The Boundaries and Remote Sensing Classification Datasets on Large Wetlands of International Importance in 2001 and 2013

Zhang, H. Y.[1]   Niu, Z. G.[1*]   Xu, P. P.[1]   Chen, Y. F.[1]   Hu, S. J.[1]   Gong, N.[2]

1. State Key Laboratory of Remote Sensing Science, Institute of Remote Sensing and Digital Earth, Chinese Academy of Sciences, Beijing 100101, China.
2. Department of Resources and Environment, Shandong Agricultural University, Tai'an 271018, China

Abstract: Wetland is one of the ecosystems with the highest biodiversity and productivity in the world. In this paper, we developed a dataset of 100 large Wetlands of International Importance (the "Ramsar Sites"), which was developed by the data integration and mining from the World Database on Protected Areas (WDPA) and remote sensing images, expecially the high resolution imagery provided on the Google Earth and the Tianditu platforms. Each site covers at least an area of 200,000 hm2. Based on the Moderate Resolution Imaging Spectro radiometer (MODIS) 16-day composite products MOD13Q1 of these sites in 2001 and 2013, we applied the Savitzky-Golay filtering and time series reconstruction algorithms to Normalized Difference Vegetation Index (NDVI) data. The method of the Support Vector Machine (SVM) classifier was used to classify and map wetlands of these Ramsar sites. The time series data based wetland classification results were compared with high resolution remote sensing imagery and the relevant literatures. This method is more advantageous than single temporal imagery based wetland monitoring technique. The dataset consists of three components which are saved as the .shp and the .kmz file formats. The total data size is 102 MB in compressed file format.

Keywords: wetlands; boundary; remote sensing classification; wetland mapping; Ramsar sites

1 Introduction

Wetland is one of the ecosystems with the highest biodiversity and productivity in the world[1]. In recent years, large scale wetland monitoring and mapping emerged as a hot research topic and has been attracting great attention of many scholars because of its ecological significance. To understand the spatio-temporal change characteristics of wetland ecosystem is the key component of global change research[2]. The fast development of remote sensing technique provides an objective means to global wetlands monitoring research. The remotely

sensed time series data play an irreplaceable role in the research of wetland changes in the growth cycle or in a calendar year of wetland vegetation[3]. It is the main data source of the remotely sensed monitoring on large wetlands of international importance at the global scale of this paper. Along with our related scientific reports published, we make this dataset available for public use.

2 Metadata

The metadata of the boundaries and remote sensing classification datasets on large wetlands of international importance in 2001 and 2013 (RamsarSites_Top100_2001/2013) is summarized in Table 1[4]. It includes the dataset full name, short name, authors, year of the dataset, temporal resolution, spatial resolution, data format, data size, data files, data publisher, and data sharing policy, et al.

Table 1 Summary of the RamsarSites_Top100_WetlandCover_2001/2013 metadata.

Items	Description
Dataset full name	The wetland cover datasets on the large wetlands of international importance in 2001 and 2013 by remote sensing data integration
Dataset short name	RamsarSites_Top100_WetlandCover_2001/2013
Authors	Zhang, H. Y. L-4985-2016, Institute of Remote Sensing and Digital Earth, CAS, zhanghy@ radi.ac.cn Niu, Z. G. L-4829-2016, Institute of Remote Sensing and Digital Earth, CAS, niuzg@ra di.ac.cn Xu, P. P. L-5064-2016, Institute of Remote Sensing and Digital Earth, CAS, 25486400 46 @qq. com Chen, Y. F. L-5003-2016, Institute of Remote Sensing and Digital Earth, CAS, 935836745 @qq.com Hu, S. J. L-6142-2016, Institute of Remote Sensing and Digital Earth, CAS, husj1989@y eah.net Gong, N. L-6422-2016, Shandong Agricultural University, gongningbaobao@126.com
Geographic region	100 large wetlands of international importance in the world, distributed in different continents.
Year	2001, 2013
Temporal resolution	Yearly
Spatial resolution	250 m
Data format	.kmz, .shp
Data size	1.54 MB in .kmz format. 100 MB in compressed
Data files	The dataset consists of three files [4]. They are as follows. 1) Top100_boundary.kmz.zip. This is the boundaries of 100 large wetlands of international importance in the world. It is stored in a KMZ file applicable to Google Earth, and the data size is 1.54MB. 2) Ramsar2001.rar. This is the remotely sensed classification results of these wetlands in 2001. Its data size is 51MB. 3) Ramsar2013.rar. This is the remotely sensed classification results of these wetlands in 2013. Its data size is 49.6MB. Note: The classification results, Ramsar2001.rar and Ramsar2013.rar, are stored as ArcGIS shapefile vector format. Each vector file includes seven fields: (1) Area refers to the area of each patch in square meters; (2) RSN stands for Ramsar Site No., namely Ramsar number or the number of wetlands of international importance; (3) Lev3 indicates wetland code of that patch in the classification system [4] (Table 2); (4) the fields Name_EN and Name_CN refer to the English names and the Chinese names of the Ramsar sites respectively; and (5) the fields Country_EN and Country_CN respectively stand for the English names and the Chinese names of the country where the Ramsar sites lie in.
Foundation	National Natural Science Foundation of China (41271423)
Data publisher	Global Change Research Data Publishing & Repository, http://www.geodoi.ac.cn.

