International Journal of Research and Review
[Indexed, Peer-Reviewed Journal]
Volume 12; Issue: 9; September 2025
Website: www.ijrrjournal.com
Research Paper
E-ISSN: 2349-9788; P-ISSN: 2454-2237

Landsat-Derived NMSI Index for Estimating Total Suspended Solids Using Linear Regression Models

Danh-tuyen Vu, Thi-thu-huong Pham

Faculty of Surveying, Mapping and Geographic Information, Hanoi University of Natural Resources and Environment, Hanoi, Vietnam.

Corresponding author: Danh-tuyen Vu

DOI: https://doi.org/10.52403/ijrr.20250946

ABSTRACT

Remote sensing provides an efficient and cost-effective approach for monitoring water quality at the watershed scale. This study assessed the potential of Landsat-derived indices for estimating total suspended solids (TSS) in the Day River Basin, Vietnam. Field-based TSS data collected during the dry (April 2025) and rainy (June 2025) seasons were integrated with the Normalized Multi-band Suspended Index (NMSI) extracted from Landsat imagery. Linear regression models were developed to link in-situ TSS measurements with NMSI values. It was found that the results demonstrated strong predictive performance. The dry-season model achieved $R^2 = 0.801$, RMSE = 2.277 mg/L, and a highly significant regression slope ($p < 0.001$). The rainy-season model performed even better, with $R^2 = 0.8447$, adjusted $R^2 = 0.8377$, RMSE = 1.889 mg/L, and MAE = 1.121 mg/L. Spatial maps of predicted TSS revealed clear seasonal contrasts and downstream amplification of suspended sediment loads, with higher concentrations observed in the middle and lower reaches during the rainy season. These findings confirm that NMSI is a reliable spectral proxy for suspended sediment, and that Landsat imagery combined with regression analysis can provide accurate, spatially explicit estimates of TSS. This approach supports the development of basin-scale monitoring frameworks and contributes to sustainable sediment and water-quality management in the Day River Basin.

Keywords: Total suspended solids, Linear Regression, remotely sensed images, Day River Basin, Vietnam.

INTRODUCTION

Suspended sediments are one of the most critical parameters of water quality, influencing aquatic ecosystems, river morphology, and the transport of nutrients and pollutants. High concentrations of total suspended solids (TSS) can reduce water transparency, alter light penetration, and negatively impact aquatic habitats, fisheries, and water supply systems (1). Traditional field-based TSS monitoring methods, while accurate, are often time-consuming, costly, and spatially limited, especially in large and dynamic river basins (2).

In recent decades, remote sensing has emerged as an efficient tool for monitoring surface water quality. Optical satellite imagery provides synoptic and repetitive coverage, enabling the retrieval of sediment-related parameters at basin to regional scales. Various spectral indices, such as the Normalized Difference Turbidity Index (NDTI), Suspended Particulate Matter Index (SPMI), and more recently the Normalized Multi-band Suspended Index (NMSI), have been employed to enhance the sensitivity of satellite data to suspended sediment concentrations (1,3).

Several studies have demonstrated the effectiveness of linear regression in estimating and predicting TSS dynamics. For instance, empirical models linking field-measured TSS with satellite-derived reflectance from Landsat or Sentinel-2 have consistently achieved high coefficients of determination (R^2 values often exceeding 0.8), indicating robust predictive power in diverse hydrological contexts (2,4). Studies in monsoon-affected basins in Southeast Asia and South Asia have revealed that linear regression against time can effectively capture the declining trend of TSS following peak runoff events, thereby supporting water resource managers in anticipating water quality conditions during dry and wet seasons (3,5). Landsat satellites, with their long temporal record (since the 1980s) and moderate spatial resolution (30 m), have been widely used for monitoring sediment and water quality in rivers, lakes, estuaries, and coastal zones worldwide (5).

