

SPATIO-TEMPORAL SALINITY MONITORING OF THE GHAGHARA RIVER USING LANDSAT TIME-SERIES IMAGERY AND MULTIPLE REGRESSION ANALYSIS

M. Gašparović^{1,*}, S. K. Singh²

¹ Chair of Photogrammetry and Remote Sensing, Faculty of Geodesy, University of Zagreb, Kačićeva 26, Zagreb, Croatia –
mgasparovic@geof.unizg.hr

² K. Banerjee Centre of Atmospheric & Ocean Studies, IIDS, Nehru Science Centre, University of Allahabad, Allahabad-211002,
Uttar Pradesh, India – sudhirinjnu@gmail.com

Commission III, WG III/4, WG III/10

KEY WORDS: Spatio-temporal, Monitoring, Water salinity, Modelling, Landsat

ABSTRACT:

Nowadays, water has become one of the most important environmental issues for our ecosystem and is facing major challenges today. During the COVID-19 pandemic, the world has understood the need for good quality of water for sanitation and hygiene. Earth observing satellites plays a critical role in near-real-time detection and monitoring of land and water change and quality. This research presents a methodology for modeling and mapping water salinity in high spatial resolution. Data for modeling were measured on the five monitoring stations (Ayodhya, Basti, Birdghat, Paliakalan, and Turtipar) along the Ghagraha River Basin in India, during the period of 28 years (1985-2013). In this research, Electrical Conductivity (EC) as water salinity parameter modeled by means of Landsat 5 satellite imagery. All available Landsat 5 imagery were acquired on the same date as the ground measurement data was utilized for the modeling. Modeling was done based on linear, 2nd and 3rd polynomial multiple regression analysis. All statistical parameters for accuracy assessment show that 3rd degree polynomial performs better EC prediction capability than 2nd degree polynomial and linear regression. The 3rd degree polynomial multiple regression model RMSE, R², MAE, p-value were 8.682, 0.993, 6.493, 0.008, respectively. The developed algorithm provides new knowledge that can be widely applied in various environmental research mapping and monitoring like water salinity. Also, this method allows rapid detection of water pollution, which has an important impact on human health, agriculture, and the environment.

1. INTRODUCTION

Water is a multi-facet, and one of the most vital elements for survival on earth has become emerging environmental issues for our ecosystem and is facing major challenges today. During the COVID-19 pandemic, the world has understood the demand for good quality of water for sanitation and hygiene. Earth observation satellites play a critical role in near-real-time detection and monitoring of vegetation, land, water change, and quality. The freshwater systems are continuously facing the threat of anthropogenic contaminants due to the direct untreated discharge of wastewaters. The water quality of any river system is affected by a wide range of natural influences such as vegetation cover, climate, topography, soils, and geologic structure of a basin (Bartram and Balance, 1996) and anthropogenic activities like land use/land cover (LULC) change, industrial wastes, domestic wastes, agricultural wastes, and atmospheric pollutants (Amin et al. 2014; Singh et al. 2014; Gašparović et al. 2018; Pilaš et al. 2019). Therefore, ecosystem services knowledge has become the most important issue in environmental policymaking and management. The concept of ecosystem services has been defined by Daily (1997) as ecological functions that sustain life and have been categorized into four service types (provisioning, regulating, supporting, and cultural). In the wake of drastically changing, LULC and climate have intensified the degradation of the aquatic environment. Hence, it is likely for ecosystems to be modified to the extent that they can no longer render services to support life in the near future. Our understanding of the functioning of the ecosystem will be challenged and will require a better

assessment of the supply-demand chain to reduce the potential negative tradeoffs and conserve the valuable ecosystem. There is an urgent need of present to address the problems of rising sediment loads and nutrients delivered to many major river basins/water bodies. Modeling approaches are needed because they are capable of accurate assessment of the catchment scale, particularly in the concentration of water quality where the in-situ monitoring is not feasible due to either location or availability of limited funds.

