

WATER SALINITY MAPPING OF KARUN BASIN LOCATED IN IRAN USING THE SVR METHOD

M. Ansari ¹, M. Akhoondzadeh ^{2,*}

¹Remote Sensing Division, School of Surveying and Geospatial Engineering, College of Engineering, University of Tehran, Iran - moh.ansari@ut.ac.ir

²Remote Sensing Division, School of Surveying and Geospatial Engineering, College of Engineering, University of Tehran, Iran - makhoz@ut.ac.ir

KEY WORDS: Water salinity, Landsat-8 OLI, SVR, GA, Remote sensing, Karun basin, Spectral signature

ABSTRACT:

Water salinity is a complex issue in coastal and estuarine areas. Currently, remote sensing techniques have been widely used to monitor water quality changes, ranging from river to oceans. The salinity of Karun River has been increasing due to some critical factors, therefore, This study aimed at building regression models to ascertain the water salinity through the relationship between the reflectance of the Landsat -8 OLI and In situ measurements. A total of 102 observed samples were divided into 70% training and 30% test from June 2013 to July 2018 along the Karun River. Spectral signature analysis showed that band 1- Coastal/Aerosol (0.433 – 0.453 μm), band 2-Blue (0.450 – 0.515 μm) and band 3 – Green (0.525 – 0.600 μm) are sensitive to salinity. Furthermore, to have a comprehensive investigation, the Support Vector Regression (SVR) method was applied. The outcomes related to the quality of the SVR depend on several factors e.g. proper setting of the SVR meta-parameters, therefore, to deal with this issue Genetic Algorithm (GA) was applied. The SVR model resulted in values of R^2 and RMSE for test data which are respectively obtained to be 0.7 and $390\mu\text{s cm}^{-1}$. Eventually, Karun water salinity maps were prepared by SVR method to demonstrate the Karun water salinity on 1 February 2015 and 5 September 2018.

1. INTRODUCTION

The Karun River is the largest river in Iran (approximately $67,500\text{ km}^2$) which is located in the south west of Iran between longitudes $48^\circ 15'$ and $52^\circ 15'$ east, latitude $30^\circ 17'$ and $33^\circ 49'$ north (Keshavarzi et al., 2015). Karun River supplies water demands of 16 cities and ultimately sink into the Persian Gulf (Naddafi et al., 2007), Figure 1 shows the location of the case study. Karun River is polluted because of adverse climate condition and regional physiography, industrial sources, domestic and urban sewerage, irrigation of agricultural land, fish hatchery, hospital sewage and high tide level of Persian Gulf (Naddafi et al., 2007).

In this study, Landsat 8 Operational Land Imager (OLI) data was used to retrieve the water salinity map for the case of Karun River since not only it is free, but it also has an acceptable spatial, temporal and spectral resolution. Moreover, (Zhao et al., 2017) reported that the Landsat-8 OLI's bands 1, 2, 3 and 4 are sensitive with salinity. In addition, previous attempts at salinity modeling by OLI have implemented different band combinations, e.g. OLI bands 2, 3, 4, and 7 (Nguyen et al., 2018). Recently, many authors provided a water salinity model, which is calculated using various regression models, e.g. Geographically Weighted Regression (GWR) technique (Xie et al., 2013), Spatially Weighted Optimization Model (SWOM) technique (Khadim et al., 2017) and Multiple Linear Regression (MLR), Decision Trees (DT) and Random Forest (RF) techniques (Nguyen et al., 2018).

Detecting of water salinity would be a challenging task when the number of observations might not be sufficient (Nguyen et al., 2018). To deal with this problem, in this study 102 In situ

measurements (Electrical Conductivity) were collected from 8 stations from June 2013 to July 2018 along the Karun River (see Figure 1). In this study the spectral signature was used to find the best band combination for modeling water salinity. Eventually, the Support Vector Regression (SVR) were applied in the present study to demonstrate the Karun water salinity on 1 February 2015 and 5 September 2018.

The structure of the manuscript consists of four sections. This section (section 1) provides some information about motivations, and comprehensive literature review. Section 2 explains data collection and investigates the research methods, and regression methods. Section 3 presents the results of the research. At last, Section 4 discusses the achieved results and provides the conclusions of the study.