Numerous international studies have highlighted the applicability of remote sensing techniques for estimating and monitoring suspended sediments in aquatic environments such as a remote sensing–based early warning system was developed for monitoring TSS concentrations in Lake Mead (USA). Their findings demonstrated the capability of satellite-derived indices to capture temporal dynamics of TSS and to support proactive water quality management (6). Remote sensing was applied to assess TSS in West Kalimantan (Indonesia) as part of a preliminary study for siting a nuclear power plant. This work emphasized the role of TSS as an important aquatic parameter influencing both environmental management and resource utilization (7). A study was carried out to investigate the use of vegetation indices such as NDVI for estimating TSS in surface waters, with a focus on scale issues. Their results showed that the relationship between spectral indices and TSS can vary depending on spatial resolution and the optical properties of different water bodies (8). A study by Adjovu provided a comprehensive review of methods for measuring TDS and TSS, comparing traditional laboratory approaches with remote sensing techniques. The review concluded that satellite-based monitoring has become increasingly popular due to its cost-effectiveness and spatial coverage, though it still requires calibration with field data (9). Deep learning models were integrated with satellite reflectance data to map TSS and dissolved organic carbon (DOC) in complex coastal waters. Their study confirmed that machine learning approaches can significantly improve prediction accuracy compared with traditional regression models (10). A recent study developed robust algorithms for estimating TSS in nearshore coastal waters, showing the stability and transferability of remote sensing–based approaches when combined with in-situ measurements. Together, these studies demonstrate the global momentum in applying remote sensing for TSS monitoring, ranging from inland rivers and lakes to complex estuarine and coastal environments. They also highlight a trend from simple regression models toward more advanced approaches such as deep learning. In Vietnam, studies have primarily focused on the Red River Delta, Mekong Delta, and selected reservoirs, but relatively few efforts have been made to apply spectral indices to monitor suspended sediment in the Day River Basin, a major tributary of the Red River system. This basin is of high socio-economic and ecological importance, supporting agriculture, aquaculture, and domestic water supply, while also being increasingly affected by upstream development, sand mining, and climate-driven hydrological variability.

This study aims to evaluate the applicability of Landsat-derived NMSI for estimating TSS concentrations in the Day River Basin. By integrating field-based TSS measurements from both the dry and rainy seasons with spectral information, linear regression models are developed and validated. The objectives of this study are to: (i) analyze the spatial distribution of NMSI during contrasting hydrological conditions; (ii) establish regression relationships between NMSI and in-situ TSS; and (iii) generate spatially explicit maps of predicted TSS concentration. The results contribute to advancing remote sensing applications for water quality monitoring in Vietnam and provide practical insights for sediment and watershed management.

MATERIALS & METHODS

Materials

The Day River Basin is located in northern Vietnam, extending across Phu Tho, Ha Noi, Ha Nam, Nam Dinh, and Ninh Binh provinces. It is one of the major tributary systems of the Red River, with a catchment area of approximately 7,600 km². The river basin has a tropical monsoon climate, characterized by a distinct dry season (November-April) and rainy season (May-October). Average annual rainfall exceeds 1,600 mm, with nearly 80% falling in the rainy season. These hydrological conditions strongly influence sediment transport and suspended matter concentrations. Human activities such as sand mining, agricultural runoff, and urban wastewater discharge further contribute to spatial variability in sediment loads.