Investigation of spatio-temporal patterns of water quality parameter is crucial to managing water resources (Breitburg et al. 2018). The difficulty in monitoring the water quality parameters in a remote or non-accessible area in a cost-effective manner can be achieved through the applications of multispectral and hyperspectral satellite data. In the last few decades, the earth observing datasets have been more frequently used to collect water quality information, particularly of lakes, ponds, and reservoirs (Goetz et al. 2008). The retrieval of water quality parameters rely on satellite bandwidth. The water-leaving radiance lies in the visible band and highly absorptive in near-infrared and infrared regions of the spectrum (Govender et al. 2007). The multispectral sensors have the capability to retrieve water quality parameters as compared to coarse bands/panchromatic band (Hestir et al. 2015; Topp et al. 2020). Linear Imaging Self Scanning System satellite data was used to model electrical conductivity (EC) in the Tawa Reservoir during the monsoon and after the monsoon. A simple linear regression model was developed to model EC using four bands, and results

* Corresponding author

explain that EC was inversely correlated with shorter wavelength bands. Whereas multiple linear regression analysis shows slightly better relations as compared to simple linear regression (Choubey et al. 1994). Hyperspectral sensors have a higher number of spectral bands and widely applied to ascertain the water quality of aquatic ecosystems (Govender et al. 2007). Abdelmalik (2018) used ASTER data of Qaroun Lake, Egypt, to retrieve the water quality parameters. The result show that the quadratic regression model has performed best in the case of electrical conductivity ($R^2=0.996$) and salinity retrieval ($R^2 = 0.985$). Visible and near-infrared bands are highly sensitive to water salinity (Pegau et al. 1997). Landsat 7 ETM+ data was used for modeling the water quality parameters (dissolved oxygen, turbidity, total hardness, alkalinity, chemical and biological oxygen demand) through multiple regression analysis of the Ganga river and results suggest that these parameters have a good correlation with the spectral radiance of bands. The turbidity shows poor correlation with spectral bands due to low suspended matter in water (Sharma et al. 2019). The stepwise multiple regression analysis among long term dissolved oxygen data sets and satellite-derived environmental variables shows a significant-good correlation and suggest that the prediction of the dissolved oxygen with the aid satellite data is accurate (Kim et al. 2020).

This research presents a methodology for modeling and mapping water salinity (EC) in high spatial resolution based on Landsat 5 imagery.

2. MATERIALS AND METHODS

2.1 Study site and data

The proposed study is investigating the spatio-temporal water salinity (EC variations). The Ghaghara River is a perennial and trans-boundary river originates near Lake Mansarovar (28.5983°N , 83.9311°E). It has a catchment area 127950 km^2 ; the majority of its catchment lies in Nepal (55%) and remaining in India (45%). It met with its tributaries Sarju at Brahmaghat in India and called as Ghaghara River. It joins the Ganges at Dorigang. The other important tributaries of the Ghaghara River are the Sarju, Rapti, and Little Gandak. The river has almost the heterogeneous topography from source to mouth and having the longest distance river in Nepal (~507 km). In the alluvial plains, Ghaghara shows the meandering pattern and oxbow lake and lateral soil erosion. The dominant land use/land cover pattern is the cropland followed by the mixed forest and grassland (Singh et al. 2017). The dominant soil is the older alluvium (Pleistocene & yellow to brown color) and the newer alluvium (Holocene & gray to black color). The average annual rainfall in the catchment ranges from 900 to 1400 mm, and evapotranspiration in the basin ranges from 1700 to 1950 mm. In India, a total of twenty-two and in Nepal five district administrative units falling in the river catchment (Mohan 2018). In India, the population density is high in the river catchment, and water is used for irrigation. It brings a huge flood during the southwest monsoon; the river also has spiritual significance.

