

# A Study of Forest Swamp Mapping in Hani Wetland Integrating Sentinel-1 and Sentinel-2 Satellite Images

Jingfa WANG<sup>1</sup>

College of Tourism and Geographical Science, Jilin Normal University, Siping, Jilin 136000, China

**Abstract.** As a unique wetland type, forest swamps play an important role in regional carbon cycling and biodiversity conservation. Taking Hani wetland in Jilin province as the research object, we integrated the application of Sentinel-1 radar and Sentinel-2 multispectral images, fully exploited the potential of Sentinel-1 multi-polarization band features and Sentinel-2 red edge index for forest swamp remote sensing identification, and applied the random forest method to realize the extraction of forest swamp distribution information of Hani wetland. The results show that when the optimal number of decision trees for forest swamp information extraction is 1200, the fusion of Sentinel-1VV and VH backscattering coefficient radar band features and Sentinel-2 red-edge band features can significantly improve the extraction accuracy of forest swamp distribution information, and the overall accuracy and Kappa coefficient of forest swamp information extraction in protected areas are as high as 89% and 0.85, respectively. The overall accuracy and Kappa coefficient of forest swamp information extraction in the protected area were 89% and 0.85, respectively. The landscape types of Hani Wetlands of International Importance are diversified, with natural wetlands, artificial wetlands and non-wetland landscape types co-existing. Among the natural wetland types, the forest swamp has the largest area of 27.1 km<sup>2</sup>, accounting for 11.2% of the total area of the reserve; the river has the smallest area of 0.7 km<sup>2</sup>, accounting for 0.3% of the total area of the reserve. The forest swamp extraction method provides data support for the sustainable management of Hani wetlands and case guidance for forest swamp mapping in other regions.

**Keywords.** forest swamp, Sentinel, random forest, mapping of remote sensing

## 1. Introduction

Wetlands are one of the most important types of terrestrial ecosystems and play an extremely important role in maintaining the balance of material cycles, protecting the diversity of biological species and maintaining ecological security [1-3]. Forested bogs are boggy wetlands with woody plants of more than 6 m as the dominant species, which are extremely rich in carbon stocks and respond significantly to global climate change [4-6]. Facing the severe challenges of both human activities and climate change, remote sensing monitoring of forest swamps can clarify the distribution boundaries of forest swamps in time, reflect the current status and transformation trends of forest

<sup>1</sup> Corresponding Author: College of Tourism and Geographical Science, Jilin Normal University, Siping, Jilin 136000, China; E-mail: jlnuwangjingfa@163.com.

swamp wetland landscape types, which is of great significance to effectively protect the health of forest swamp ecosystems and maintain the stability of regional ecosystems [7, 8]. Multi-source remote sensing data integration can maximize the fusion of different types of sensor remote sensing data, effectively improve the accuracy of feature type identification and extraction, and multi-source remote sensing data fusion is widely used in the extraction of wetland cover information research [9, 10]. With the rapid development of satellite remote sensing technology, the types of remote sensing data sources are constantly enriched. Sentinel-1 synthetic aperture radar (SAR) is different from passive optical remote sensing, which can effectively avoid the influence of clouds, rain, fog and no light at night, and has the advantages of high resolution, dual polarization, short period, strong penetration, multi-mode, all-weather, all-day and rapid product production, which can accurately determine the satellite position and attitude angle, and realize the high-resolution monitoring of global land, coastal zone, airline routes and large area coverage monitoring of global ocean [11, 12]. Sentinel-2 A and B satellite systems are equipped with Multispectral Instrument (MSI) sensors, which can carry out ground monitoring in 13 bands of visible, near-infrared, short-wave infrared and red-edge with a 5d cycle, especially the red-edge range of three bands has great advantages for vegetation growth and health monitoring, and the high spatial and temporal resolution and wide spectral resolution make Sentinel-2 remote sensing images commonly used in monitoring the evolution of terrestrial ecological environment, wetland landscape mapping, vegetation crop growth monitoring and disaster warning mapping [13, 14].