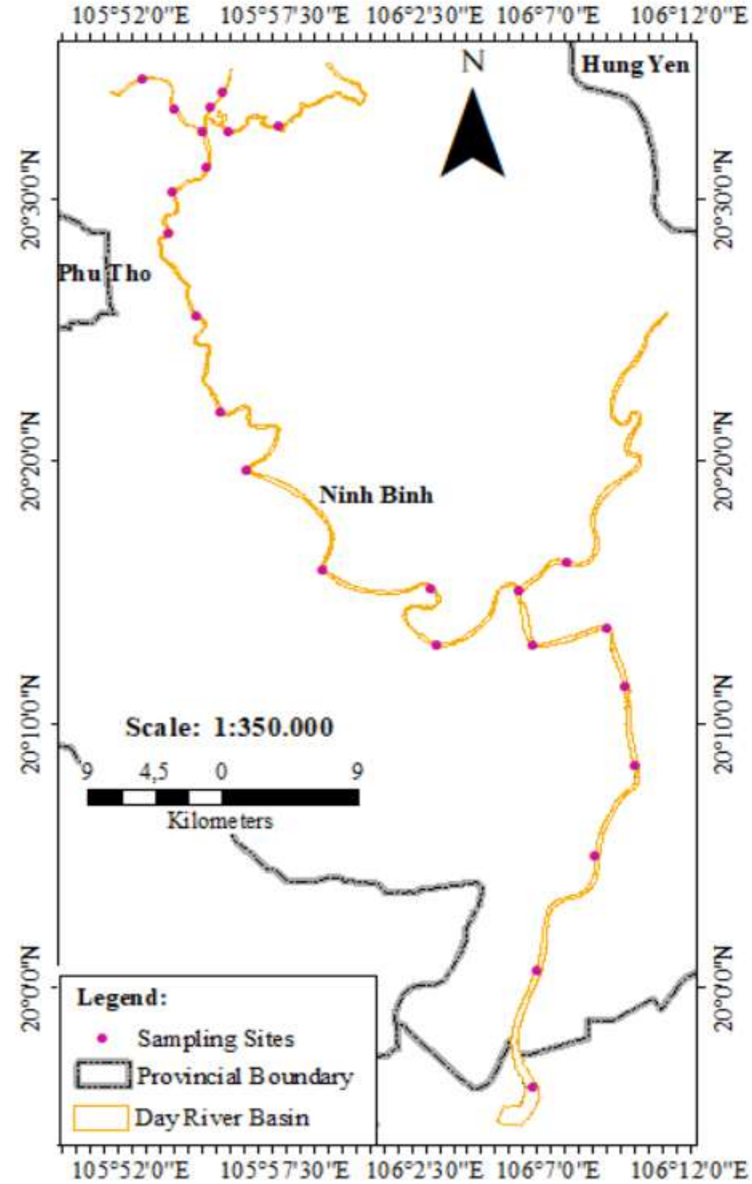

Figure 1. Study area of Day River Basin.

In this study, Landsat 8 OLI/TIRS and Landsat 9 OLI-2/TIRS-2 images covering the Day River Basin were acquired from the USGS Earth Explorer portal. Cloud-free scenes were selected for 16 April 2025 (dry season) and 11 June 2025 (rainy season). Standard Level-2 surface reflectance products were used, ensuring atmospheric correction through the Landsat Surface Reflectance Code (LaSRC). The Normalized Multi-band Suspended Index (NMSI) was calculated from visible and near-infrared bands following established formulas (e.g., Thuy et al., 2022). In addition, in-situ TSS measurements were collected on 30 March 2025 (dry season) and 18 June 2025 (rainy season) at 25 sampling stations distributed along the main stem of the Day River and selected tributaries. Water samples were analyzed in the laboratory using the gravimetric method (APHA Standard Methods, 2017), providing TSS concentrations in mg/L. These observations were temporally aligned with the Landsat overpasses to minimize mismatches between remote sensing and field measurements.

METHODS

Landsat data pre-processing and calculation of NMSI index

Level-2 surface reflectance products from Landsat 8 OLI/TIRS and Landsat 9 OLI-2/TIRS-2 were obtained from the USGS Earth Explorer. Pre-processing steps included atmospheric correction: implemented using the Landsat Surface Reflectance Code (LaSRC), cloud and shadow masking: performed with the Quality Assessment (QA) band. Only pixels flagged as cloud-free were used. Geometric correction: images were referenced to WGS84/UTM Zone 48N to ensure spatial consistency with in-situ sampling points. Image subsetting: restricted to the Day River Basin boundary using GIS tools. The Normalized Multi-band Suspended Index (NMSI) was calculated from Landsat reflectance bands according to the formula:

$$NMSI = \frac{(B_{Red} + B_{NIR}) - (B_{Green} + B_{Blue})}{(B_{Red} + B_{NIR}) - (B_{Green} + B_{Blue})} \quad (1)$$

where B_{Red}, B_{NIR}, B_{Green}, B_{Blue} denote the surface reflectance of Landsat bands 2, 3, 4, and 5, respectively. The index enhances the sensitivity of spectral data to suspended particulate matter by emphasizing reflectance differences between visible and near-infrared wavelengths.

Linear regression:

To predict TSS concentration, we applied a linear regression model, which is one of the most widely used statistical approaches for examining the relationship between a dependent variable and one or more independent variables. To establish the statistical relationship between observed TSS and NMSI, simple linear regression models were constructed separately for the dry and rainy seasons:

$$TSS_i = \alpha + \beta \times NMSI_i + \varepsilon_i \quad (2)$$

where TSS_i is the observed suspended solids concentration (mg/L) at site iii, $NMSI_i$ is the corresponding Landsat-derived index value, α is the intercept, β\betaβ is the regression slope, and ε_i is the residual error term. Regression parameters were estimated using the ordinary least squares (OLS) method. Both R^2 and adjusted R^2 were calculated to assess explanatory power. Statistical significance of the slope coefficient (β) was tested using the student's t-test ($p < 0.05$ threshold). Model performance was evaluated using multiple indicators including coefficient of determination (R^2, adjusted R^2): measuring variance explained and root Mean Square Error (RMSE):

$$NMSI = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(TSS_i - \widehat{TSS_i}\right)^2} \quad (3)$$

Mean Absolute Error (MAE):

$$MAE = \frac{1}{n}\sum_{i=1}^{n}\left|TSS_i - \widehat{TSS_i}\right| \quad (4)$$

The model's predictive performance was assessed using standard statistical indicators, including the Coefficient of Determination (R^2), which explains the proportion of variance in TSS accounted for by the regression, and the Root Mean Square Error (RMSE), which quantifies the average prediction error. Visual comparison between observed and forecasted TSS values was also conducted using line plots and histograms to assess model fit and the distribution of residuals. The rationale for applying linear regression lies in its ability to identify and quantify temporal trends in TSS dynamics.