For this research, EC/water salinity data was monitored on five stations (Ayodhya, Basti, Birdghat, Paliakalan, and Turtipar) along the Ghaghara River during the period of 28 years (1985-2013) (Figure 1). In this research, Electrical Conductivity (EC) as a water salinity parameter was tested and correlated with Landsat 5 satellite imagery (Figure 2). In India, the water quality of rivers is monitored since the late 1950s, and the

monitoring network is spread all over India, and the water quality network of Central Water Commission is incidental to a hydrological observation network. The Central Water Commission, New Delhi, India, is the agency for monitoring of water quality and discharge. The Paliakalan sampling station is located at 28.3928°N latitude and 80.5306°E longitude (the reference code GGU6016). The site is classified as trend station and Gauge, Discharge, Silt, and Water Quality (GDSQ) type of station. This station is also used for Hydrological Observation/flood forecasting (HO/FF). The trend station shows how the monitoring point varies over the period of time due to both anthropogenic and geogenic activities (GWQM, 2017). The trend station quality is monitored once in a month before southwest monsoon (pre-monsoon), and a total of twenty-five parameters were determined in general. The twenty-five parameters are categorized into General, Nutrients, Demand Parameters, Major Ions, Other Inorganic, and Microbiological parameters. The long-term monthly water quality data (1984-2018) was collected from the Central Water Commission. The EC is measured through Conductivity Meter and Water Analysis Kit. The monitoring agency follows the Bureau of Indian Standard (BIS) Methods for Testing Water and Wastewater-methods of sampling and testing (physical and chemical) (IS:3025).

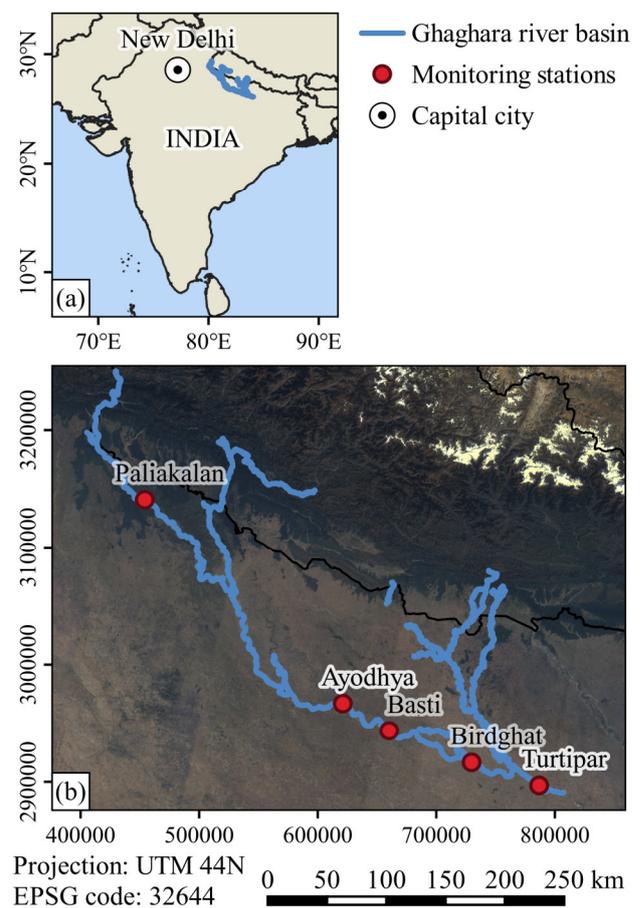


Figure 1. (a) Location of the study area, (b) study area and monitoring stations on Ghaghara River basin (Landsat 5 'true color' composite (B3–B2–B1) for 2010)

In order to allow a large temporal overlap of satellite data with ground truth data, this research is based on clear-sky Landsat 5 data (Figure 2). The atmospherically pre-corrected surface reflectance data were used from the Landsat 5 sensor (USGS

Landsat 5 Surface Reflectance Tier 1). Satellite data contains seven bands: four visible and near-infrared (VNIR) bands, two short-wave infrared (SWIR) of 30-m spatial resolution, and one thermal infrared TIR band of 120-m spatial resolution (Table 1). Values for all Landsat 5 bands were extracted via Google Earth Engine.