Items	Description
Address	No. 11A, Datun Road, Chaoyang District, Beijing, 100101, China.
Data sharing policy	***Data*** from the Global Change Research Data Publishing & Repository includes metadata, datasets (data products), and publications (in this case, in the *Journal of Global Change **Data** & Discovery*). ***Data*** sharing policy includes: (1) ***Data*** are openly available and can be free downloaded via the Internet; (2) End users are encouraged to use *Data* subject to citation; (3) Users, who are by definition also value-added service providers, are welcome to redistribute ***Data*** subject to written permission from the GCdataPR Editorial Office and the issuance of a ***Data*** redistribution license, and; (4) If ***Data*** are used to compile new datasets, the 'ten per cent principal' should be followed such that ***Data*** records utilized should not surpass 10% of the new dataset contents, while sources should be clearly noted in suitable places in the new dataset [5].

3 Methods

3.1 The objective and indicators of wetland monitoring by remote sensing

Based on some principles such as wetlands with an area larger than 200,000 hm^2, with established geographic extent and map information, 100 wetlands of international importance all over the world distributed at different ecological-climatic zones and continents were selected from the database of Ramsar sites (https://rsis.ramsar.org/) (Figure 1).

Indices for monitoring large wetlands of international importance mainly include the types and the area of land cover (Table 2). The conversion between wetlands and non-wetlands, the conversion in different wetland types and the changes on the area of different types, were used as important indicators to evaluate the ecological and environmental conditions of the wetlands. In general, if wetlands are converted into non-wetlands, or natural wetlands are converted into artificial wetlands, they usually lead to a deteriorated environmental condition and pose some threats to wetland biodiversity conservation. On the contrary, it often causes healthy development on the wetland environments.
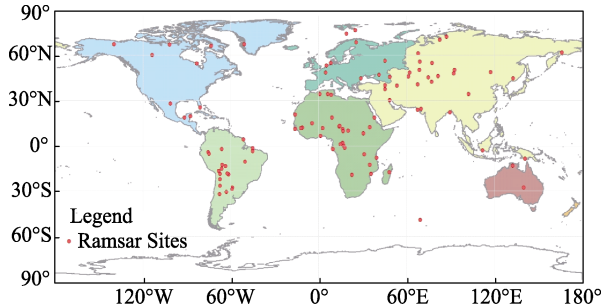

Figure 1 Spatial distribution of large wetlands of international importance

3.2 Methodology of wetland boundaries and classification dataset development

Most of the wetland boundaries were delineated by reference to the scanned paper maps published on the Ramsar website (https://rsis.ramsar.org/. Retrieved on July 23, 2014)[6]. Another small part of them were derived from World Database on Protected Areas (https://www.iucn.org/; http://www.protectedplanet.net/. Retrieved on July 23, 2014) and the other products of literature and document. The maps were georeferenced in ArcGIS. Because there are many differences of wetland management and monitoring techniques in each country, some inconsistencies of the boundaries often occur in such occasion. Therefore we made modification and verification by using high resolution remote sensing images on Google Earth platform, based on the physiographic forms of wetlands such as walking along rivers, roads, and the ridge lines. The MODIS product MOD13Q1 (16-day composite, 250 m)

Table 2 Land cover classification system adopted in remote sensing based mapping on wetlands of international importance

