318 western regions. In the south-central part of study area, the C factor of woodlands is obviously
319 smaller than that of valleys. The main reason is that valleys are affected by human cultivation, so the
320 C factor is improved. Using the C factor based on measured coverage, the C factor of the survey
321 sample area based on remote sensing interpretation is verified (Figure 8).

322

323   Figure 8 Verifying the C factor based on remote sensing interpretation

324   It can be observed in Figure 8 that the C factor of the survey sample area calculated via remote sensing

325   is consistent with the one based on measured coverage. The two data are distributed near the 1:1 line,

326   with $R^2 = 0.36$ and correlation coefficient = 0.7. It can be low R^2 if we do not predict and 1:1 line is able

327   to judge the result of measured and interpretation. Because the resolution or interval is different

328   among the TM image and the field measured, so we get a low R^2. This shows that the C factor

329   measured in this study can reflect the vegetation cover in the study area and can be used to calculate

330   the C factor in the soil erosion equation.

331   5   Discussion

332   *5.1   Relationship between line-point estimated vegetation coverage and intercept vegetation coverage in the*

333   *km-sized grid*

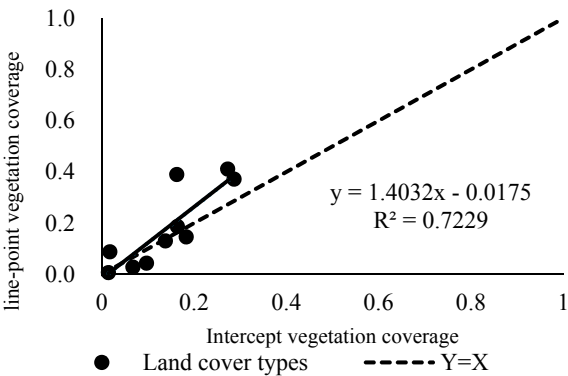
334   With the eCognition software package, we used the object-oriented classification method to

335   interpret the high-precision images (Gaofen-2 satellite images (GF-2)) of the sample area, and obtain

336   the coverage of various land cover types. The cropland, grassland and woodland coverage obtained

337   from high-precision images was compared with the line-point intercept estimated data, as shown in

338   Figure 9.

339

340   Figure 9 The correlation between the measured Fractional Vegetation Cover (FVC) from the line-point

341   intercept method and the high-resolution image interpretation.

342   It can be seen from Figure 9 that the correlation between the FVC from the line-point intercept

343   method and the high-resolution image interpretation is very high. The two groups of data are mainly

344   distributed in the vicinity of the 1:1 line. The nonparametric test of the relevant sample Wilcoxon,

345   shows that the two groups of data are significantly consistent at a significance level of $p<0.05$. The

346   value of R^2 is 0.72 and the slope is 1.40, indicating that the cropland, grassland and woodland

347   coverage obtained from the line-point intercept method is larger than that from image interpretation,

348   but can still be used to reflect the proportion of cropland, grassland and woodland in the study area,

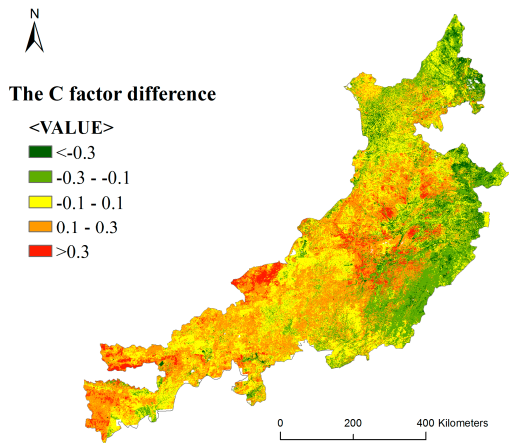
349   with an interpretation rate of 72.29%. This result is basically consistent with the findings of previous

350   studies [26,27], indicating that the line-point intercept method can effectively reflect the coverage of

351   cropland, grassland and woodland in the survey sample area.