Band	Wavelength (µm)	Description
B1	0.45 – 0.52	Blue
B2	0.52 – 0.60	Green
B3	0.63 – 0.69	Red
B4	0.77 – 0.90	Near Infrared
B5	1.55 – 1.75	Shortwave infrared 1
B6	10.40 – 12.50	Thermal infrared
B7	2.09 – 2.35	Shortwave infrared 2

Table 1. Spectral bands of the Landsat 5 imagery (Gašparović et al. 2017)

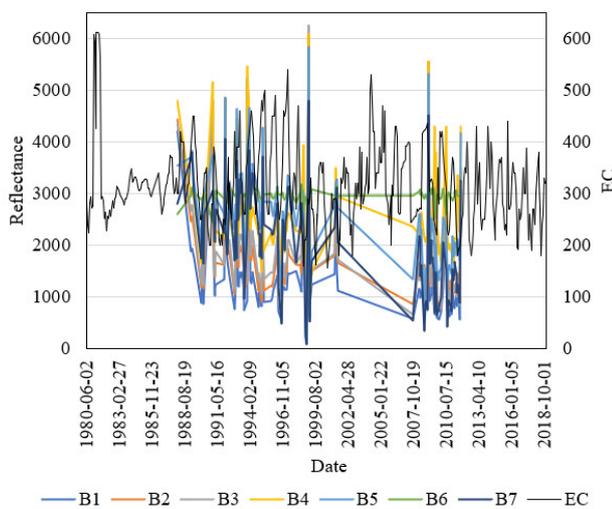


Figure 2. Spectral reflectance of Landsat 5 bands and ground measured EC for Paliakalan station, India

For further research, all available imagery were acquired on the same date as the ground measurement data was used for the research. That defined data was prepared and selected for the stations (Table 2). The satellite data was used for regression analysis. The number of ground-truth measurements, satellite measurements, and final measurements used in modeling were shown in Table 2. Final measurements represent combined ground-truth and satellite measurements acquired on the same day.

Station	Ground-truth	Landsat 5	Final
Ayodhya	238	175	1
Basti	421	547	19
Birdghat	451	201	8
Paliakalan	452	130	8
Turtipar	407	209	8
Total	1969	1262	44

Table 2. Number of measurements used in this research

2.2 Methods

Based on the Landsat 5 (B1-B7) time-series and multiple regression analysis with ground-truth data from stations were used for the modeling of the water salinity. Three types of

multiple regression analyses were tested: linear, 2nd degree polynomial, and 3rd degree polynomial. The developed models enable water salinity mapping and monitoring during the entire Ghagraha River flow based on the clear-sky Landsat 5 satellite imagery collected on the same date as the measurement was done on the ground station. For modeling, all bands and various spectral indices were used (Gholizadeh et al. 2016; Abdelmalik 2018, Gašparović et al. 2019). Additionally, in this research, three types of Normalized difference water index (NDWI₁, NDWI₂, NDWI₃), Enhanced vegetation index (EVI), Normalized difference vegetation index (NDVI) and one ratio (ratio₅₄) were applied. Calculation of indices mentioned above was done based on the following equations (1-6):

$$NDWI_1 = \frac{B4 - B5}{B4 + B5} \quad (1)$$

$$NDWI_2 = \frac{B4 - B7}{B4 + B7} \quad (2)$$

$$NDWI_3 = \frac{B2 - B4}{B2 + B4} \quad (3)$$

$$EVI = 2.5 \cdot \frac{B4 - B3}{B4 + 6 \cdot B3 - 7.5 \cdot B1 + 1} \quad (4)$$

$$NDVI = \frac{B4 - B3}{B3 + B3} \quad (5)$$

$$ratio_{54} = \frac{B5}{B4} \quad (6)$$

where B1, B2, B3, B4, B5, B7 are Landsat 5 bands and specifications are defined in Table 1.