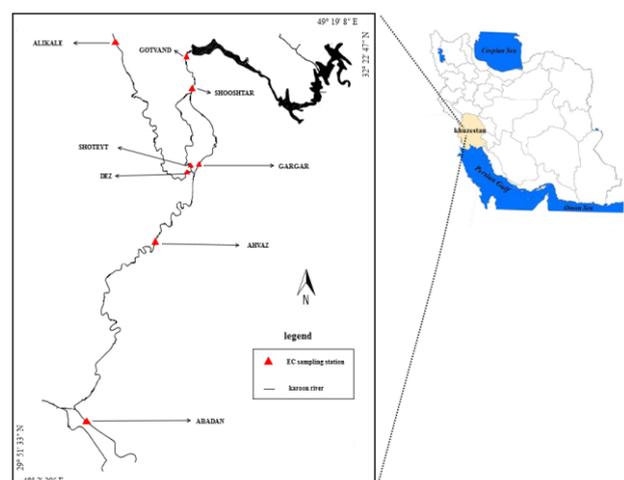


Figure 1. The study area and location of 8 EC sampling station (southwest of Iran) (Karamouz et al., 2009; Keshavarzi et al., 2015)

2. METHODOLOGY

2.1 Data Collection

The Landsat-8 OLI has an entire Earth coverage every 16 days and 30 m × 30 m of spatial resolution. 36 Landsat-8 OLI images were used consisting of mostly cloud free scenes June 2013 to July 2018. Scenes of images between 165/38 and 165/39 and 166/38 paths/rows. These images were used to extract reflectance data for a single pixel and were obtained from (US Geological Survey (USGS) 2019). First, radiometric correction was implemented to normalize satellite images for factors such as sensor degradation, Earth-Sun distance variation, incidence angle, view angle and time of data gathering were applied to the image data. This process convert Digital Number (DN) into radiance using calibration parameters that accompanied with the images metadata. Second, in order to obtain the surface reflectance values, the process of atmospheric correction was executed using the Fast Line-of-sight Atmospheric Analysis of Spectral Hypercubes (FLAASH) (Cooley et al., 2002). The FLAASH module can convert the raw quantized calibrated pixel values to surface reflectance (Hu and Tang, 2012).

Figure 1 shows the Electrical Conductivity (EC) sampling stations along the Karun River. Noteworthy, salinity was measured by Iran Water and Power Recourses Development Company, Using Portable conductivity meter Profile Cond 3310. Eight stations are located in the critical point for EC measuring: ALIKALE, GOTVAND, SHOOSHTAR, SHOTEYT, GARGAR, DEZ, AHVAZ, and ABADAN. Average EC of each station are listed in Table 1. Iran Water and Power Recourses Development Company obtained 102 observed EC samples from June 2013 to July 2018 along the Karun River. A total of 102 observed samples were divided into 70% training and 30% test from June 2013 to July 2018 along the Karun River. The time difference between satellite images and In situ acquisitions was at most two days and data is provided in different seasons and months.

Table 1. The average of EC at each station

EC sampling station	Average EC $\mu\text{s cm}^{-1}$
ALIKALE	470
GOTVAND	1115
SHOOSHTAR	1269
SHOTEYT	1656
GARGAR	3349
DEZ	2554
AHVAZ	2265

2.2 Band Selection

Recognition bands were important for setting the model to match with the accuracy of the predictive water salinity (Nguyen et al., 2018). Therefore it is essential to determine which combination of bands would conclude the best outcomes. In order to address this issue, the Spectral signature of salty water was employed. Table 1 shows the average EC of each station and it shows that the Karun River is salty from top to bottom.

Landsat-8 (OLI) contains 11 bands but at the present study, however, only 7 bands have been used: Band 1- Coastal/Aerosol (0.433 - 0.453 μm), Band 2-Blue (0.450 - 0.515 μm), Band 3 - Green (0.525 - 0.600 μm), Band 4 -Red (0.630 - 0.680 μm), Band 5 - Near Infrared (0.845 - 0.885 μm), Band 6 - Shortwave Infrared 1 (1.560 - 1.660 μm), Band 7 - Shortwave Infrared 2 (2.08 - 2.35 μm). Figure 2 shows the average reflectance for each station in the Karun River. Figure 2 shows that approximately with increasing salinity, the reflectance values also increase, it also shows that band 3 has the highest reflectance. As can be seen from Figure 2, Bands 1, 2 and 3 have changed with increasing salinity more regularly and meaningful Compared to other bands. Therefore, the combination of bands 1, 2 and 3 have been selected for modeling water salinity.