Level 1	Level 2	Level 3	Code	Characteristics of time series NDVI	Definition or spatial distribution characteristics
Wetlands	Water		11	$NDVI \leqslant 0.1$	Including rivers, lakes, estuarine wa ter, reservoirs and urban entertainment landscape, wastewater treatment plant, etc. It mainly refers to permanent water body.
	Forested/shrub-dominated wetlands	Forested/shrub-dominated wetlands (evergreen)	121	$NDVI \geqslant 0.4$. There is slight amplitude of variation on NDVI in a year.	Including the forested/shrub-dominated wetlands with evergreen vegetation located at the inland and coastal zone.
		Forested/shrub-do min ated wetlands (deciduous)	122	$NDVI \geqslant 0.3$. It presents a pattern of periodicity.	Including the forested/shrub-dominated wetlands with deciduous vegetation located at the inland and coastal zone.
	Marshes	Permanent marshes	131	$0.1 \leqslant NDVI \leqslant 0.5$. There is a relative slight amplitude of variation on NDVI in a year. In high-water season, there may be minus value of NDVI.	Wetlands with oversaturated soil and permanently flooded herbaceous plants.
		Seasonal marshes	132	$0.1 \leqslant NDVI \leqslant 0.5$. There is a periodic pattern of NDVI in a year. In high-water season, $NDVI < 0$	Herbaceous marshes flooded regularly or seasonally in leaf-on season.
	Flooded flats/Intertidal zone / Estuarine deltas		14	$NDVI \leqslant 0.2$. In high-water season, $NDVI < 0$.	They refer to inundated periodically and oversaturated soils distributed near lakes, rivers and estuary, with the vegetation coverage lower than 30%. They are often not covered with open water.
	Paddy fields		15	$NDVI < 0$, at the early stage of crop growth. There is a steep decreasing trend of NDVI after crops harvesting.	Arable land used for growing rice.
	Tundra wetlands		16	In leaf-on season, $NDVI < 0.6$. They are usually distributed at high latitudes.	
Non-wetlands	Snow		21		Perennial snow
	Natural vegetation	Forests/Scrublands	22	In the leaf-on season, $NDVI > 0.5$	Forests and scrublands
		Grasslands	23	In the leaf-on season, $0.3 \leqslant NDVI \leqslant 0.6$	
	Construction / Barren lands		24	$0 \leqslant NDVI \leqslant 0.2$	Including artificial construction such as buildings, roads, etc. and other forms of barren lands, e.g., rocks, barren sands.
	Drylands		25	The occurrence of the slow growth and the steep drop-off of the NDVI value in a calendar year.	Farms cultivated with dry-land crops as a result of their scarcity of water.
	Cloud cover		32		These regions are covered by clouds.
	Ocean		33		

in 2001 and 2013 downloaded from NASA (https://ladsweb.nascom.nasa.gov/data/) was used as the main data source. There are 23 periods of the images in each year. In this study, the images on the public service platform Tianditu (http://www.tianditu.com/map/index.html)

powered by the National Geomatics Center of China (NGCC) and other geographic data were also used as auxiliary data. The Normalized Difference Vegetation Index (NDVI) can reflect the growth status of different vegetation types. By comparing the NDVI differences including the maximum, the minimum, and the range of annual NDVI, the emergence date of the positive NDVI value, and the growing season length, etc. between different vegetation types in the growth cycle or within a calendar year, and analyzing plants growth pattern [7],

the accuracy of interpretation on different types of land cover types could be effectively improved. We applied Savitzky-Golay filtering to reconstruct NDVI time series dataset [8]. Based on the clustering results of time series data and the analysis on the features of different land cover types in high resolution remote sensing images, we collected 10 – 20 training samples for each type and used SVM classifier to delineate and map different types of wetlands. The minimum of the map unit is about 56.25 hm^2 for 3×3 pixels. The main methods and the technical flow are described in Figure 2.
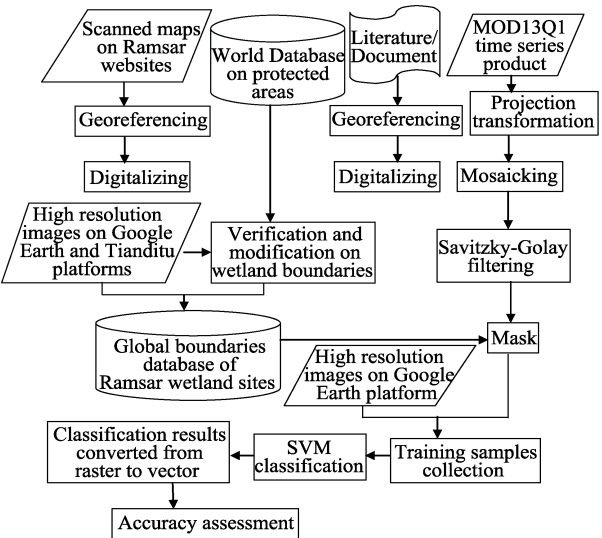