Accuracy assessment of the regression analysis was performed based on various standard statistical parameters: Residual standard error (RSE), F-statistic, and p-value. Furthermore, an independent accuracy assessment based on the Leave-one-out cross-validation (LOOCV) approach was applied. Based on the LOOCV approach, Root Mean Square Error (RMSE), Coefficient of determination (R²), and Mean Absolute Error (MAE) was also calculated.

3. RESULTS

For the purpose of observing the dependence of the modeling variables, the correlation matrix of all used variables was employed to know the correlation among the studied variables (Figure 3).

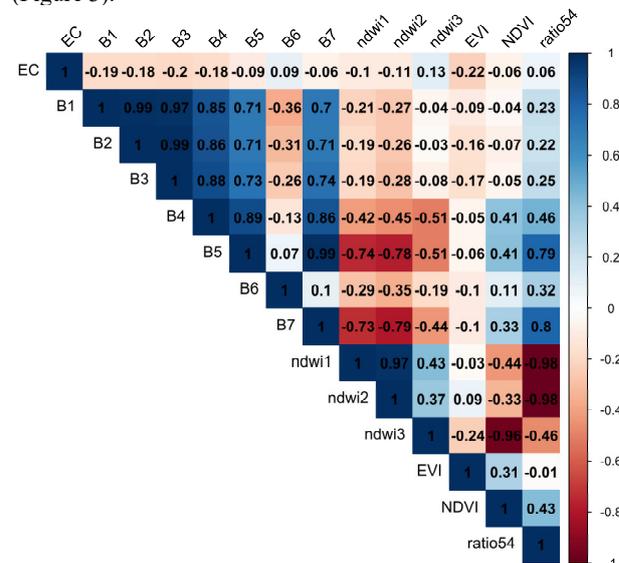


Figure 3. Correlation matrix of all variables

Accuracy assessment of three regression analysis was carried out based on the linear, 2nd degree polynomial, and 3rd degree polynomial, results are shown in Table 3. All three regressions are provided on all 44 final measurements and 14 variables. As variables, seven Landsat bands (B1-B7), five spectral indices (NDWI₁, NDWI₂, NDWI₃, NDVI, EVI), and one spectral ratio (ratio₅₄) were used. The developed regression models enable water salinity mapping in high resolution (30 m) and based on the remote sensing satellite data.

Statistical parameter	Linear	Poly 2 nd	Poly 3 rd
RSE	97.980	80.190	27.320
F-statistic	1.408	2.120	15.830
p-value	0.213	0.056	0.008
RMSE	84.258	52.540	8.682
R ²	0.328	0.738	0.993
MAE	64.504	39.062	6.493

Table 3. Statistical parameters for accuracy assessment of three multiple regression analysis