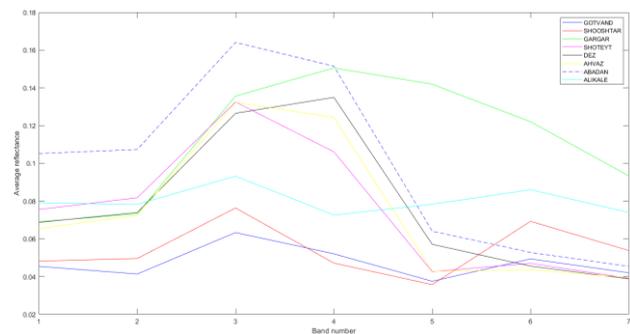


Figure 2. Spectral signature of salty water for each station

2.3 Modeling of Water Salinity

The Support Vector Machine (SVM) was originally used for classification issue, i.e. Support Vector Classification (SVC), but thereafter extended for using along with regression issue, namely Support Vector Regression (SVR) (Vapnik, 1995). The outcomes related to the quality of the SVR depend on several factors e.g. proper setting of the SVR meta-parameters: the loss function ϵ , the error penalty factor C and the kernel function parameters (Wang et al., 2016). Commonly, radial basis kernel function (RBF), $k(x, x') = k(x, x') = \exp(-||x - x' / \sigma^2)$, has been used in similar studies, likewise, it is implemented in the present study. Finally, the Genetic Algorithm (GA) is employed to optimize some parameters including C, ϵ and σ .

GA is an optimization technique developed by (Holland, 1975).(Goldberg and Holland, 1988; Michalewicz and Hartley, 1996) discussed the mechanism and robustness GA in solving nonlinear optimization problems. Three processes of reproduction, crossover and mutation are applied probabilistically to discrete decision variable that are coded into binary or real number strings. Figure 3 shows flowchart of the GA algorithm.

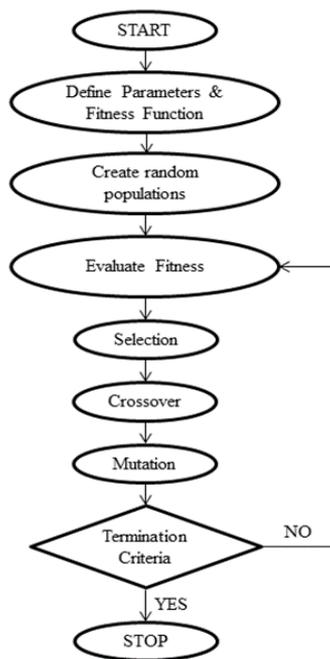


Figure 3. The flowchart of the process cycle for the GA

3. RESULTS

Spectral signature analysis proved that bands 1, 2 and 3 are the best for modeling water salinity (see figure 2). In the present study, the GA is used to determine the SVR meta-parameters including the loss function ϵ , the error penalty factor C and σ parameters, which are obtained to be 1×10^{-9} , 1099 and 0.96, respectively. Figure 4 shows how the GA converges to optimize the SVR meta-parameters. In this study, the parameters of GA is: the crossover operator rate is 0.7, the mutation operator's rate is 0.05, the population size is 20 and the maximum iterations is 70 (Termination Criteria); then by putting the target function $1-R^2$ in the genetic algorithm, attempts have been made to minimize it. The SVR model resulted in values of R^2 and RMSE for test data, which are respectively obtained to be 0.7 and $390 \mu\text{s cm}^{-1}$ (Figure 5).