All mapping and spatial analyses were conducted using ENVI 6.2 and ArcGIS 10.2, ensuring integration of remote sensing and field data in a consistent geospatial framework.

RESULTS & DISCUSSION

Analysis of spatial distribution of NMSI

The spatial distribution of the Normalized Multi-band Suspended Index (NMSI) along the Day River Basin exhibited clear seasonal variability between April (dry season) and June (rainy season) of 2025 as shown in Figure 2. In both periods, NMSI values showed heterogeneous patterns, reflecting differences in suspended sediment concentrations influenced by hydrological conditions, land use, and upstream inputs.

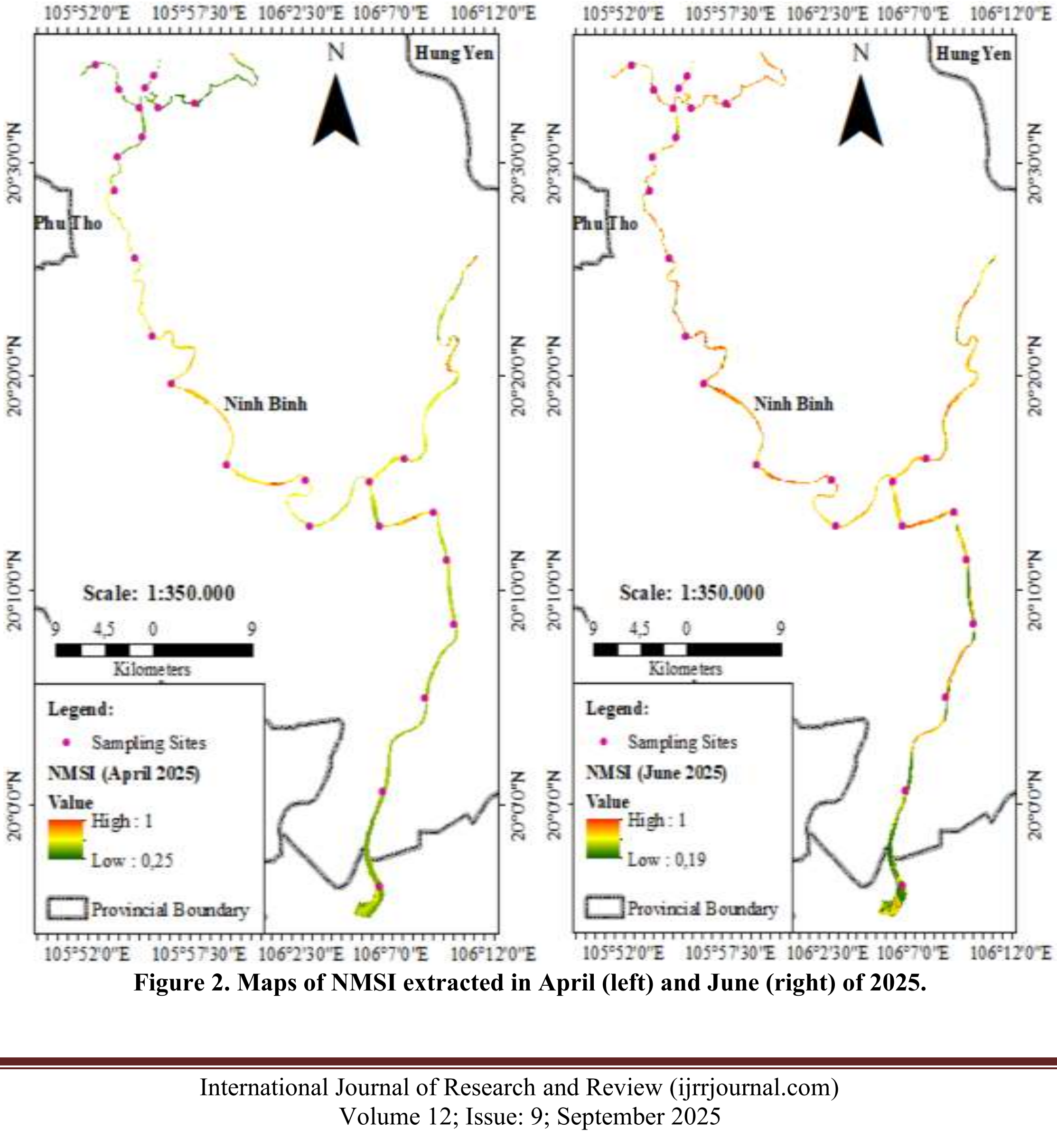

Figure 2. Maps of NMSI extracted in April (left) and June (right) of 2025.