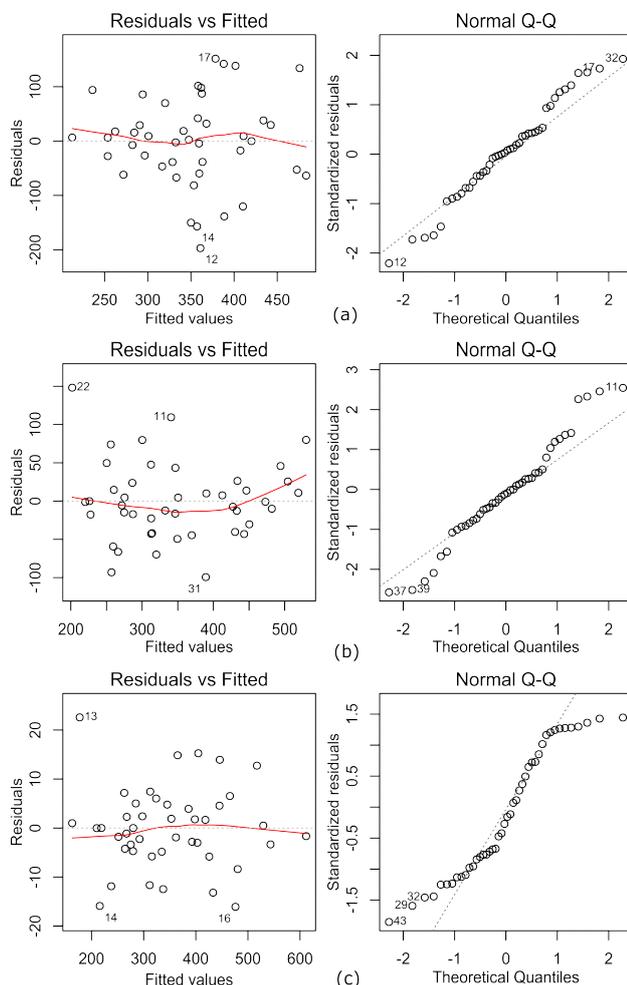


Figure 4. Statistical visualization of (a) linear, (b) 2nd degree polynomial and (c) 3rd degree polynomial type of three multiple regression analysis

All statistical parameters (Table 3), as well as statistical visualization (Figure 4), show that linear regression performed worse prediction capability compared to polynomial regressions. The results show that 3rd degree polynomial has the best prediction capability as the p-value (<0.05). All statistical results also confirmed that 3rd degree polynomial has better EC prediction capability than 2nd degree polynomial. The residual

versus fitted plot shows whether the plot is normal (no change), heteroscedasticity (constant variance), and nonlinearity (Figure 4). The first plot indicates uncorrelated values between residuals and the fitted means homoscedastic linear model (Figure 4a). The Q-Q plot shows the presence of outliers and inequality (Helsel and Hirsch 1992). The 3rd order polynomial shows high prediction capability as R² (0.993) and RMSE (8.682) as compared to multiple linear regression and 2nd polynomial (Table 3). The normal Q-Q plot does not support normality, and the distribution of multiple linear regression shows left-skewed, whereas the 2nd polynomial shows a similar trend. The 3rd polynomial shows the right-skewed data.

Similar to our previous study (Gudelj et al. 2018; Gašparović et al. 2019), water body mapping was done based on the NDWI₃ (Figure 5a). Modeled EC calculated based 3rd degree polynomial multiple regression coefficients was calculated for a water body and showed in Figure 5b.

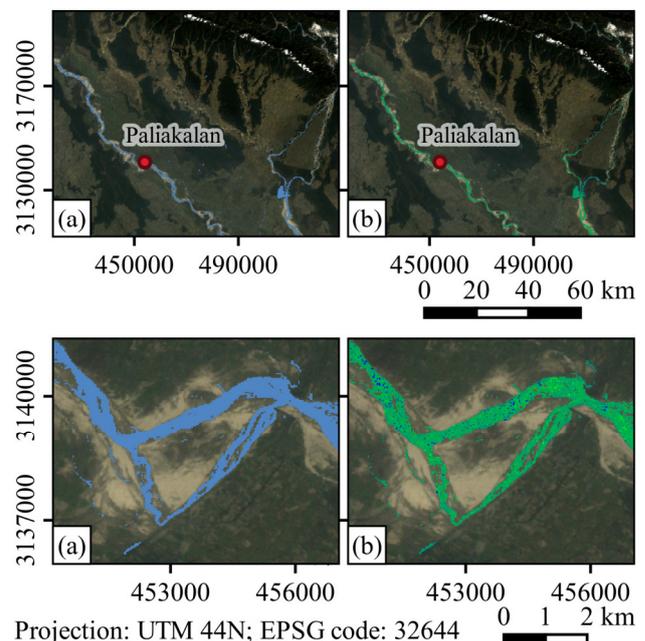


Figure 5. (a) River body mapped based NDWI₃, (b) modeled EC based 3rd degree polynomial multiple regression analysis for Paliakalan station surroundings, India