352   *5.2   Comparison of our Method and the C factor Algorithm Based on Remote Sensing Vegetation Index*

353   We compared the C factor derived from the vegetation index (see Appendix A), with the C factor

354   determined by our method, and the difference of two C factors is shown in Figure 10.

355
356   Figure 10 The difference of C factors based on the vegetation index and remote sensing data.

357   It can be seen from Figure 7 and Figure 10 that the two C factors are basically consistent (Due to
358   the system errors of two sets data, we define that the difference between −0.1 and 0.1, which is 10%
359   of the difference, is basically consistent.) in 22% of the study area. Compared with C factor derived
360   from image interpretation, the C factor based on the vegetation index is higher (up to 0.3) for western
361   sandy lands, while lower (up to 0.3) for eastern croplands in our study area. The main reason for the
362   differences is that the C factor formula of the standard plot is different. The C factor algorithm based
363   on the vegetation index assumes only one vegetation type in the km-sized grid, and calculates the
364   coverage of the vegetation to obtain the C factor. Therefore, in the sandy area where the complete
365   distribution of sand is assumed, the C factor is overestimated, while in cropland areas where the
366   complete distribution of cropland is assumed, the C factor is underestimated. As such, the C factor
367   based on the remote sensing data is more precise than that based on the vegetation index.

368   *5.3   Uncertainties of the Calculated C Factors*

369   5.3.1   The empirical parameters of the C factor need to be further confirmed

370   The water erosion equation used in this paper is a small-scale method, and although many
371   previous studies have applied it on regional scales, many of its parameters are empirical, and their
372   rationality has not been confirmed: 1) The measured parameters of C factor are for selected field
373   survey time and area. The C factor in this paper was obtained from remote sensing and field survey
374   data. The summer months (mainly July and August) in the study area have the strongest water
375   erosion throughout the year, and there was almost no water erosion in other months. Besides, the
376   vegetation coverage is higher in summer. The summer C factor determined the accuracy of water
377   erosion simulation. Therefore, the field investigation time and TM image acquisition time of this
378   study were from July to August. 2) Empirical parameters of the C factor obtained from remote sensing
379   data. The coverage of cropland, woodland and grassland was obtained by interpreting the remote
380   sensing images and then calculating the proportion of land use/cover. The grasslands can be divided
381   into high coverage, medium coverage and low coverage grassland[34]. Among them, the high
382   coverage grassland has a coverage of over 50%, assumed as 80%; the medium coverage grassland has
383   a coverage of 20%–50%, assumed as 50%; while the low coverage grassland has a coverage of 5%–20%
384   of the grass, assumed as 20%. The croplands are classified as irrigated and dry lands. The woodlands
385   are divided into woodlands and shrubs and the coverages of cropland and woodland vegetation are
386   assumed to be 100%. Swamps are low-lying wetlands, which are different from other unused lands

387 (bare land, quicksand, saline and alkaline land), while similar to high coverage grassland, with an
388 assumed coverage of 80%. These vegetation coverage parameters are empirical to some degree and
389 should be further confirmed in the further study.

390 5.3.2 Differences between the remote sensing image coverage and the measured coverage

391   Based on high-precision field survey and remote sensing image interpretation, this study
392 constructed a km-sized grid C factor algorithm. Although the land cover data obtained from TM
393 image interpretation had already the highest resolution on a regional scale, the coverage was still
394 different from the measured coverage because TM images contained mixed pixels. In addition, based
395 on TM images we can only measure the canopy coverage, rather than surface coverage that has
396 impacts on soil erosion [42]. Furthermore, currently the regional surface coverage can only be
397 determined based on the surface coverage of representative sites through upscaling.