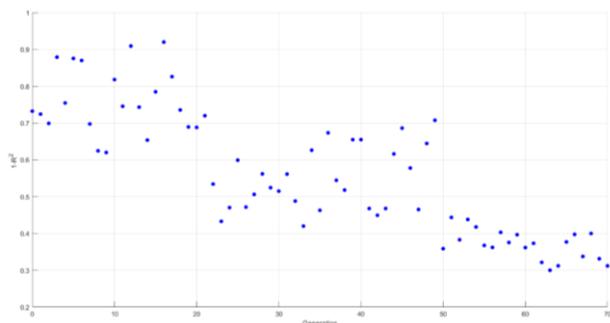


Figure 4. Converging the GA to optimize the SVR meta-parameters

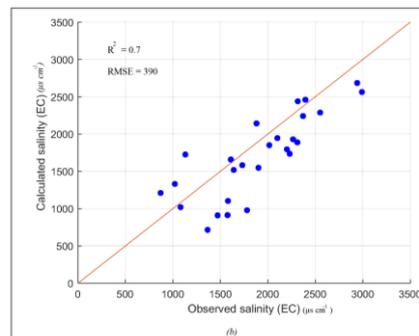
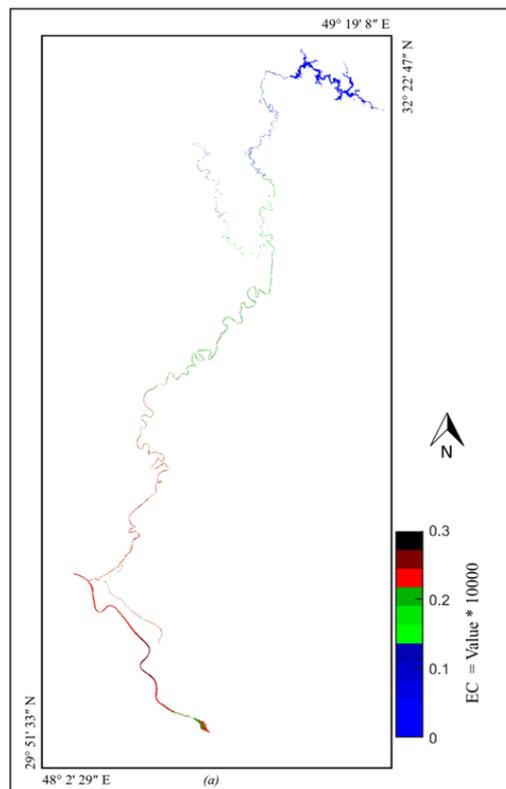


Figure 5. Relation plot of the calculated salinity vs. measured ones (test data)

Finally, SVR model was employed to achieve a water salinity map for the whole study area in two Landsat-8 OLI images: on 1 February 2015 and 5 September 2018. Figure 6 illustrates the water salinity map achieved by implementing the SVR models.



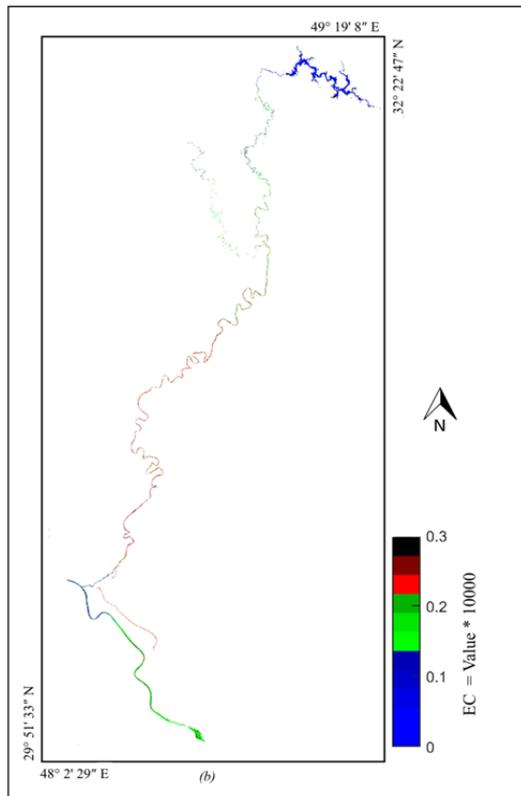


Figure 6. The water salinity map from the SVR model, (a) on 1 February 2015, (b) on 5 September 2018.

4. DISCUSSION AND CONCLUSIONS

This study investigated the relationship between reflectance wavelength and water salinity retrieved from Landsat-8 OLI in the Karun River. SVR model had a good operation by considering the large size, geographic complexity of the study domain and the wide range of field data that change between 385 and $4310\mu\text{s cm}^{-1}$.

(Jahad Agricultural Organization of Khuzestan, 2018) reported that in 2014 and 2016, 8200 and 3722 hectares of agricultural land along the Karun River were allocated to be used for sugar cane, respectively. Therefore, reducing sugar cane production can decrease water salinity from 1 February 2015 to 5 September 2018. Furthermore (Iran Water Resources Management Company 2018) reported an increased precipitation from 2015 to 2018 in the Karun river basin which was measured to be 228.3 mm and 379.8 mm, respectively. Therefore, the maps are consistent with the reports, so the models are implemented correctly and the results are acceptable.

