

HYPERSPECTRAL CHANGE DETECTION IN WETLAND AND WATER-BODY AREAS BASED ON MACHINE LEARNING

Marjan Ahangarha ^a, Seyd Teymoor Seydi ^a, Reza Shahhoseini ^{a*}

^a School of surveying and Geospatial Engineering, University of Tehran, Tehran, Iran
(ahangarha.marjan, seydi.teymoore, rshahosseini)@ut.ac.ir

Commission VI, WG VI/4

KEY WORDS: Wet land, Machine learning, Change Detection, Random forest, Hyperspectral

ABSTRACT:

Wetlands and water-bodies are transitional lands between terrestrial and aquatic ecosystems that give many advantages. The presence of phenomena activities and artificial human activities cause changes on the Earth's surface. This process causes the changes in land cover type especially in wetlands and water-bodies' area. For monitoring and assessing resources like wetlands and water-bodies, it is necessary to be aware of these changes. Change detection and attribution of wetlands and water-bodies change over time present more challenges for correctly analysing remote sensing imagery. Hyperspectral images now have potential applications in many scientific areas due to their high spectral resolution and so their good information contents. The aim of this study is to propose a procedure for determining land surface changes within the semi-arid wetland and surrounding upland areas using the new method by combining machine Learning method for detecting change using EO-1 Hyperion satellite hyperspectral imagery. The study area is Shadegan wetlands in the south-west of Iran in Khuzestan province. The most critical water resources of the province are depleted and contain unprecedented levels of toxic waste. In addition, the results show the superiority of the implemented method to extract change map with overall accuracy by a margin of nearly 94% using multi-temporal hyperspectral.

1. INTRODUCTION

Remote sensing data has specializad capabilities because of its integrity and comprehensive range, spectral variety, continuous coverage and low cost, compared to other methods of data acquisition, which is the first factor in studying the surface area today. (Brando & Dekker, 2003) Remote sensing data has a lot of application such as monitoring, classification, target detection, change detection etc. Nowadays, with the progressive of technology and it has been possible to hyperspectral datasets with many bands in order to, there are many sensors such as Hyperion, and AVIRIS (Singh, 1989). A new generation of these sensors such as Enmap and Hydice will be launch in the future, as well as improving data quality and improving temporal resolution (Rickard et al, 1993). Hyperspectral images now have potential applications in many scientific areas due to their high spectral resolution and so their good information contents (Hasanlou et al., 2018). Change detection is one of the most essential applications of satellite images. Change detection is done with the purpose of analyzing two or more images of a region that has been obtaining at different times (Marinelli et al., 2019). We can consider two categories methods for changing: 1) supervised based on classification supervised and 2) unsupervised method. In supervised methods need to training data or threshold selection while unsupervised methods don't need to threshold selection or training data (Hasanlou and Seydi, 2018b).

Wetlands are areas where water covers the soil or is present either at or near the surface of the earth all year or for varying periods during the year, including during the growing season. Change detection and attribution of wetlands and water-bodies change over time present more challenges for correctly analysing remote sensing imagery. According to the climate and weather changes, decreasing water sources, earth warm up, Environmental puddles in the world, keeping, and protection of wetland and water bodies is very important (Hakdaoui et al., 2019; Hasanlou and Seydi,