398 5.3.3 The accuracy of the TM image interpretation

399   In this study, we randomly generated points in the TM images to verify the accuracy of our land
400 use interpretation. A total of 2,000 verification points (50 in each image) were generated in the study
401 area, where each verification point was a single pixel. Based on the high-precision Google Earth
402 images, we determined the land use type of each verification point to obtain the verification data. In
403 order to ensure the accuracy of the verification data and to reduce the mixed pixel problem caused
404 by resolution difference and time inconsistence of the verification data, we moved the random points
405 at the edge of the land use units to the center of land use type. Using the randomly generated 2,000
406 verification points, we tested the interpreted 2010 land use data. The results are shown in Table 5.

407 Table 5 Accuracy of land use interpretation in the study area

		True surface						
		Water body	Built-up land	Grassland	Cropland	Woodland	Unused land	User's Accuracy(%)
Land use classification results	Water body	40	0	1	4	1	1	85.1
	Built-up land	0	61	7	17	8	5	62.24
	Grassland	1	9	551	120	41	61	70.37
	Cropland	0	10	53	443	60	18	75.86
	Woodland	0	3	39	49	234	10	69.85
	Unused land	0	3	14	14	4	111	76.03

408   It can be seen from Table 5 that the interpretation accuracy is 72.25% and the Kappa coefficient
409 is 0.62. Landis and Koch [43] point out that a Kappa coefficient larger than 0.6 indicates good
410 accuracy, thus the interpretation accuracy of this paper is good. Many papers has proved that the
411 remote sensing interpretation of large region of low accuracy[36,37]. For example, the accuracy of
412 many automatic classification methods, is basic less than 65%[37], and this paper is study the
413 fragmented landscape areas for the ecological transition zone, so 72.25% accuracy has been enough
414 high.

415   There are several problems in the interpretation of land use in the study area: 1) The problem of
416 mixed pixel, and different spatial resolutions among Landsat TM and Google earth decreases the
417 interpretation accuracy. 2) Problems arising from different types of lands having the same type of
418 spectrum. The spectra of unused lands and construction lands are relatively close, and on large
419 regional scale the proportion of construction lands is relatively low, so erosion is rarely influenced.
420 Unused lands such as bareland will cause erosion. However, this paper studied the vegetation

cover(C) factor which is related to vegetation (cropland, grassland and woodland), so the misclassification of construction land and bareland, though resulted in low accuracy, did not affect the vegetation coverage calculation. So the error can be ignored since this paper focuses on the proportion of vegetation (i.e. croplands, woodlands and grasslands). 3) The mosaic problem. Due to the phase differences of different TM images, there exist edge snap problems between different TM images. These problems reduced the accuracy of the TM image interpretation, so that the results had some uncertainty in the edge areas. However, the overall interpretation accuracy of cropland, woodland and grassland can meet the requirement of vegetation coverage and C factor calculation.

6   Conclusions

The vegetation cover (C) factor is one of the most influential factors in the soil erosion model. The C factor derived from a standard plot cannot be directly used on regional scale. Based on remote sensing data and field investigations, we designed a C factor conversion method from the standard plot to a km-sized grid based on large sample theory. It can be concluded that: 1) Compared with existing C factor algorithms, our algorithm improves the applicable range of the C factor formula of the standard plot, and can be used to simulate soil erosion in large areas. 2) The vegetation coverage obtained by remote sensing interpretation is significantly consistent (paired samples t-test, t = −0.03, df = 0.12, 2-tail significance $p < 0.05$) and significantly correlated with the measured vegetation coverage. Meanwhile, the line-point intercept method can be used to effectively obtain the vegetation coverage of cropland, woodland and grassland in the survey sample area ($p < 0.05$). 3) The C factor of the study area is smaller in the middle, southern and northern regions, and larger in the eastern and western sections. The main reason for this is the distribution of woodlands, the Hunshandake and Horqin sandy lands and human cultivation that affects the valleys; 4) The C factor conversion method based on large sample theory is better than the one based on vegetation indices.