Data from Figure 2 (left) demonstrates NMSI values in April of 2025 ranged from 0.25 to 1.0, with relatively higher values distributed in the upstream reaches near Phu Tho and Hung Yen provinces. These elevated values may be attributed to reduced flow velocity during the dry season, combined with anthropogenic activities such as sand mining and agricultural runoff, which can locally increase suspended material. Midstream and downstream sections, particularly around Ninh Binh, showed moderate to low NMSI values, suggesting relatively clearer water conditions in this period. In contrast, data from Figure 2 (right) in June of 2025 (Figure 2, right) presented a wider range of NMSI values (0.19-1.0), with a noticeable increase in suspended sediment in the middle and lower reaches of the river basin. The rainy season is characterized by higher river discharge and stronger surface runoff, which enhances soil erosion and transports larger amounts of particulate matter into the river system. As a result, higher NMSI values were concentrated in the central and southern sections of the basin, whereas the upstream sections displayed more mixed values with localized hotspots of suspended solids.

Overall, the comparison between April and June highlights the strong influence of seasonal hydrological dynamics on the spatial distribution of suspended sediments in the Day River Basin. The dry season is marked by localized high NMSI values in the upstream reaches, while the rainy season significantly amplifies suspended sediment loads in downstream areas. These findings confirm the applicability of NMSI for detecting spatial and temporal variations in suspended solids and underline the importance of integrating remote sensing indices with in-situ measurements for effective water quality monitoring.

Analysis of the accuracy of TSS estimation

The relationships between measured TSS concentrations and the NMSI extracted from Landsat images in April and June 2025 are presented in Figure 3. Both scatter plots reveal significant positive correlations, confirming that NMSI is a reliable spectral proxy for suspended sediment concentration in the Day River Basin.

For the dry season (April 2025), the regression model demonstrated a strong linear relationship, with $R^2 \approx 0.80$ and RMSE $\approx$ 2.3 mg/L. The regression line closely follows the data cloud, and most sampling points cluster around the fitted line, indicating good predictive performance. The low RMSE value suggests that the model provides accurate estimates of TSS with relatively small residual deviations. The p-value of the slope (< 0.001) further confirms the statistical significance of the relationship. This implies that under low-flow conditions, spectral reflectance captured by NMSI can effectively capture variability in suspended solids.

$$TSS_{April,2025} = 179.6133 \times NMSI_{April,2025} - 35.1931 \quad (3)$$

For the rainy season (June 2025), the regression model performed even better after refining the dataset. The results show $R^2 = 0.8447$ and adjusted $R^2 = 0.8377$, meaning that approximately 84% of the variability in TSS was explained by NMSI. The regression slope for NMSI was 252.66 with p-value = 2.30×10^{-10}, confirming very strong statistical significance. Error metrics also indicated high predictive accuracy, with RMSE = 1.889 mg/L and MAE = 1.121 mg/L, both relatively small compared to the full observed TSS range (10-30 mg/L). These results indicate a strong positive correlation between NMSI_110625 and TSS_180625, and the linear model is well-suited for TSS estimation in the rainy season.

$$TSS_{April,2025} = 252.6556 \times NMSI_{June,2025} - 43.0937 \quad (4)$$

Overall, the analysis demonstrates that both seasonal models are statistically significant and capable of predicting TSS reliably. While the dry-season model captures suspended sediment dynamics effectively under lower flow conditions, the rainy-season model explains a larger portion of the variability and provides smaller prediction errors, confirming the robustness of the NMSI-TSS relationship.