2018a; Kaplan et al., 2019; Li et al., 2019a). Remote sensing systems and satellites are growing to a high spectral resolution that their suitable for investigating changes of wetland. So, the wetland and waterbody areas are sensitive areas and need to accurately monitor. for this end, the many methods are proposed for wetland monitoring based on remote sensing imagery (Klema, 2011). Since the Shadegan, the wetland is one of the most crucial wet lands in the world that registered at UNESCO. For environmental changing detection with the goal of protecting and reviving, them in the long that is important. ZHAO et al investigated wetland change detection in the Pearl River Estuary from 1979 to 2009. They used classifier images for generating maps of wetland degradation. This area was complicated because natural wetlands changed to artificial wetlands and vice versa. They just got characteristics of estuarine wetland and the connection between wetland change and urban land. Moreover, they used Landsat dataset and their highest Kappa accuracy was 85% (Zhao et al., 2010). Ludwig et al used a method based on automated monitoring wetland by using multi-temporal optical satellite data. They investigated three sites. The results show their way is appropriate for identifying water and wetness areas. Since they used a different satellite sensor is a limitation because deviations occur in their conclusion (Ludwig et al., 2019). Barducci et al monitored wetlands by using hyperspectral sensors. They do not have any accuracy assessment, their datasets have not a high spectral and spatial resolution that is a limitation, and you need a way that has more information and precision for management the wetland areas (Barducci et al, 2009). Fickas et al used Landsat images based on monitoring of seasonal wetland from 1972 to 2012. Their result shows 91% accuracy but Landsat time series dataset does not have a high spectral resolution and they interpreted google earth images in addition to Landsat. Another limitation was about wetland classes because Landsat does not have the same spatial resolution and cause of mixed type (Fickas et al,

* Corresponding author

2016). Baker et al used Landsat images in the Gallatin valley from 1988 to 2001. They used SGB and CVA methods for classifying and finding wetland areas. They had some errors in the classification of the wetland boundaries because they used Landsat data that it does not have a high spectral resolution. The overall accuracy of the CVA method was 76% that is not adequate and their processes were abolished (Baker et al, 2007). Kuleli et al used multi-temporal satellite image for shoreline change detection with an automatic way. They investigated Landsat images for shoreline changes on coastal wetlands in Turkey. Their results have not any quantity assessment and they just used a digital shoreline analysis system as a statistical way for changing coastline rate. (Kuleli et al, 2011). Hyperspectral sensors can be monitor objects from long distance for lots of purposes like improving detection.

In this way, machine learning is one of the best algorithms for modern image analysis. The definition of machine learning from 1960 started that has considerable usages in the real world such as, Speech recognition, Computer vision, Bi-surveillance, Robot control, Accelerating empirical sciences, Medical Diagnosis, Prediction, Classification etc (Mitchell, 2006). We named some of the advantages in machine learning like, quickly identifies trends and patterns, No human intervention needed, Continues improvement, Handling multi-dimensional and multi-variety data (Guo et al, 2017). Machine learning methods have five categories and we choose random forest that a supervised approach (Gewali et al, 2018).

The need for changes continuous monitoring of using is felt by human and inhuman. Since the dimension reduction is one of the challenges of the hyperspectral dataset for solving this problem, they already used methods based on shallow learning machine such as a support vector machine and that have limited.

Generally, they methods had been many challenges such as: (1) the used dataset focused on multispectral that has low spectral

resolution, (2) the mentioned proposed methods have low accuracy, (3) more techniques are supervised and need to threshold selection, training data while collecting training data is time consuming and complicated, and (4) some ways are very sophisticated and has high computing. For this end, current research presented a novel framework for wetland and waterbody change detection using hyperspectral imagery. Our proposed method is applied in three main steps: the first step are image differencing. In the second step, pseudo training data is generated by Expectation Maximization (EM) based classification algorithm. In phase 2 optimization of the random forest classifier parameters are performed, and generation binary change map.

This paper is organized as follows: Section 2 the details of the proposed method. Section 3 introduces study areas and hyperspectral datasets. Chapter 4 provides the evaluation results, and the final Section 5 includes the conclusion of experimentation.

2. PROPOSED METHOD

The proposed method has been done in three phases. At first, the pre-processing is applied to bi-temporal hyperspectral imagery. In the next step, the difference images (DI) is used to the hyperspectral dataset. This process caused to the change pixels give high intensity, also, no-change pixels get low value and nearby zero. The third step is named pseudo training data that it is obtained based on the summation of DI and EM based classification. Due to the increasing reliance on training sample is applied a hierarchical classification based on EM algorithm. Then, the sample data divided into two part, test and training example that is needed to optimize the random forest (RF) parameters. Finally, the change map is obtained based on RF classifier on DI pixels. The RF parameters are optimized by the grid search algorithm.