4. CONCLUSIONS

Regarding that water has become one of the most important environmental issues for our ecosystem and is facing major challenges today, this research presents the methodology for modeling EC based on Landsat 5 imagery. This research allows water salinity mapping in high spatial resolution. Accuracy assessment calculated based on the five stations monitoring stations (Ayodhya, Basti, Birdghat, Paliakalan, and Turtipar) along the Ghagraha River during the period of 28 years (1985-2013). All statistical result show that 3rd degree polynomial performs better EC prediction capability than 2nd degree polynomial and linear regression. The 3rd degree polynomial multiple regression model RMSE, R², MAE, p-value were 8.682, 0.993, 6.493, 0.008, respectively. The developed algorithm provides new knowledge that can be widely applied in various environmental research mapping and monitoring like water salinity. Also, this method allows the rapid detection of water pollution, which has an important impact on human

health, agriculture, and the environment. The developed algorithm can be applied to many other areas around the Earth and based on the other hyperspectral and optical-based satellite data, e.g., Sentinel-2, RapidEye, and PlanetScope.

ACKNOWLEDGEMENTS

The authors would like to thank the University of Zagreb, Croatia that funded the project entitled: “Advanced photogrammetry and remote sensing methods for environmental change monitoring” (RS4ENVIRO), the Ministry of Science and Education, Croatia for supporting for space and space technology initiatives, and the Science & Engineering Research Board, New Delhi, India (sanction order no CRG/2019/003551) under which this research was conducted. The authors also thank Central Water Commission, New Delhi, India, for providing the long-term water quality data.