In this paper, a method for estimating the regional C factor was proposed by combining the interpretation of remote sensing data and data obtained from field investigations. Our method is limited by the resolution of the remote sensing data, and the accuracy of the TM image interpretation. Thus, in the future, we aim to develop our method by: 1) further confirming the empirical parameters of the C factor; 2) establishing a database of C factors in different seasons; 3) studying the differences between remote sensing and high precision measured vegetation coverage.

Acknowledgments: This study was supported by the National Natural Science Foundation of China (Project No.: 41271286), and also supported by the National Key Research and Development Program (No. 2016YFA0602402) and also supported by the Natural Science Foundation of Zhejiang Province, China(LQ18D010003). Thanks to the people from Beijing Normal University: Bo Chen, Gangfeng Zhang, LiLi Lv, Fengfang Chen, Feng Kong, Yongchang Meng, Zhu Wang, Mengyang Li, Mengjie Li, Zhao Liu, Jiayi Fang, Fan Liu, Xu Yang, ZhuoRong Ying, Hao Guo, Yibo Luan, Bank Hu, Jie Zhang and others, to take part to investigate the vegetation Coverage in Eastern Section of Farming-pasture Ecotone of Northern China.

Author Contributions: Degen Lin conceived the entire paper, analyzed the data and designed the experiments. Yuan Gao, Yaoyao Wu and Huiming Yang analyzed the data and polished the text. Peijun Shi provided technical support, polished the text and provided project support. Jing'ai Wang provided technical support, polished the text and provided project support.

Conflicts of Interest: Declare conflicts of interest or state "The authors declare no conflict of interest." Authors must identify and declare any personal circumstances or interest that may be perceived as inappropriately influencing the representation or interpretation of reported research results. Any role of the funding sponsors in the design of the study; in the collection, analyses or interpretation of data; in the writing of the manuscript, or in the decision to publish the results must be declared in this section. If there is no role, please state "The

466   founding sponsors had no role in the design of the study; in the collection, analyses, or interpretation of data; in
467   the writing of the manuscript, and in the decision to publish the results".

468   **Appendix A: The C Factor Algorithm Based on the Remote Sensing Vegetation Index**

469   1.   Calculate vegetation coverage based on the Vegetation index

470   In order to express the vegetation coverage of the study area, the MODIS EVI vegetation index
471   needs to be partitioned according to the measured vegetation coverage to carry out the regression
472   modeling. These regression models are generally applicable to the area where the relational
473   expression is constructed. Therefore, the difference between the EVI vegetation index of MODIS and
474   the measured vegetation coverage (hereinafter referred to as FVCD) was used to divide the vegetation
475   index and the measured coverage model. Based on the natural geographical conditions and factors,
476   we analyzed the relationship between the dryness grade, landform type, vegetation type and sand
477   land, and used the Pearson correlation of FVCD, as shown in Table A1.

478   Table A1 The Pearson Relationship between Different Geographical Elements and FVCD

Different Geographical Elements	dryness degree	Sand land or not	landform type	Vegetation type
Pearson correlation	.386*	-.206	.114	-.283
Significance (2-tail)	.012	.190	.473	.069

479   It can be seen from Table A1 that the degree of dryness has the highest correlation, where the
480   Pearson's correlation coefficient of FVCD is 0.386, which is significant at the $p<0.05$ level. Therefore,
481   we have established a regression relationship between the vegetation index and the measured
482   vegetation coverage under different dryness. The study area is an arid and semi-arid region. The
483   dryness can be subdivided into dryness 1 (0.2–0.4), dryness 2 (0.4–0.44) and dryness 3 (0.44–0.5).
484   According to the degree of dryness, the relationship between vegetation index and measured
485   vegetation coverage was established, as shown in Table A2 and Figure A1.