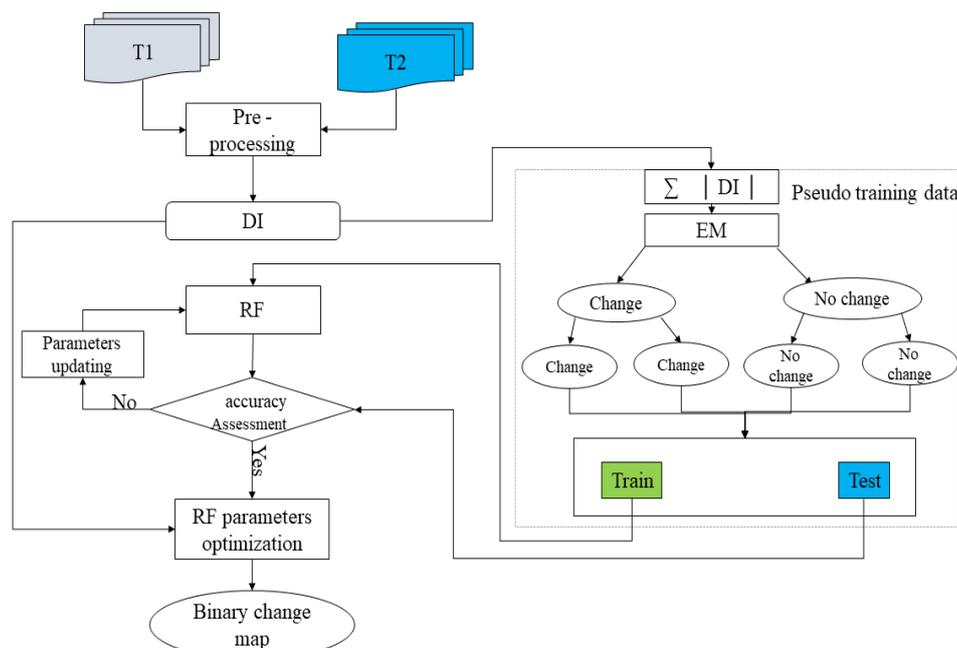


Figure 1. Proposed method flowchart

2.1 Image Differencing Algorithm

The ID algorithm is one most common algorithm in the change detection process. This algorithm used pure mathematical for measurement of spectral properties for two different times. For a grayscale image $H(i, j)$ $P \times Q$ pixels in size, a difference image $DI(i, j)$, $P \times (Q-1)$ pixels in size can be generated from the original image H by using Formula 1 (Lin et al, 2008):

$$DI(i, j) = |H(i, j) - H(i, j + 1)|, \quad 0 \leq i \leq P - 1, \\ 0 \leq j \leq Q - 2, \quad (1)$$

2.2 Expectation Maximization based Classifications

The EM is an iteration method that extracts density in an unsupervised framework. In this way, estimated model initial parameters and then subsequent iterations their updated. This process progresses until convergence time (Rohlfing et al, 2004)

2.3 Pseudo Training Data

The primary purpose of this section, is to generate the training sample in an automatic framework. For this end, at first, the image-differencing algorithm is applied. This process caused to change areas differs from no-change areas. At the next step, the EM algorithm is used as hierarchically. The use of hierarchically EM caused to increase the reliability of the training data. The extracted training sample is divided into two parts: training data and testing data. The all of the training data used for the optimization of RF classifier.

2.4 Random Forest

Random Forest algorithm proposed by Breiman in 2001 (Breiman, 2001). This algorithm is a new kind of none parametric machine learning method that is better than many single forecasting methods. Random forest is a method that used tree-stretcher classifier which gets maximum the predict results with average or judgment on the predict values of a number of them. The RF method consists of two sections training and classification. The training step is to get a classification model by using sample data and a decision tree (Pino-Mejías et al., 2010). The RF is a combine of a decision tree that each decision gives a vote to assign the highest class to an input vector that to allocate each pixel to a class based on the maximum number of the votes, which receives, from classifier group. The critical point of the random forest algorithm is to complex calculations simplify by reducing the number of input features per node. In the random forest algorithm, two parameters need to be determined: the number of decision trees that need to be created and the number of features to each node in a tree (Pal, 2005). This algorithm increases the variety of decision trees to enhance them by modifying the training set. The decision tree divides the training samples into smaller nodes at each node by using decision rules.