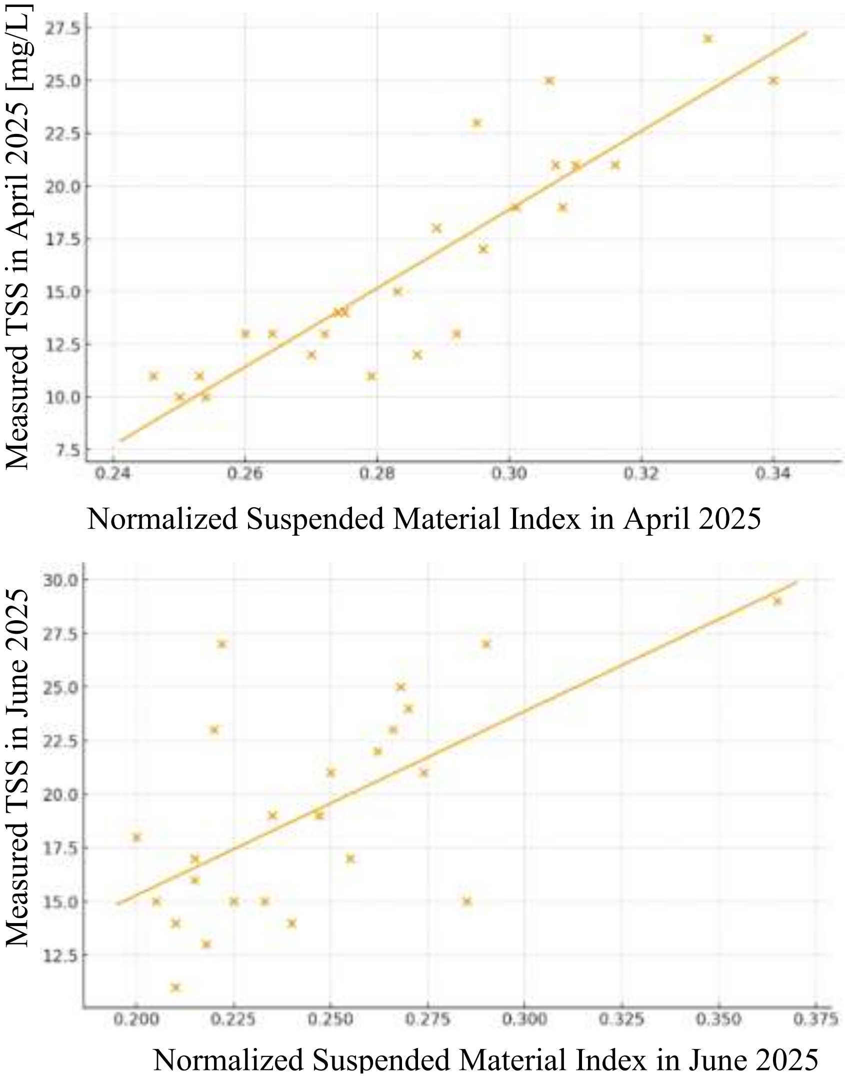

Figure 3. Scatter plots of measured TSS and NMSI in April (upper) and June (lower) of 2025.

Analysis of spatial distribution of predicted TSS concentration

Figure 4 presents the reach-wise maps of predicted TSS derived from the NMSI-TSS regression models for April 2025 (dry season) and June 2025 (rainy season). Predicted concentrations span 10.1-31.1 mg/L in April and 10.8-31.7 mg/L in June, revealing coherent seasonal and longitudinal patterns across the Day River main stem. It can be seen that there exists a seasonal contrast. Both maps show broadly similar spatial structures, but June exhibits a slightly higher upper bound (31.7 mg/L) and more extensive areas of elevated TSS, consistent with enhanced runoff and bank erosion during the monsoon onset. In April, high-TSS segments are patchier and more localized; in June, high values expand and link into longer continuous reaches.

In addition, there is a longitudinal gradient. A clear downstream increase is visible in both months. The lower river, particularly near the confluence/estuarine sections in

Ninh Binh, contains the largest contiguous stretches of high predicted TSS. This likely reflects (i) cumulative upstream inputs, (ii) hydraulic deceleration and meandering that favor sediment resuspension and settling-re-entrainment cycles, and (iii) possible tidal influence near the river mouth that keeps fine particles in suspension. The midstream shows predominantly moderate TSS with intermittent peaks, while the upper reaches (Phu Tho and Hung Yen) remain mostly in the low-moderate range, punctuated by short hotspots at bends and confluences.