The Hani Wetland is the third wetland of international importance in Jilin Province and is extremely rich in freshwater, forest swamp and other wetland resources. It has the largest peat swamp deposit in northeast China, which is of great significance in terms of water conservation and provision, climate regulation and scientific research on peat swamp wetlands. This study takes Hani wetland in Jilin Province as the research object, integrates Sentinel-1 radar and Sentinel-2 multispectral images, fully exploits the potential of Sentinel-1 multi-polarization band feature and Sentinel-2 red edge index for forest swamp remote sensing identification, applies random forest method to realize the extraction of forest swamp distribution information in Hani wetland, and the forest swamp The extraction method can provide case guidance for forest swamp mapping in other regions and provide data support for sustainable management of Hani wetlands.

## 2. Study Area and Data Processing

### 2.1. Study Area

Jilin Hani Wetland ( $42^{\circ} 4' 12'' \sim 42^{\circ} 14' 30''$  N,  $126^{\circ} 4' 9'' \sim 126^{\circ} 33' 30''$  E) is located in the southeast of Liuhe County, Tonghua City, Jilin Province, adjacent to Tonghua County in the south, and bordered by Baishan City and Jingyu County in the east, with an approximate "northeast-southwest" strip distribution, including three parts: core area, buffer area and pilot area. The study area is located in the temperate humid and semi-humid monsoon climate zone, with four distinct seasons, high temperature and humidity in summer, cold and dry in winter, rain and heat in the same period, annual average temperature  $5.01^{\circ}\text{C}$ , annual average precipitation 755.52 mm,

annual frost-free period 125-145 d. The geomorphological type is mainly basalt lava mountains, with high terrain in the east and low terrain in the west, with an average altitude of 884.50 m.

## 2.2. Data Source and Processing

The Sentinel-1 Ground Range Detected (GRD) polarized radar image data applied in this study were downloaded from ESA Openeus, the image acquisition instrument was SAR-C, the collection mode was IW, the data size was about 1.64 G, and the data quality was high [15]. Sentinel-2 remote sensing images with rich spectral resolution are widely used in the fields of terrestrial vegetation, soil and water monitoring, grain yield estimation and wetland landscape mapping research, etc. The Sentinel-2 level-1C product was atmospherically corrected and radiometrically calibrated, and each band was super-resolved and resampled to 10 m [16].

In this study, 1942 training samples were obtained through visual interpretation and field GPS survey, of which 502 were the largest number of training samples from forested areas and 105 were the smallest number of training samples from paddy fields, which generally met the conditions of "uniform spatial distribution and sufficient number" and satisfied the experimental classification requirements. Based on the selected number of training samples, 50% of the training samples were selected as the set of validation points to test the final accuracy of the wetland classification results, with a total of 971 validation points, including 91 forest swamps.

With reference to the definition of wetlands in the international Convention on Wetlands and related literature, a classification system was established in line with the forest swamp wetland classification of Jilin Hani National Nature Reserve (Table 1).

**Table 1.** Classification system of swamp in the Hani Ramsar site.

The primary classification	The secondary classification	Definition and basis of division
Natural wetlands	River	A river having annual, seasonal, or intermittent river runoff
	Lake	A permanent, seasonal, or intermittent lake consisting of fresh, brackish, salt, or salt water
	Swamp	A freshwater marsh dominated by thickets or trees
	Marsh	A freshwater marsh consisting of dominant communities of aquatic or bog herbs
Artificial wetland	Paddy field	A ridge is enclosed to store water and plant paddy fields of one, two or three seasons
Non-wetlands	Dry farmland	Cultivated land without irrigation facilities that relies mainly on natural precipitation for xerophytic crops
	Woodland	Land covered by patches of natural, secondary, and artificial forests
	Grassland	Land on which herbaceous and shrub plants are predominant and suitable for animal husbandry production
	Construction land	Land for artificial construction of buildings and structures

### 3. Research Methodology

#### 3.1. Random Forest

Random Forest (RF) classification algorithm is a newly emerged, flexible and highly accurate predictive machine learning classification algorithm for calculators, which is commonly used in the extraction of wetland information from geological remote sensing images. Random Forest classification algorithm has the advantages of no worry about over-fitting, suitable for large data sets and a large number of unknown features, autonomous estimation of optimal feature sets, direct processing of high-dimensional feature input samples without dimensionality reduction, strong resistance to fabrication, easy to understand the algorithm content and the possibility of merging processing, etc. [17, 18].