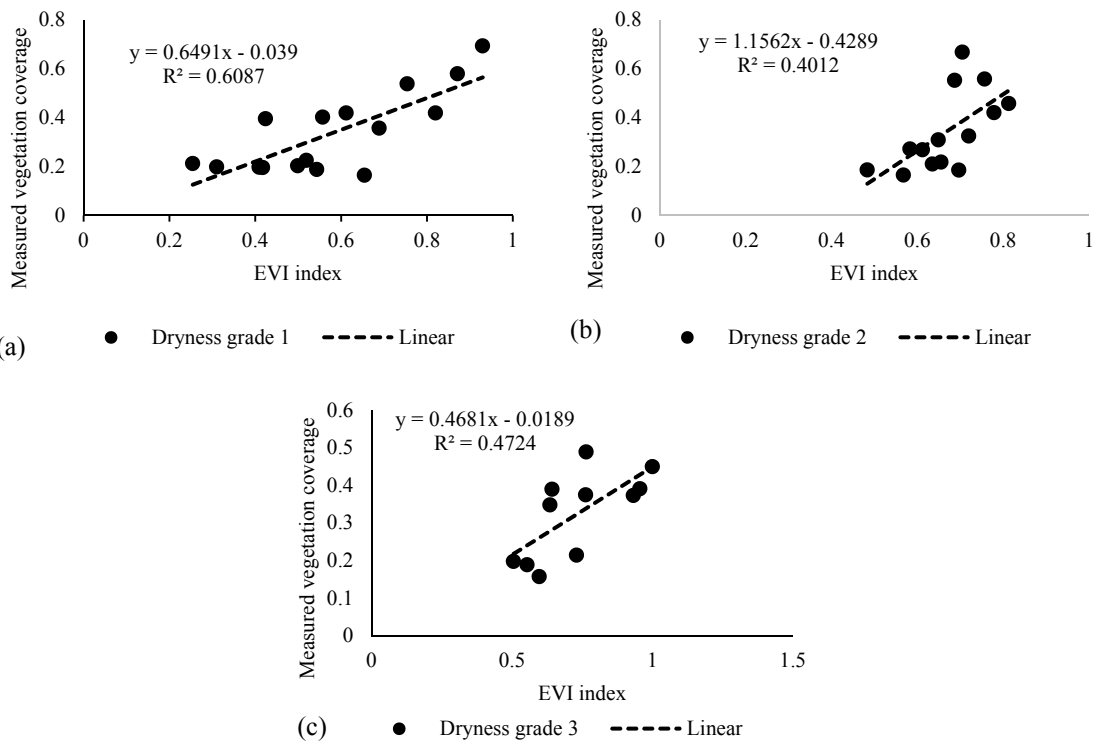
486   Table A2 The EVI coefficient and measured vegetation coverage under different dryness grades.

Dryness grade		Unstandardized coefficient		Normalization coefficient	t	Significance
		B	Standard error	Beta		
Dryness 1	(constant)	.263	.075		3.516	.003
	EVI coefficient	.938	.201	.780	4.667	.000
Dryness 2	(constant)	.548	.046		11.903	.000
	EVI coefficient	.347	.122	.633	2.836	.015
Dryness 3	(constant)	.406	.122		3.336	.009
	EVI coefficient	1.009	.356	.687	2.839	.019

487   Note: The dependent variable is the measured vegetation coverage

488   Table A2 expresses the regression equation of the EVI and measured vegetation coverage relative
489   to the dryness classification and its significance. When the dryness is grade 1, the significance level is
490   0.003 ($p<0.05$); when the dryness is grade 2, the significance level is 0.015 ($p<0.05$); when the dryness
491   is grade 3, the significance level is 0.019 ($p<0.05$).