2.5 Pre-processing hyperspectral data

Preprocesses of hyperspectral images is performed in two levels Geometric and spectral correction. Spectral correction is related to digital number (DN) included de stripping, noise reduction, Correction of spectral curvature effect, radiometric conversion, Atmospheric correction etc. Geometric correction is related to spectral location pixels that used control points between 2 images to georeferencing them. In addition, with the linear interpolation

method the mean square error is about 0.4 pixels (Hasanlou et al, 2018).

3. EXPERIMENTAL RESULT

3.1 Study area

This section shows details of the case study and data used for wetland changing detection. Shadegan wetland with a surface area of 5377 km², which is a vast ecological system consists of salty and sweet marshes, with the tidal bay of Khormosi and that islands. This wetland is located downstream of the Jarahi River between the cities of Shadegan, Abadan, Mahshahr, Ramhormoz and Ahvaz in Khuzestan province. The swamp is of ecosystem, aesthetic, environmental, economic and employment importance to residents. Figure 2 shows the location of our study area.

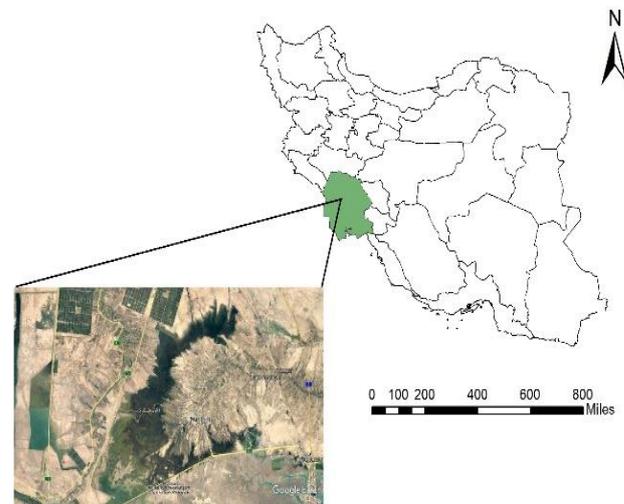
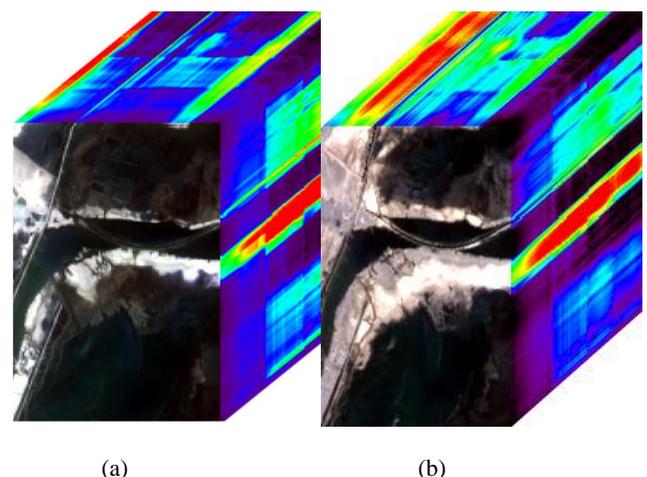


Figure 2. The Study area of Shadegan

3.2 Remote Sensing Data

In this study, hyperspectral image dataset is used for analysing changes in wetlands and water bodies illustrated in (Figure 3).