Figure 2 Flowchart for wetlands classification

4 Data product

The dataset consists of three components, namely Top100_boundary kmz (1470.14KB), Ramsar2001.rar (52245.46KB) and Ramsar2013.rar (50858.42KB). They represent the boundaries of 100 large wetlands of international importance in the world, and the wetland classification results in 2001 and 2013, respectively (http://www.geodoi.ac.cn/WebEn/doi.aspx?Id=243).

4.1 Statistics on large wetlands of international importance

The total area of 100 large wetlands of international importance is 138,392,121.05 hm^2. Okavango Delta System (Ramsar Site No.879), located in Botswana, ranks first. The statistical results are listed in Table 3.

Table 3 Statistics on large wetlands of international importance

Name of Ramsar Sites	Area (hm^2)	Country	Ramsar Site No.
Okavango Delta System	6703463.94	Botswana	879
Queen Maud Gulf	6263006.27	Canada	246
Ngiri-Tumba-Maindombe	6194581.31	Democratic Republic of the Congo	1784
Plaines d'inondation des Bahr Aouk et Salamat	6187728.80	Chad	1621
Grands affluents	5787900.96	Republic of Congo	1742
Plaines d'inondation du Logone et les dépressions Toupouri	4017225.78	Chad	1560
Complejo de humedales del Abanico del río Pastaza	3781662.25	Peru	1174
Malagarasi-Muyovozi Wetlands	3705800.43	Tanzania	1024
Río Yata	3448554.92	Bolivia	2094
Palmar de las Islas y las Salinas de San José	3434589.67	Bolivia	1088

(*To be continued on the next page*)

Name of Ramsar Sites	Area (hm^2)	Country	Ramsar Site No.
Delta Intérieur du Niger	3143537.63	Mali	1365
Río Blanco	2909863.42	Bolivia	2092
Plaine de Massenya	2905803.26	Chad	1839
Reentrancias Maranhenses	2671201.38	Brazil	640
Area between the Pura & Mokoritto Rivers	2620671.33	Russia	697
Sudd	2470516.39	South Sudan	1622
Gueltas et Oasis de l'Aïr	2413275.61	Niger	1501
Réserve Naturelle Nationale des Terres Australes Francaises	2332885.13	French	1837
Polar Bear Provincial Park	2231833.08	Canada	360
Pacaya-Samiria	2194600.24	Peru	546
Coongie Lakes	2142810.22	Australia	376
Río Matos	2116721.12	Bolivia	2093
Kakadu National Park	1939734.28	Australia	204
Parapolsky Dol	1826613.22	Russia	693
Baixada Maranhense Environmental Protection Area	1811854.16	Brazil	1020
Los Lípez	1729203.22	Bolivia	489
Lake Niassa and its Coastal Zone	1697149.57	Mozambique	1964
Partie tchadienne du lac Tchad	1686175.47	Chad	1134
Whooping Crane Summer Range	1616992.61	Canada	240
Parc National du Banc d'Arguin	1477460.61	Mauritania	250
Suakin-Gulf of Agig	1450427.12	Sudan	1860
Site Ramsar Odzala Kokoua	1385476.02	Republic of Congo	2080
Lago Titicaca	1366536.56	Bolivia	959
Ba?ados del Río Dulce y Laguna de Mar Chiquita	1351673.35	Argentina	1176
Mamirauá	1319593.43	Brazil	623
Lagunas altoandinas y puneñas de Catamarca	1311834.01	Argentina	1865
Tobol-Ishim Forest-steppe	1278538.61	Russia	679
Gambie-Oundou-Liti	1246779.36	Guinea	1579
Volga Delta	1199249.85	Russia	111
Lagunas de Guanacache, Desaguadero y del Bebedero	1157839.02	Argentina	1012
Zambezi Delta	1140940.11	Mozambique	1391
Lagos Poopó y Uru Uru	1136273.65	Bolivia	1181
Archipel Bolama-Bijagós	1065174.91	Guinea-Bissau	2198
Ili River Delta and South Lake Balkhash	973699.53	Kazakhstan	2020
Lake Uvs and its surrounding wetlands	885360.38	Mongolia	1379
Har Us Nuur National Park	859526.41	Mongolia	976
Dinder National Park	855072.75	Sudan	1461
Alakol-Sasykkol Lakes System	784807.01	Kazakhstan	1892
Dewey Soper Migratory Bird Sanctuary	782498.00	Canada	249
Le Lac Alaotra : Les Zones Humides et Bassins Versants	781472.69	Madagascar	1312
Everglades National Park	754573.06	United States	374
Brekhovsky Islands in the Yenisei estuary	737894.04	Russia	698
Área de Protección de Flora y Fauna Laguna de Términos	734350.69	Mexico	1356
Rufiji-Mafia-Kilwa Marine Ramsar Site	734194.39	Tanzania	1443