As shown in Figure 6, although upstream of the river is saltier on 5 September 2018 than 1 February 2015, however, due to more rainfall and a reduction in the cultivation of sugar cane, downstream of the river is saltier in 1 February 2015. February is in the winter and rainfall is higher than in September, which is in the summer and it was expected that salting would be lower in February but due to the reasons, water salinity has fallen from 2015 to 2018.

Improving field data is the vital priority work for future study to probe more thoroughly the relationship between water salinity and satellite images. In addition, the current study discloses that the future study investigating the effect of seasons changing and temperature changes on the reflectance of the Landsat-8 OLI images, In addition, the contribution of thermal bands of the Landsat-8 can help to increase accuracy.

ACKNOWLEDGEMENTS

Authors would like to acknowledge the Iran Water and Power Recourses Development Company for providing the in situ water salinity data and the US Geological Survey (USGS) for providing Landsat-8 OLI images. This work was supported financially by Iran National Science Foundation (INSF).

REFERENCES

- Cooley, T., Anderson, G. P., Felde, G. W., Hoke, M. L., Ratkowski, A. J., Chetwynd, J. H., Gardner, J. A., Adler-Golden, S. M., Matthew, M. W. & Berk, A. FLAASH, a MODTRAN4-based atmospheric correction algorithm, its application and validation. IEEE International Geoscience and Remote Sensing Symposium, 2002. IEEE, 1414-1418.
- Goldberg, D. E. & Holland, J. H. 1988. Genetic algorithms and machine learning. Machine learning, 3(2), pp 95-99.
- Holland, J. H. 1975. Adaptation in natural and artificial systems Ann Arbor. The University of Michigan Press, 1(975).
- Hu, C. & Tang, P. 2012. Automatic algorithm for relative radiometric normalization of data obtained from Landsat TM and HJ-1A/B charge-coupled device sensors. Journal of Applied Remote Sensing, 6(1), pp 063509.
- Iran Water Resources Management Company 2018. precipitation [Online]. Available: <http://www.wrbs.wrm.ir/>.
- Jahad Agricultural Organization of Khuzestan. 2018. Sugar cane [Online]. Available: <https://ajkhz.ir/main/>.
- Karamouz, M., Kerachian, R., Akhbari, M. & Hafez, B. 2009. Design of river water quality monitoring networks: a case study. Environmental Modeling & Assessment, 14(6), pp 705.
- Keshavarzi, B., Mokhtarzadeh, Z., Moore, F., Mehr, M. R., Lahijanzadeh, A., Rostami, S. & Kaabi, H. 2015. Heavy metals and polycyclic aromatic hydrocarbons in surface sediments of Karoon River, Khuzestan Province, Iran. Environmental Science and Pollution Research, 22(23), pp 19077-19092.

Khadim, F. K., Su, H. & Xu, L. 2017. A spatially weighted optimization model (SWOM) for salinity mapping in Florida Bay using Landsat images and in situ observations. *Physics and Chemistry of the Earth, Parts A/B/C*, 101(86-101).

Michalewicz, Z. & Hartley, S. J. 1996. Genetic algorithms+ data structures= evolution programs. *Mathematical Intelligencer*, 18(3), pp 71.

Naddafi, K., Honari, H. & Ahmadi, M. 2007. Water quality trend analysis for the Karoon River in Iran. *Environmental monitoring and assessment*, 134(1-3), pp 305-312.

Nguyen, P. T., Koedsin, W., McNeil, D. & Van, T. P. 2018. Remote sensing techniques to predict salinity intrusion: application for a data-poor area of the coastal Mekong Delta, Vietnam. *International journal of remote sensing*, 39(20), pp 6676-6691.

US Geological Survey (USGS) 2019. Science for a changing world [Online]. Available: <https://earthexplorer.usgs.gov/>.

Vapnik, V. N. 1995. *The nature of statistical learning. Theory.*

Wang, L. a., Zhou, X., Zhu, X., Dong, Z. & Guo, W. 2016. Estimation of biomass in wheat using random forest regression algorithm and remote sensing data. *The Crop Journal*, 4(3), pp 212-219.

Xie, Z., Zhang, C. & Berry, L. 2013. Geographically weighted modelling of surface salinity in Florida Bay using Landsat TM data. *Remote sensing letters*, 4(1), pp 75-83.

Zhao, J., Temimi, M. & Ghedira, H. 2017. Remotely sensed sea surface salinity in the hyper-saline Arabian Gulf: Application to landsat 8 OLI data. *Estuarine, Coastal and Shelf Science*, 187(168-177).