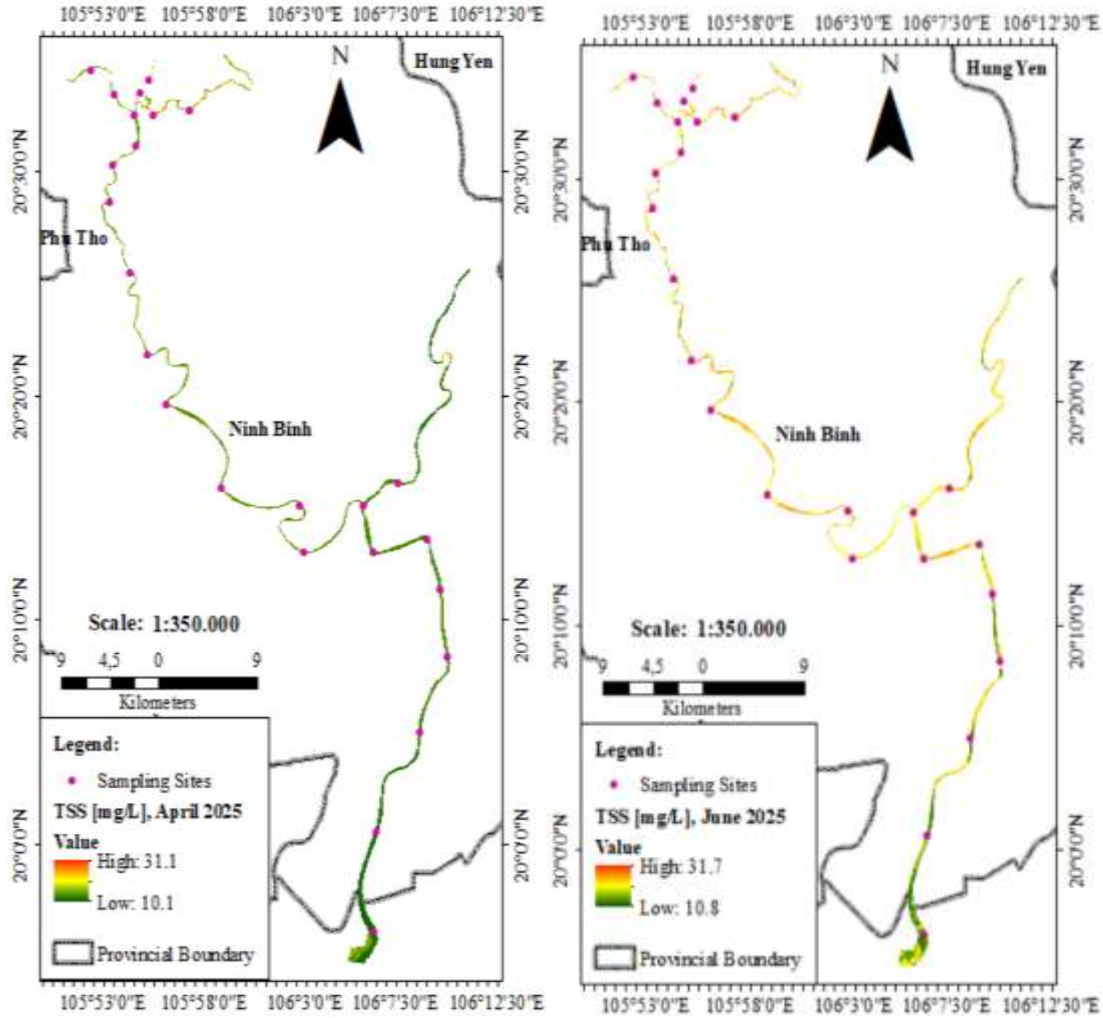

Figure 4. Maps of predicted TSS in April (left) and June (right) of 2025.

Consistency with NMSI and regression results. The mapped TSS patterns mirror the NMSI distributions as discussed above, with high-NMSI zones translating into high predicted TSS. They are also consistent with the statistically significant linear relationships reported previously (dry season $R^2 \approx 0.80$, RMSE ≈ 2.3 mg/L; rainy season models significant but more variable), lending confidence to the spatial extrapolation along the channel network.

Implications for management. The concentration of high predicted TSS in the mid- to lower basin highlights priority zones for sediment-source control (e.g., bank protection at outer bends, riparian buffers at tributary mouths) and for targeted monitoring during early monsoon events when transport increases rapidly. The maps also provide a synoptic baseline for evaluating the effects of dredging, sand mining, and land-use change on suspended sediment dynamics.

Uncertainties and caveats. While the spatial patterns are robust, several factors can affect local predictions: (i) the 30-m Landsat pixel relative to channel width (mixed water-land pixels along narrow reaches), (ii) adjacency effects near banks and small tributaries, (iii) residual atmospheric or bidirectional reflectance artifacts, and (iv) temporal mismatch between image acquisition and field sampling. These limitations argue for combining satellite-based predictions with in-situ spot checks in hotspot reaches and, where feasible, multi-date image compositing to reduce noise.

Overall, Data from Figure 4 demonstrates that the NMSI-based linear models capture the seasonal amplification and downstream concentration of suspended sediments in the Day River. The resulting maps furnish spatially explicit evidence to guide sediment management and water-quality protection at basin scale.