#### 3.2. Waveform Extraction and Feature Set Combination

Based on Sentinel-1 radar and Sentinel-2 multispectral remote sensing images, Sentinel-1 VV and VH polarized backscatter coefficients were extracted, and band features such as Sentinel-2 spectral features, red-edge index features, vegetation index features, water body index features and texture features were obtained by Band Maths band calculation. .

#### 3.3. Precision Evaluation

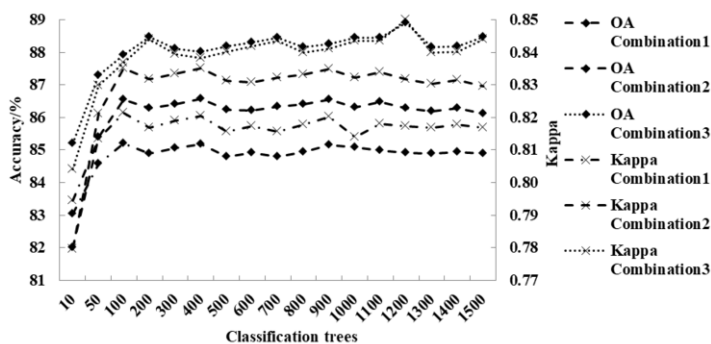
The study used the data of 971 validation points obtained from field measurements and the classification results to establish the confusion matrix of wetland information extraction in Hani Reserve 2020, and calculated four evaluation indexes of overall accuracy (*OA*), kappa coefficient, producer accuracy (*PA*) and user accuracy (*UA*) to accurately and reasonably evaluate the accuracy of this forest swamp information extraction.

## 4. Results and Analysis

#### 4.1. Optimal Random Forest Classification Tree

The overall accuracy and Kappa coefficient of random forest extraction of wetland cover information in Hani Reserve showed a trend of fluctuating and stabilizing after first increasing at the stage of 10~100 classification trees (Figure 1), and the overall accuracy of wetland information extraction in the reserve increased by 2.6%, 5.5%, and 3.2% at the stage of 10~100 classification trees for combinations 1, 2, and 3, respectively, and the Kappa coefficients increased by 0.03, 0.06, and 0.03, respectively. 0.06 and 0.03, respectively, and the kappa coefficients increased by 0.03, 0.06, and 0.03, respectively, with the increase of the number of classification trees, the wetland information extraction time of protected areas also increased accordingly. The optimal number of random forest classification trees for forest swamp information extraction was different for the three combinations of protected areas, and the optimal random

forest classification trees were 100, 400, and 1200 for combinations 1, 2, and 3, respectively.



**Figure 1.** Varied classification accuracies among the different numbers of classification trees for different feature combinations.

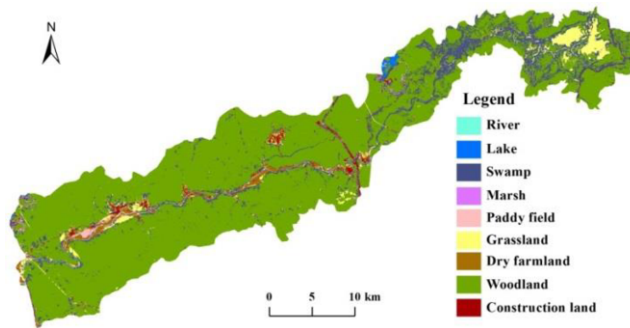
#### 4.2. Optimal Combination of Features