(a)

(b)

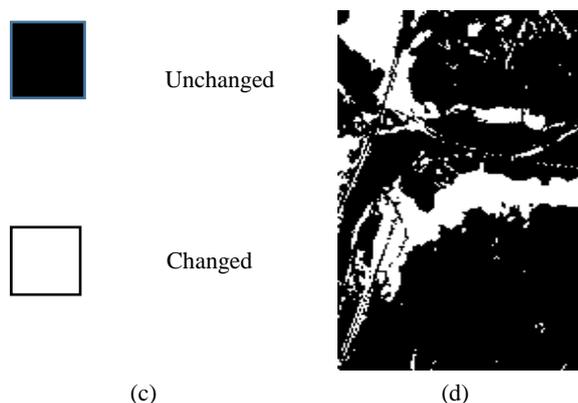


Figure 3. The (a) and (b) presented false-colour composite of the original hyperspectral images acquired in 2006 and 2006 of Shadegan wetland in Iran respectively. The (c) and (d) showed Ground truth with the legend.

The Hyperion sensor is located on the EO-1 satellite and is a space board device. The specifications of this sensor are given in (Table 1) Shadegan wetland is one of the largest wetland in Iran. The datasets acquired on 6 June 2006 and 29 June 2006. The number of bands was 154.

Parameters	Hyperion	
	Spectral Range	0.4-25
Spatial Resolution	30	m
Swath Width	7.5	Km
Spectral Resolution	10	nm
Spectral Coverage	Continuous	
Pan Band Resolution	N/A	
Total Number of Bands	220	

Table 1. Specifications of Hyperion Hyperspectral sensor

3.3 Results and Discussion

In this section, you can see quantitative results that were used overall accuracy and kappa coefficient for analysing. We used a popular change detection method in comparison proposed method such as principal component analysis (PCA) and iteratively regularized multivariate alteration detection (IR-MAD).

Figure 4 shows the result of the proposed method with other methods. Based on this figure, all of the methods are detected main change areas that this show potential hyperspectral imagery for change detection. In addition, the proposed method is detected more changes pixels while other methods have many miss-detect pixels or false pixels.

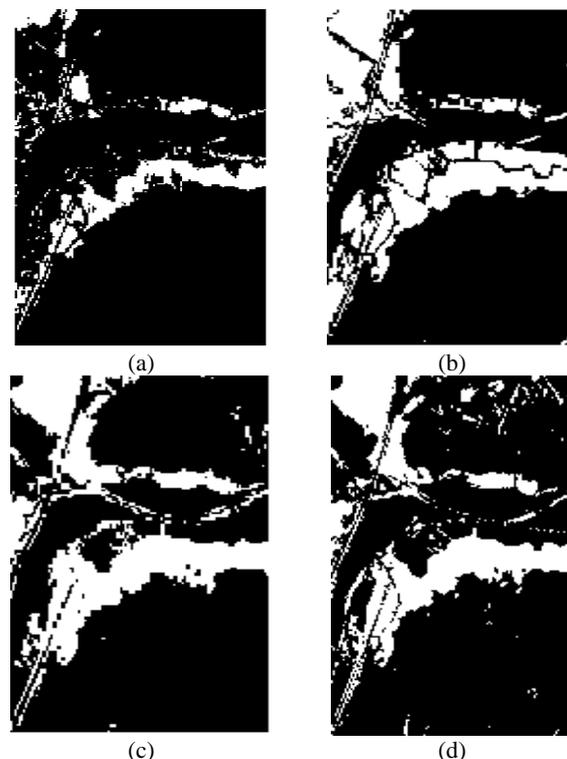


Figure 4. The results of CD methods for Shadegan wetland data set. (a) IR-MAD, (b) PCA, (c) Proposed method-binary change the map, (d) Ground truth.

In order to prove the efficiency of the proposed method, Table 2 shows the results of quantitative analysis of each technique. According to the results of the proposed method, it is more accurate than other methods. In addition, use of the IR-MAD way can improve change detection in this area. In addition, the PCA method has an excellent performance but the quantitative and qualitative result of the proposed method shows outstanding performance.

Table 2. Performance of the proposed method and other conventional CD methods for all data sets.

Data-sets	Indices	IR-MAD	PCA	Proposed Method
Shadegan	Overall	86.06	81.81	93.33
Wetland	Kappa	0.49	0.485	0.811

The both of analysis shows the proposed method has excellent ability for monitoring wetland and waterbody areas. In addition, the hyperspectral imagery can be used as a suitable dataset for this purpose but need to a unique technique for information extraction.