(*Continued*)

Name of Ramsar Sites	Area (hm^2)	Country	Ramsar Site No.
Dalai Lake National Nature Reserve, Inner Mongolia	733235.89	China	1146
Sian Ka'an	655727.26	Mexico	1329
Aydar-Arnasay Lakes system	636595.85	Uzbekistan	1841
Isyk-Kul State Reserve with the Lake Isyk-Kul	629536.18	Kyrgyzstan	1231
Chott El Jerid	616315.19	Tunisia	1699
Lake Chad Wetlands in Nigeria	611134.90	Nigeria	1749
Rio Sabinas	606876.07	Mexico	1769
Eqalummiut Nunaat and Nassuttuup Nunaa	594195.32	Denmark	386
Ilha do Bananal	592910.39	Brazil	624
Site Ramsar du Complexe W	581917.66	Benin	1668
Sangha-Nouabalé-Ndoki	552358.22	Republic of Congo	1858
Wadden Sea	541257.30	Netherlands	289
Cabo Orange National Park	523199.02	Brazil	2190
Bafing-Falémé	522191.82	Guinea	1719
Chott Ech Chergui	519421.40	Algeria	1052
Bañados del Izozog y el río Parapetí	516491.19	Bolivia	1087
Runn of Kutch	513386.48	Pakistan	1285
Sundarbans Reserved Forest	499921.44	Bangladesh	560
Upper Dvuobje	499157.08	Russia	678
Jaaukanigás	491725.87	Argentina	1112
Humedales Chaco	491620.32	Argentina	1366
Schleswig-Holstein Wadden Sea and adjacent areas	481776.92	Germany	537
Lake Sevan	474652.91	Armenia	620
Lake Urmia [or Orumiyeh]	456996.73	Iran	38
Danube Delta	451407.18	Romania	521
Indus Delta	435103.46	Pakistan	1284
Wasur National Park	411337.03	Indonesia	1624
Chany Lakes	387239.92	Russia	680
Tanjung Puting National Park	378350.88	Indonesia	2192
Zones Humides du Littoral du Togo	368011.51	Togo	1722
Shadegan Marshes & mudflats of Khor-al Amaya & Khor Musa	334645.27	Iran	41
Turkmenbashy Bay	324761.47	Turkmenistan	1855
Hopen	316453.18	Norway	1957
Lesser Aral Sea and Delta of the Syrdarya River	313371.18	Kazakhstan	2083
Bear Island	296711.28	Norway	1966
Lemmenjoki National Park	285550.79	Finland	1521
Lakes of the lower Turgay and Irgiz	263261.44	Kazakhstan	108
Tengiz-Korgalzhyn Lake System	261593.75	Kazakhstan	107
Gansu Gahai Wetlands Nature Reserve	247747.47	China	1975
Etangs de la Champagne humide	241348.36	France	514
Kama-Bakaldino Mires	223709.33	Russia	670
Lake Khanka	192382.13	Russia	112
Chott Melghir	192282.62	Algeria	1296
Veselovskoye Reservoir	185046.81	Russia	672
Petit Loango	150542.08	Gabon	352
Old Crow Flats	31461.26	Canada	244

4.2 Remote sensing classification results on large wetland of international importance in 2001 and 2013

Based on the remote sensing classification results on 100 large wetlands of international importance in 2001 and 2013, we calculate the subtotal area of different land cover types. The statistical results are shown in Table 4.

In 2001, the largest three wetland classes are forested/shrub-dominated wetlands (deciduous), seasonal marshes, and permanent marshes, respectively. Between 2001 and 2013, the total area of wetland reduced from 58134960 hm2 to 57,734,032 hm^2, the big three losers were water, forested/shrub-dominated wetlands (deciduous), and tundra wetlands, respectively.