The combination of Sentinel-1 radar VV and VH backscatter coefficient bands with Sentinel-2 red-edge bands3 forest swamp information extraction accuracy is the highest, the overall accuracy and Kappa coefficient are as high as 88.9%, 0.85, where the producer accuracy of forest land, forest swamp, dry land, and grassland are as high as 97.9%, 93.8%, 90.6%, and 90.1%, and the user accuracies were as high as 96.6%, 89.5%, 89.4%, and 86.9% for woodland, paddy field, forest swamp, and grassland, respectively, with relatively high extraction quality for forest swamp, followed by combination 2 which only incorporates Sentinel-2 red-edge bands, and the smallest overall accuracy and Kappa coefficient for combination 1 wetland information extraction.

#### 4.3. Spatial Patterns of Hani Internationally Important Wetlands

The total area of Hani Reserve is about 242.5 km<sup>2</sup> (Figure 2), among which, the natural wetland area is 30.6 km<sup>2</sup>, and the forest swamp is the largest among the natural wetland types, 27.1 km<sup>2</sup>, accounting for 88.6% of the natural wetland area and 11.2% of the total area of the reserve, mainly distributed in the relatively high altitude area in the northeast of the reserve; the river area is the smallest, 0.7 km<sup>2</sup>, accounting for only 2.2% of the natural wetland area and 0.3% of the total area of the reserve. The area of rivers is the smallest, 0.7 km<sup>2</sup>, accounting for only 2.2% of the area of natural wetlands and 0.3% of the total area of the reserve, running from northeast to southwest; the area of lakes and herbaceous marshes is 1.3 km<sup>2</sup> and 1.5 km<sup>2</sup> respectively, accounting for 4.4% and 4.9% of the area of natural wetlands and 0.6% and 0.6% of the total area of the reserve respectively. The area of artificial wetland paddy field is 0.9 km<sup>2</sup>, which is 0.4% of the total area of the reserve, mainly distributed in the gentle terrain on both sides of the river; the area of non-wetland types in the reserve is relatively large, and the area of forest land is the largest, up to 192.3 km<sup>2</sup>, which is 79.3% of the total area of the reserve; the area of construction land is the smallest, 3.6 km<sup>2</sup>, which is 1.5% of the total area of the reserve; the area of dry land and grassland is 5.2 km<sup>2</sup>, respectively. The area

of dry land and grassland is 5.2 km<sup>2</sup> and 9.9 km<sup>2</sup> respectively, accounting for 2.2% and 4.1% of the total area of the reserve.



**Figure 2.** Land cover distribution patterns in the Hani Ramsar wetland site.

## 5. Conclusion

The accuracy of the random forest method for extracting forest swamp and wetland landscape information is governed by various factors such as the type of study area, the number of classification trees, and the selection of training and validation samples, etc. The optimal random forest classification tree for extracting forest swamp and wetland landscape information of Hani International Important Wetland is 1200.

The fusion of Sentinel-1 VV and VH backscatter polarization and Sentinel-2 red-edge feature bands can significantly improve the accuracy of forest swamp landscape information extraction, the overall accuracy of Hani internationally important wetland landscape information extraction is 88.9%, the Kappa coefficient is 0.85, and the accuracy of forest swamp extraction producers and users are as high as 93.8% and 89.4%, respectively.

The wetland landscape types in the reserve are diverse, with natural wetlands, artificial wetlands and non-wetland landscape types co-existing. Among the natural wetland types, the forest swamp has the largest area of 27.1km<sup>2</sup>, accounting for 11.2% of the total area of the reserve, and is mainly distributed in the relatively high altitude area in the northeast of the reserve, while the river has the smallest area.

## Acknowledgements

This research was funded by National Natural Science Foundation of China (NO.41871022).