4. CONCLUSION

Since in Iran, wetland and water bodies are going to become dry in the last decade, thus it is vital to preserve them. In this paper, we have presented based on machine learning-based change detection framework for wetland and water bodies change monitoring. One of the challenges in wetland areas is spectral class separability which we have solved this problem to some extent by using the proposed method. Also, existence a robust plan for monitoring change area. The aim of this study was to have a highly automated and novel framework method. Based on evaluation it can be seen that, hyperspectral images have a good potential for monitoring changes in water bodies and wetland. Comparison to other change detection approaches that have been used for wetland detection shows that proposed method are many performances such as: Do not need to training data and threshold selection, high potential for extraction change information, high accuracy, and is pure for implementation. In addition, the proposed method has presented the accuracy of more than 90% that it is excellent compare to other mentioned change detection methods.

Future work

Although in this research used the hyperspectral dataset also they have excellent capability for monitoring wetland and water body but they have a low spatial resolution. This challenge can be dissolved with a new generation of hyperspectral datasets.

REFERENCES

- Hakdaoui, S., Emran, A., Pradhan, B., Lee, C.-W., Fils, N., Cesar, S., 2019. A Collaborative Change Detection Approach on Multi-Sensor Spatial Imagery for Desert Wetland Monitoring after a Flash Flood in Southern Morocco. *Remote Sensing* 11, 1042.
- Hasanlou, M., Seydi, S.T., 2018a. Automatic change detection in remotely sensed hyperspectral imagery (Case study: wetlands and waterbodies).
- Hasanlou, M., Seydi, S.T., 2018b. Hyperspectral change detection: An experimental comparative study. *International journal of remote sensing* 39, 7029-7083.
- Hasanlou, M., Seydi, S.T., Shah-Hosseini, R., 2018. A Sub-Pixel Multiple Change Detection Approach for Hyperspectral Imagery. *Canadian Journal of Remote Sensing* 44, 601-615.
- Kaplan, G., Avdan, Z.Y., Avdan, U., 2019. Mapping and Monitoring Wetland Dynamics Using Thermal, Optical, and SAR Remote Sensing Data. *Wetlands Management: Assessing Risk and Sustainable Solutions*, 87.
- Li, N., Li, L., Lu, D., Zhang, Y., Wu, M., 2019a. Detection of coastal wetland change in China: a case study in Hangzhou Bay. *Wetlands Ecology and Management* 27, 103-124.
- Li, S., Song, W., Fang, L., Chen, Y., Ghamisi, P., Benediktsson, J.A., 2019b. Deep Learning for Hyperspectral Image Classification: An Overview. *IEEE Transactions on Geoscience and Remote Sensing*.
- Marinelli, D., Bovolo, F., Bruzzone, L., 2019. A Novel Change Detection Method for Multitemporal Hyperspectral Images Based on Binary Hyperspectral Change Vectors. *IEEE Transactions on Geoscience and Remote Sensing*.
- Baker, C., Lawrence, R. L., Montagne, C., & Patten, D. (2007). Change detection of wetland ecosystems using Landsat imagery and change vector analysis. *Wetlands*, 27(3), 610.
- Barducci, A., Guzzi, D., Marcoionni, P., & Pippi, I. (2009). Aerospace wetland monitoring by hyperspectral imaging sensors: A case study in the coastal zone of San Rossore Natural Park. *Journal of environmental management*, 90(7), 2278-2286.
- Bhatt, A., Ghosh, S., & Kumar, A. (2015). Automated change detection in satellite images using machine learning algorithms for Delhi, India. Paper presented at the 2015 IEEE International Geoscience and Remote Sensing Symposium (IGARSS).
- Bhavsar, H., & Ganatra, A. (2012). A comparative study of training algorithms for supervised machine learning. *International Journal of Soft Computing and Engineering (IJSCE)*, 2(4), 2231-2307.
- Brando, V. E., & Dekker, A. G. (2003). Satellite hyperspectral remote sensing for estimating estuarine and coastal water quality. *IEEE transactions on geoscience and remote sensing*, 41(6), 1378-1387.
- Breiman, L. (2001). Random forests. *Machine learning*, 45(1), 5-32.
- Fickas, K. C., Cohen, W. B., & Yang, Z. (2016). Landsat-based monitoring of annual wetland change in the Willamette Valley of Oregon, USA from 1972 to 2012. *Wetlands ecology and management*, 24(1), 73-92.
- Gewali, U. B., Monteiro, S. T., & Saber, E. (2018). Machine learning based hyperspectral image analysis: a survey. arXiv preprint arXiv:1802.08701.
- Guo, M., Li, J., Sheng, C., Xu, J., & Wu, L. (2017). A review of wetland remote sensing. *Sensors*, 17(4), 777.
- Klemas, V. (2011). Remote sensing of wetlands: case studies comparing practical techniques. *Journal of Coastal Research*, 27(3), 418-427.
- Kotsiantis, S. B., Zaharakis, I., & Pintelas, P. (2007). Supervised machine learning: A review of classification techniques. *Emerging artificial intelligence applications in computer engineering*, 160, 3-24.
- Kuleli, T., Guneroglu, A., Karsli, F., & Dihkan, M. (2011). Automatic detection of shoreline change on coastal Ramsar wetlands of Turkey. *Ocean Engineering*, 38(10), 1141-1149.
- Lin, C.-C., Tai, W.-L., & Chang, C.-C. (2008). Multilevel reversible data hiding based on histogram modification of difference images. *Pattern Recognition*, 41(12), 3582-3591.
- Ludwig, C., Walli, A., Schleicher, C., Weichselbaum, J., & Riffler, M. (2019). A highly automated algorithm for wetland detection using multi-temporal optical satellite data. *Remote Sensing of Environment*, 224, 333-351.
- Mitchell, T. M. (2006). *The discipline of machine learning* (Vol. 9): Carnegie Mellon University, School of Computer Science, Machine Learning
- Muro, J., Canty, M., Conradsen, K., Hüttich, C., Nielsen, A., Skriver, H., . . . Menz, G. (2016). Short-term change detection in