492

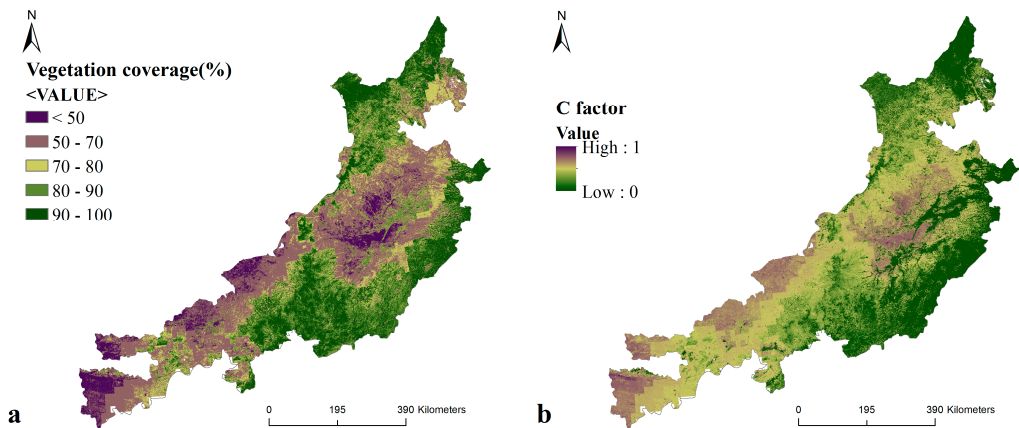
493

494   Figure A1 Relationship between the EVI and measured vegetation coverage under different dryness

495   2.   C factor Based on the Vegetation Index

496   According to the fitting equation for different dryness classifications, the FVC in the study area

497   was calculated (Figure A2a). Based on the MODIS land cover data, the C factor (Figure A2b) of the

498   study area was calculated according to the C factor formula of the standard plot.

499

500   Figure A2 Vegetation coverage and C cover factor based on the vegetation index (a: FVC map; b: C

501   factor map).

502   **References**

503   1.   EEA. *Environment in the european union at the turn of the century. European environment*

504   *agency*. Office des Publications: Luxembourg, 1999.

505   2.   EEA. *Climate change, impacts and vulnerability in Europe 2012: An indicator-based report*.

506   Office for Official Publ. of the Europ. Union: 2012.

507   3.   Lin, D.; Guo, H.; Lian, F.; Gao, Y.; Yue, Y.; Wang, J.a. A quantitative method for long-term

508   water erosion impacts on productivity with a lack of field experiments: A case study in huaihe

509   watershed, china. *Sustainability* **2016**, *8*, 675.

510   4.   Panagos, P.; Borrelli, P.; Poesen, J.; Ballabio, C.; Lugato, E.; Meusburger, K.; Montanarella,

511   L.; Alewell, C. The new assessment of soil loss by water erosion in Europe. *Environmental*

512   *science & policy* **2015**, *54*, 438-447.

513   5.   Panagos, P.; Ballabio, C.; Borrelli, P.; Meusburger, K. Spatio-temporal analysis of rainfall

514   erosivity and erosivity density in greece. *Catena* **2016**, *137*, 161-172.

515   6.   Nachtergaele, F.; Petri, M.; Biancalani, R.; Van Lynden, G.; Van Velthuizen, H. *An information*

516   *database for land degradation assessment at global level*; Global land degradation

517   information system (GLADIS): 2010.

518   7.   Wischmeier, W.H.; Smith, D.D. Predicting rainfall erosion losses-a guide to conservation

519   planning. *Predicting rainfall erosion losses-A guide to conservation planning* **1978**.

520   8.   Renard, K.G.; Foster, G.R.; Weesies, G.A.; McCool, D.K.; Yoder, D.C. *Predicting soil erosion*

521   *by water: A guide to conservation planning with the revised universal soil loss equation (rusle)*.

522   United States Department of Agriculture: Washington, DC, 1997; Vol. 703.

523   9.   Renard, K.G.; Foster, G.R.; Weesies, G.A.; Porter, J.P. Rusle: Revised universal soil loss

524   equation. *Journal of soil and Water Conservation* **1991**, *46*, 30-33.

525   10.   Liu, B.; Zhang, K.; Xie, Y. In *An empirical soil loss equation*, Proceedings 12th international

526   soil conservation organization conference, Beijing, 2002; Tsinghua University Press: Beijing,

527   p 15.