## References

- [1] Jin BCA., Jiang MY, Bai WR et al. Fluorescence spectra characteristics of reclaimed water to replenish constructed wetlands using EEM-PARAFAC, Spectroscopy and Spectral Analysis 41(2021), 1240-1245.
- [2] Islam AMT, Talukdar S, Mahato S, et al. Machine learning algorithm-based risk assessment of riparian wetlands in Padma River Basin of Northwest Bangladesh. Environmental Science and Pollution Research 28(2021), 34450-34471.

- [3] Song F, Su FL, Mi CX, et al. Analysis of driving forces on wetland ecosystem services value change: A case in Northeast China, *Science of The Total Environment* 751(2021), 141778.
- [4] Xulu S, Mbatha N, Peerbhay K, et al. Detecting harvest events in plantation forest using sentinel-1 and-2 data via google earth engine, *Forests* 11(2020), 1283.
- [5] Dong D, Wang C, Yan JH, et al. Combing Sentinel-1 and Sentinel-2 image time series for invasive *Spartina alterniflora* mapping on Google Earth Engine: a case study in Zhangjiang Estuary. *Journal of Applied Remote Sensing* 14(2020), 044504.
- [6] Mohammadi A, Kamran KV, Karimzadeh S, et al. Flood detection and susceptibility mapping using sentinel-1 time series, alternating decision trees, and bag-adtree models. *Complexity*, (2020), 4271376.
- [7] Pyankov SV, Maximovich NG, Khayrulina EA et al. Monitoring acid mine drainage's effects on surface water in the kizel coal basin with sentinel-2 satellite images. *Mine Water and the Environment*, 40(2021), 606-621.
- [8] Yang QQ, Yuan QQ, Li TW, et al. 2020. Mapping PM2.5 concentration at high resolution using a cascade random forest based downscaling model: Evaluation and application, *Journal of Cleaner Production*, 277(2020), 123887.
- [9] Braun A. Retrieval of digital elevation models from Sentinel-1 radar data - open applications, techniques, and limitations, *Open Geosciences*, 13(2021), 532-569.
- [10] Rumora L, Miler M, Medak D. Contemporary comparative assessment of atmospheric correction influence on radiometric indices between Sentinel-2A and Landsat 8 imagery, *Geocarto International*, 36(2021), 13-27.
- [11] Yahiaoui I, Bradai A, Douaoui A, et al. Performance of random forest and buffer analysis of Sentinel-2 data for modelling soil salinity in the Lower-Chelif plain (Algeria), *International Journal of Remote Sensing*, 42(2021), 128-151.
- [12] Harper KA, Gray L, Querry ND. Spatial patterns of vegetation structure and structural diversity across edges between forested wetlands and upland forest in Atlantic Canada, *Canadian Journal of Forest Research*, 51(2021), 1189-1198.
- [13] Santos CF, Teixeira RC, Raizer J, et al. Post-fire phyllostomid assemblages in forest patches of the Pantanal wetland, *Mammalia*, 85(2021), 155-158.
- [14] Bonansea M, Ledesma M, Bazan R, et al. Evaluating the feasibility of using Sentinel-2 imagery for water clarity assessment in a reservoir, *Journal of South American Earth Sciences*, 95(2019), 102265.
- [15] Yousef B, Hossein H, Abbas M. Investigating the capabilities of multispectral remote sensors data to map alteration zones in the Abhar area, NW Iran, *Geosystem Engineering*, 24(2021), 18-30.
- [16] Clauss K, Ottinger M, Leinenkugel P, et al. 2018. Estimating rice production in the Mekong Delta, Vietnam, utilizing time series of Sentinel-1 SAR data, *International Journal of Applied Earth Observation and Geoinformation*, 73(2018), 574-585.
- [17] Chen ZT, Wang JF, Gao DY, et al. Dynamic spatial fusion of cloud vertical phase from CALIPSO and cloudsat satellite data. *Photogrammetric Engineering and Remote Sensing*, 87(2021), 61-67.
- [18] Chen R 2021. Application of UAV-low altitude remote sensing system in sea area supervision, *Earth Sciences Research Journal*, 25(2021), 65-68.