CONCLUSIONS

This study demonstrated the successful application of Landsat imagery and the NMSI index to estimate suspended sediment concentrations in the Day River Basin. By integrating in-situ TSS measurements with regression models, the study achieved reliable estimation accuracy in both dry and rainy seasons. The results showed that the dry-season model explained ~80% of TSS variability, with relatively small errors, showing that NMSI effectively captures suspended sediment dynamics under low-flow conditions. The rainy-season model explained ~84% of variability (adjusted R^2 = 0.8377), with very low RMSE and MAE, confirming the robustness of the NMSI-TSS relationship even under hydrologically dynamic conditions. Spatial distribution maps of predicted TSS highlighted both seasonal contrasts and longitudinal gradients, with increased concentrations in downstream sections during the rainy season. These patterns are consistent with hydrological processes such as runoff, erosion, and resuspension. The results validate NMSI as a reliable proxy for TSS and demonstrate the capability of Landsat imagery to deliver basin-scale, cost-effective monitoring of water quality. While the models performed well, some uncertainties remain due to spatial resolution limits, atmospheric effects, and temporal mismatches between image acquisition and field sampling. Future research should incorporate multi-temporal imagery, additional spectral indices, and advanced statistical or machine learning models to further enhance prediction accuracy. Overall, the study provides strong evidence that satellite-based regression models can complement conventional field sampling, offering timely and spatially comprehensive insights into suspended sediment dynamics for improved management of the Day River Basin.

Declaration by Authors

Acknowledgement: This study received financial support from the Ministry-level Basic Science Program in Chemistry, Life Science, Earth Science, and Marine Science for the period 2017-2025 (Grant No. TNMT.2023.562.07).

Source of Funding: Vietnam Ministry of Agriculture and Environment.

Conflict of Interest: The authors declare no conflict of interest.

REFERENCES

1. Zhang Y, Lin S, Liu J, Qian X, Ge Y. Time-series MODIS image-based retrieval and distribution analysis of total suspended matter concentrations in Lake Taihu (China). Int J Environ Res Public Health. 2010;7(9):3545–60.
2. Arisanty D, Saputra AN. Remote sensing studies of suspended sediment concentration variation in Barito Delta. In: IOP Conference Series: Earth and Environmental Science. IOP Publishing; 2017. p. 12058.
3. Thuy VTT, Ha NTT, Koike K, Thao NTP, Trung PN, Thanh DX. Estimating Water Content and Grain Size of Intertidal Flat Sediments Using Visible to Shortwave-Infrared Reflectance and Sentinel 2A Data: A Case Study of the Red River Delta, Vietnam. IEEE J Sel Top Appl Earth Obs Remote Sens. 2022; 15:2696–708.
4. Warren MB, Bullard SA. First elucidation of a blood fluke (Electrovermis zappum n. gen., n. sp.) life cycle including a chondrichthyan

or bivalve. Int J Parasitol Parasites Wildl. 2019; 10:170–83.

5. Duong HH, Phong ND, Ha TL, Tang TD, Trinh TN, Nguyen TM, et al. Application of Machine Learning to Forecast Drought Index for the Mekong Delta. 2024;
6. Imen S, Chang N-B, Yang YJ. Developing the remote sensing-based early warning system for monitoring TSS concentrations in Lake Mead. J Environ Manage. 2015; 160:73–89.
7. Susiati H, Dede M, Widiawaty MA, Risko, Udiyani PM. TSS in West Kalimantan based on remote sensing data: A preliminary study for siting nuclear power plant. In: AIP Conference Proceedings. AIP Publishing LLC; 2022. p. 20005.
8. Chen W, Wang J, Cao X, Ran H, Teng D, Chen J, et al. Possibility of using multiscale normalized difference vegetation index data for the assessment of total suspended solids (TSS) concentrations in surface water: A specific case of scale issues in remote sensing. Environ Res. 2021; 194:110636.
9. Adjovu GE, Stephen H, James D, Ahmad S. Measurement of total dissolved solids and total suspended solids in water systems: A review of the issues, conventional, and remote sensing techniques. Remote Sens. 2023;15(14):3534.
10. Unnithan SLK, Cherukuru N, Ingleton T, Lehmann E, Paget M, Guo Y, et al. Mapping total suspended solids (TSS) and dissolved organic carbon (DOC) in complex coastal waters using deep learning enhanced remote sensing. Ecol Inform. 2025;103276.

How to cite this article: Danh-tuyen Vu, Thi-thu-huong Pham. Landsat-Derived NMSI index for estimating total suspended solids using linear regression models. *International Journal of Research and Review*. 2025; 12(9): 473-482. DOI: *https://doi.org/10.52403/ijrr.20250946*

******