Table 4 Area of different land cover types in large wetlands of international importance

Code	Level 2	Level 3	Area (hm^2) 2001	Area (hm^2) 2013
Wetlands			58134960.36	57734032.42
11	Water		7962751.49	7220121.10
121	Forested/shrub-dominated wetlands	Forested/shrub-dominated wetlands (evergreen)	5670938.35	5372163.35
122		Forested/shrub-dominated wetlands (deciduous)	17989514.92	17264579.94
131	Marshes	Permanent marshes	8852939.87	9019623.32
132		Seasonal marshes	11446049.25	12970340.19
14	Flooded flats/Intertidal zone / Estuarine deltas		3677715.96	3755553.93
15	Paddy fields		161536.44	118356.86
16	Tundra wetlands		2373514.09	2013293.73
Non-wetlands			80256924.04	80657761.93
21	Snow		203466.14	161471.34
22		Forests/Scrublands	49795219.42	49793411.47
23	Natural vegetation	Grasslands	10253074.54	10202998.58
24	Construction / Barren lands		10255547.25	9634365.16
25	Drylands		2071076.67	3200432.37
32	Cloud cover		54188.66	53787.91
33	Ocean		7624351.36	7611295.11

5 Data validation

Validation on the world's large wetlands classification results accuracy was applied by using high resolution remotely sensed images. We randomly selected 10 wetlands of international importance in different continents as the testing regions except for the Antarctic. About 2386 validation samples randomly collected were interpreted based on high resolution images of China's satellites GF-1, ZY-3. Accuracy of all land cover types was assessed. In general, the overall accuracy for all types is 88% and 89% in 2001 and 2013 respectively. And the Kappa coefficients are 0.86 and 0.87 respectively.

The dataset contains the boundaries of the world's 100 large wetlands of international importance and wetland classification results on them in 2001 and 2013 based on 250m

spatial resolution, 16-day composite product MOD13Q1. It provides important reference data for research on global wetlands and climate change.

Author contributions

Six coauthors were engaged in the design and development of the boundaries and remote sensing classification datasets on large wetlands of international importance in 2001 and 2013. Zhang, H. Y. and Niu, Z. G. participated in all the technological processes of the design and methodological development of boundaries extraction, data collection and processing, data analysis and validation, etc. Zhang, H. Y., Xu, P. P., Chen, Y. F., Hu, S. J. and Gong, N. took part in the boundaries extraction and wetlands classification. And the author list was arranged in their workload order. Zhang, H. Y. wrote this paper, and Niu, Z. G. was in charge of the paper proofreading, editing and revision.

References

[1] Costanza R., D'Arge R., De Groot R., *et al*. The value of the world's ecosystem services and natural capital [J]. *Nature*, 1997, 387: 253-260.
[2] Kloiber S. M., Macleod R. D., Smith A. J., *et al*. A semi-automated, multi-source data fusion update of a wetland inventory for East-Central Minnesota, USA [J]. *Wetlands*, 2015, 35: 335-348.
[3] Dong J., Xiao X., Kou W., *et al*. Tracking the dynamics of paddy rice planting area in 1986–2010 through time series Landsat images and phenology-based algorithms [J]. *Remote Sensing of Environment,* 2015, 160: 99-113.
[4] Zhang H. Y., Niu Z. G., Gong N., *et al*. The wetland cover datasets on the large wetlands of international importance in 2001 and 2013 by remote sensing data integration [DB]. *Global Change Research Data Publishing & Repository,* DOI:10.3974/geodb.2015.02.09.V1
[5] Global Change Research Data Publish & Repository Editorial Office. Global change research data publish & repository data sharing policy. DOI:10.3974/dp.policy. 2014.05.
[6] Ramsar Convention Secretariat. The Ramsar convention manual: A guide to the convention on wetlands (Ramsar, Iran, 1971), 4th edition [Z]. 2006.
[7] Jin Z., Xu B. A novel compound smoother-RMMEH to reconstruct MODIS NDVI time series [J]. *IEEE Geoscience and Remote Sensing Letters,* 2013, 10: 942-946.
[8] Chen J., Jönsson P., Tamura M., *et al*. A simple method for reconstructing a high-quality NDVI time-series data set based on the Savitzky-Golay filter [J]. *Remote Sensing of Environment,* 2004, 91: 332-344.