wetlands using Sentinel-1 time series. *Remote Sensing*, 8(10), 795.

Pal, M. (2005). Random forest classifier for remote sensing classification. *International Journal of Remote Sensing*, 26(1), 217-222.

Pino-Mejías, R., Cubiles-de-la-Vega, M. D., Anaya-Romero, M., Pascual-Acosta, A., Jordán-López, A., & Bellinfante-Crocci, N. (2010). Predicting the potential habitat of oaks with data mining models and the R system. *Environmental Modelling & Software*, 25(7), 826-836.

Rickard, L. J., Basedow, R. W., Zalewski, E. F., Silverglate, P. R., & Landers, M. (1993). HYDICE: An airborne system for hyperspectral imaging. Paper presented at the Imaging Spectrometry of the Terrestrial Environment.

Rohlfing, T., Russakoff, D. B., & Maurer, C. R. (2004). Performance-based classifier combination in atlas-based image segmentation using expectation-maximization parameter estimation. *IEEE transactions on medical imaging*, 23(8), 983-994.

Singh, A. (1989). Review article digital change detection techniques using remotely-sensed data. *International journal of remote sensing*, 10(6), 989-1003.

Wuest, T., Weimer, D., Irgens, C., & Thoben, K.-D. (2016). Machine learning in manufacturing: advantages, challenges, and applications. *Production & Manufacturing Research*, 4(1), 23-45.

Xin, W., Can, T., Wei, W., & Ji, L. (2019). Change Detection of Water Resources via Remote Sensing: An LV-NSCT Approach. *Applied Sciences*, 9(6), 1223.

Zhao, H., Cui, B., Zhang, H., Fan, X., Zhang, Z., & Lei, X. (2010). A landscape approach for wetland change detection (1979-2009) in the Pearl River Estuary. *Procedia Environmental Sciences*, 2, 1265-1278.