528   11.   Jones, C.; Lowe, J.; Liddicoat, S.; Betts, R. Committed terrestrial ecosystem changes due to

529   climate change. *Nature Geoscience* **2009**, *2*, 484.

530   12.   Benkobi, L.; Trlica, M.; Smith, J.L. Evaluation of a refined surface cover subfactor for use in

531   rusle. *Journal of Range Management* **1994**, 74-78.

532   13.   Biesemans, J.; Van Meirvenne, M.; Gabriels, D. Extending the rusle with the monte carlo

533   error propagation technique to predict long-term average off-site sediment accumulation.

534   *Journal of Soil and Water Conservation* **2000**, *55*, 35-42.

535   14.   Risse, L.M.; Nearing, M.A.; Laflen, J.M.; Nicks, A.D. Error assessment in the universal soil

536   loss equation. *Soil Science Society of America Journal* **1993**, *57*, 825-833.

537   15.   Nachtergaele, F.; Petri, M.; Biancalani, R.; Van Lynden, G.; Van Velthuizen, H. In *Global land*

538   *degradation information system (gladis)*, Scientific Conference, 2010; Unccd, Ed.

539   16.   Suriyaprasit, M.; Shrestha, D.P. Deriving land use and canopy cover factor from remote

540   sensing and field data in inaccessible mountainous terrain for use in soil erosion modelling.

541   *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information*

542   *Sciences* **2008**, *37*, 1747-1750.

543   17.   Warren, S.D.; Mitasova, H.; Hohmann, M.G.; Landsberger, S.; Iskander, F.Y.; Ruzycki, T.S.;

544   Senseman, G.M. Validation of a 3-d enhancement of the universal soil loss equation for

545   prediction of soil erosion and sediment deposition. *Catena* **2005**, *64*, 281-296.

546   18.   Wang, G.; Wente, S.; Gertner, G.; Anderson, A. Improvement in mapping vegetation cover

547   factor for the universal soil loss equation by geostatistical methods with landsat thematic

548   mapper images. *International Journal of Remote Sensing* **2002**, *23*, 3649-3667.

549   19.   Lu, D.; Li, G.; Valladares, G.S.; Batistella, M. Mapping soil erosion risk in rondonia, brazilian

550   amazonia: Using rusle, remote sensing and gis. *Land Degradation & Development* **2004**, *15*,

551   499-512.

552   20.   Patric, J.H. Soil erosion in the eastern forest. *Journal of Forestry* **1976**, *74*, 671-677.

553   21.   Gertner, G.; Wang, G.; Fang, S.; Anderson, A.B. Mapping and uncertainty of predictions

554   based on multiple primary variables from joint co-simulation with landsat tm image and

555   polynomial regression. *Remote Sensing of Environment* **2002**, *83*, 498-510.

556   22.   Fu, B.; Xu, Y.; Lv, Y. Scale characteristics and coupled research of landscape pattern and

557   soil and water loss. *Advances In Earth Science* **2010**, *25*, 673-681.

558   23.   Lehmann, E.L. *Elements of large-sample theory*. Springer Science & Business Media %\

559   2017-04-06 21:59:00: Netherlands, 2004.

560   24.   Shannon, C.E. A mathematical theory of communication. *ACM SIGMOBILE Mobile*

561   *Computing and Communications Review* **2001**, *5*, 3-55.

562   25.   Ji, L.; Wang, J.; Geng, X.; Gong, P. Probabilistic graphical model based approach for water

563   mapping using gaofen-2 (gf-2) high resolution imagery and landsat 8 time series. *arXiv*

564   *preprint arXiv:1612.07801* **2016**.

565   26.   Floyd, D.A.; Anderson, J.E. A comparison of three methods for estimating plant cover. *The*

566   *Journal of Ecology* **1987**, *75*, 221-228.

567   27.   Godínez-Alvarez, H.; Herrick, J.E.; Mattocks, M.; Toledo, D.; Van Zee, J. Comparison of three

568   vegetation monitoring methods: Their relative utility for ecological assessment and monitoring.