REFERENCES

- Abdelmalik, K.W. 2018. Role of statistical remote sensing for Inland water quality parameters prediction. *The Egyptian Journal of Remote Sensing and Space Science*, 21(2), 193-200.
- Amin, A., Fazal, S., Mujtaba, A., Singh, S.K. 2014. Effects of land transformation on water quality of Dal Lake, Srinagar, India. *Journal of the Indian Society of Remote Sensing*, 42(1), 119-128.
- Bartram J., Balance R. 1996. *Water quality monitoring—a practical guide to the design and implementation of freshwater quality studies and monitoring programmes*. United Nations Environment Programme and the World Health Organization Geneva.
- Breitburg, D., Levin, L.A., Oschlies, A., Grégoire, M., Chavez, F.P., Conley, D.J., Garçon, V., Gilbert, D., Gutiérrez, D., Isensee, K., Jacinto, G.S., Limburg, K.E., Montes, I., Naqvi, S.W.A., Pitcher, G.C., Rabalais, N.N., Roman, M.R., Rose, K.A., Seibel, B.A., Telszewski, M., Yasuhara, M., Zhang, J. 2018. Declining oxygen in the global ocean and coastal waters. *Science*, 359(6371), eaam7240.
- Choubey, V.K. 1994. Monitoring surface water conductivity with Indian remote sensing satellite data: A case study from central India. *IAHS Publications-Series of Proceedings and Reports-Intern Assoc Hydrological Sciences*, 219, 317-326.
- Daily, G.C. 1997. *Nature's services* (Vol. 3). Island Press, Washington, DC.
- Gašparović, M., Medak, D., Pilaš, I., Jurjević, L., Balenović, I. 2018. Fusion of Sentinel-2 and PlanetScope Imagery for Vegetation Detection and Monitoring. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, XLII-1, 155-160.
- Gašparović, M., Zrinjski, M., Gudelj, M. 2017. Analysis of Urbanization of Split. *Geodetski list*, 71(3), 189-202.
- Gašparović, M., Zrinjski, M., Gudelj, M. 2019. Automatic cost-effective method for land cover classification (ALCC). *Computers, Environment and Urban Systems*, 76, 1-10.
- Gholizadeh, M.H., Melesse, A.M., Reddi, L. 2016. A comprehensive review on water quality parameters estimation using remote sensing techniques. *Sensors*, 16(8), 1298.
- Goetz, S.J., Gardiner, N., Viers, J.H. 2008. Monitoring freshwater, estuarine and near-shore benthic ecosystems with multi-sensor remote sensing: An introduction to the special issue. *Remote Sensing of Environment*, 112(11), 3993-3995.
- Govender Brando, V.E., Bresciani, M., Giardino, C., Matta, E., Villa, P., Dekker, A.G. 2015. Measuring freshwater aquatic ecosystems: The need for a hyperspectral global mapping satellite mission. *Remote Sensing of Environment*, 167, 181-195.
- Govender, M., Chetty, K., Bulcock, H. 2007. A review of hyperspectral remote sensing and its application in vegetation and water resource studies. *Water Sa*, 33(2), 145-152.
- Gudelj, M., Gašparović, M., Zrinjski, M. 2018. Accuracy Analysis of the Inland Waters Detection. *18th International Multidisciplinary Scientific GeoConference: SGEM*, 18(1.5), 203-210.
- GWQM (2017). *Guidelines Water Quality Monitoring*. The Central Water Commission, New Delhi, India. https://cpcb.nic.in/wqm/Guidelines_Water_Quality_Monitoring_2017.pdf (7 April 2020)
- Helsel, D.R., Hirsch, R. M. (1992). *Statistical methods in water resources* (Vol. 49). Elsevier Science, Amsterdam, Netherlands.
- Kim, Y.H., Son, S., Kim, H.C., Kim, B., Park, Y.G., Nam, J., Ryu, J. 2020. Application of satellite remote sensing in monitoring dissolved oxygen variabilities: A case study for coastal waters in Korea. *Environment International*, 134, 105301.
- Mohan, R. 2018. Ghaghara River System—Its Current Status and Value to Society. doi.org/10.1007/978-981-10-2984-4_12.
- Pegau, W.S., Gray, D., Zaneveld, J.R.V. 1997. Absorption and attenuation of visible and near-infrared light in water: dependence on temperature and salinity. *Applied Optics*, 36(24), 6035-6046.
- Pilaš, I., Gašparović, M., Đodan, M., Balenović, I., Dugački, I. 2019. Possibilities of the Application of the Medium-Resolution Landsat 8 and the High-Resolution RapidEye Optical Imagery in Visualization and Detection of Changes in the Forest Cover by Windthrows. *Geodetski list*, 73(3), 261-276.
- Sharma, B., Kumar, M., Denis, D.M., Singh, S.K. 2019. Appraisal of river water quality using open-access earth observation data set: a study of river Ganga at Allahabad (India). *Sustainable Water Resources Management*, 5(2), 755-765.
- Singh, H., Singh, D., Singh, S.K., Shukla, D.N. 2017. Assessment of river water quality and ecological diversity through multivariate statistical techniques, and earth observation dataset of rivers Ghaghara and Gandak, India. *International Journal of River Basin Management*, 15(3), 347-360.
- Singh, S.K., Srivastava, P.K., Gupta, M., Thakur, J.K., Mukherjee, S. 2014. Appraisal of land use/land cover of mangrove forest ecosystem using support vector machine. *Environmental Earth Sciences*, 71(5), 2245-2255.
- Topp, S.N., Pavelsky, T.M., Jensen, D., Simard, M., Ross, M.R. 2020. Research Trends in the Use of Remote Sensing for Inland Water Quality Science: Moving Towards Multidisciplinary Applications. *Water*, 12(1), 169.