569   *Ecol Indic* **2009**, *9*, 1001-1008.

570   28.   Hanley, T.A. A comparison of the line-interception and quadrat estimation methods of

571   determining shrub canopy coverage. *Journal of Range Management* **1978**, 60-62.

572   29.   Edwards, T.A. Monitoring plant and animal populations. *Pacific Conservation Biology* **2002**,

573   *8*, 219-219.

574   30.   Duncan, J.; Stow, D.; Franklin, J.; Hope, A. Assessing the relationship between spectral

575   vegetation indices and shrub cover in the jornada basin, new mexico. *International Journal*

576   *of Remote Sensing* **1993**, *14*, 3395-3416.

577   31.   Purevdorj, T.; Tateishi, R.; Ishiyama, T.; Honda, Y. Relationships between percent vegetation

578   cover and vegetation indices. *International Journal of Remote Sensing* **1998**, *19*, 3519-3535.

579   32.   Chen, J.; Chen, J.; Liao, A.; Cao, X.; Chen, L.; Chen, X.; He, C.; Han, G.; Peng, S.; Lu, M.*,*

580   *et al.* Global land cover mapping at 30 m resolution: A pok-based operational approach.

581   *ISPRS Journal of Photogrammetry and Remote Sensing* **2015**, *103*, 7-27.

582   33.   Gong, P.; Wang, J.; Yu, L.; Zhao, Y.; Zhao, Y.; Liang, L.; Niu, Z.; Huang, X.; Fu, H.; Liu, S.

583   Finer resolution observation and monitoring of global land cover: First mapping results with

584   landsat tm and etm+ data. *International Journal of Remote Sensing* **2013**, *34*, 2607-2654.

585   34.   Liu, J.; Liu, M.; Tian, H.; Zhuang, D.; Zhang, Z.; Zhang, W.; Tang, X.; Deng, X. Spatial and

586   temporal patterns of china's cropland during 1990　2000: An analysis based on landsat tm

587   data. *Remote Sensing of Environment* **2005**, *98*, 442-456.

588   35.   Jengo, C. *Rulegen v1. 02 extension in envi*; ITT Visual Information Solutions User-

589   Contributed Library: 2004.

590   36.   Zhao, Y.; Gong, P.; Yu, L.; Hu, L.; Li, X.; Li, C.; Zhang, H.; Zheng, Y.; Wang, J.; Zhao, Y.

591   Towards a common validation sample set for global land-cover mapping. *International*

592   *Journal of Remote Sensing* **2014**, *35*, 4795-4814.

593   37.   Chen, J.; Chen, J.; Liao, A.; Cao, X.; Chen, L.; Chen, X.; He, C.; Han, G.; Peng, S.; Lu, M.

594   Global land cover mapping at 30m resolution: A pok-based operational approach. *ISPRS*

595   *Journal of Photogrammetry and Remote Sensing* **2015**, *103*, 7-27.

596 38. Wischmeier, W.H.; Smith, D.D. *Predicting rainfall erosion losses-a guide to conservation*

597   *planning*. U. S. Dep. Agric. , Agric. Handb: 1978; p 537.

598 39. Bollinne, A. Adjusting the universal soil loss equation for use in western Europe. **1985**.

599 40. Jin, Z.; Shi, P.; Hou, F. *Soil erosion system model and control model in huangfuchuan*

600   *watershed of the yellow river*. Ocean Press: Beijing, 1992.

601 41. Jiang, z.; WANG, Z.; Liu, Z. Quantitative study on spatial variation of soil erosion in a small

602   watershed in the loess hilly region. *Journal of soil erosion and soil conservation* **1996**, *2*, 1-9.

603 42. Liu, B. *Soil loss equation of beijing*. Science press: 2010.

604 43. Landis, J.R.; Koch, G.G. Measurement of observer agreement for categorical data.

605   *Biometrics* **1977**, *33*, 159-174.

606

607