Applied Water Science (2026) 16:201
https://doi.org/10.1007/s13201-026-02883-1

ORIGINAL ARTICLE

Coupling computational fluid dynamics with Decision Tree and k-Nearest Neighbours classifications for stagnation and water quality assessment in arid rivers: evidence from the Tigris

Mustafa Hathal[1] · **Viktor Sebestyén**[1] · **Stanisław Anweiler**[2] · **Viola Somogyi**[1]

Received: 11 February 2026 / Accepted: 25 May 2026 / Published online: 14 June 2026

Abstract

Urban river systems in arid climates are experiencing extreme hydrological shifts driven by evolving drought patterns and upstream flow regulation. Despite these transformations, the precise physical mechanisms linking reach-scale stagnation to water quality deterioration remain inadequately resolved. In this study, a Comprehensive Water Assessment Framework was developed by integrating k–ε turbulence-based Computational Fluid Dynamics (CFD) with interpretable machine learning (ML) classifiers—Decision Tree and k-Nearest Neighbour. This framework was applied to diagnose flow dynamics in the Tigris River, Baghdad, over the period 2014–2023. It was revealed that flow stagnation responds non-linearly to discharge, with the stagnant fraction expanding to ~72% of the reach length during extreme drought, coinciding with critical Water Quality Index (WQI) collapses (51–55). An empirically derived safe flow threshold of ≈200 m^3/s was identified, below which hydraulic connectivity and self-purification capacity dropped sharply (R^2 = 0.98). Furthermore, WQI variability was shown to be strongly governed by a minimal set of drivers, including stagnation extent, turbidity, pH, TDS, Ca^{2+}, and $PO_4^{3-}$with a dominant stagnation × turbidity interaction highlighted via multivariate response-surface analysis (R^2 = 0.97). As CFD-derived diagnostics were replicated with ML models at <2% error, a high-accuracy, computationally efficient tool for real-time management is provided. These findings offer a scalable outline for drought operations, allowing high-risk urban hotspots to be identified and abstraction strategies to be optimized in water-stressed environments.

Keywords Tigris River · Stagnation area · Water quality · CFD · DT · k-NN · Flow threshold

Introduction

As primary sources for domestic, industrial, and agricultural water, rivers require robust ecological and hydrodynamic management to balance the competing demands of human development and environmental sustainability (Qi et al. 2024). A comprehensive evaluation of river water quality and quantity is the basis of sustainable water resource management, especially in deteriorating climatic variability and anthropogenic stressors (Seo et al. 2024). Such assessments are central for developing robust evidence to inform adaptive policies and ensuring resilient and equitable water governance (Tyagi et al. 2024; Zheng et al. 2025). This need is particularly acute in arid and semi-arid regions, such as the Tigris–Euphrates Basin, where water-scarcity, pollution, and ecosystem degradation collectively threaten the long-term sustainability of riverine systems (Mensoor et al. 2022). Evaluation techniques can be categorized as follows: (1) real-time sensors for hydrodynamics (e.g., measuring river levels and flow rates), (2) remote sensing for detecting pollution sources (Ahmed et al. 2023), and (3) laboratory testing of samples, which includes physical, chemical, and biological analyses (Ewaid et al. 2020). Despite advances in real-time monitoring, predicting river water quality remains difficult due to its nonlinear, dynamic, and non-stationary behaviour (Chen et al. 2024). Techniques such

Communicated by Ahmed M. A-Areeq.

✉ Mustafa Hathal
mus.mhathal@phd.mk.uni-pannon.hu

[1] Faculty of Engineering, Sustainability Solutions Research Laboratory, University of Pannonia, Egyetem Str. 10, Veszprém H 8200, Hungary

[2] Opole University of Technology, Opole, Poland

as Computational Fluid Dynamics (CFD) (Bao et al. 2022) and Machine Learning (ML) (Najah Ahmed et al. 2019) can complement traditional sampling. CFD and ML (supported by GIS) can complement traditional sampling. Still, their integration remains limited forjointly capturing water quantity–quality dynamics under highly variable hydrologic and climatic forcing. Simplified hydrodynamic indices and models can indicate stagnant zones, yet they are coarse and become less reliable when discharge fluctuates strongly (Higashino 2011). Tracer-based investigations have been used to explain pollutant transport in stationary or low-flow regions; nevertheless, these methodologies are often limited to confined experimental environments and brief temporal scales (Abessi et al. 2012). Remote sensing has been used to identify stagnant surface waters and assess their quality; however, it often provides surface-only imagery that does not capture underlying dynamics or temporal fluctuations (Rahman et al. 2011). Previous regional studies of the Tigris River (e.g., Torabi Haghighi et al. 2023) primarily focused on alterations in the flow regime and climatic or upstream regulatory impacts, often neglecting the geometric and hydrodynamic characteristics that control stagnation, recirculation, and pollutant retention.

Computational Fluid Dynamics (CFD) has significantly advanced the understanding of aquatic hydrodynamics and pollutant transport by resolving complex flow fields and concentration gradients (Alfwzan et al. 2024). In riverine research, CFD applications range from simplified one-dimensional (1D) momentum balances to fully three-dimensional (3D) simulations. For instance, Mannina and Viviani (2010) used 1D momentum and mass-balance equations in the Savena River (Italy) to evaluate how discharge affects biochemical oxygen demand (BOD_5). Moving to multi-dimensional analysis, Ismail et al. (2022) employed a 2D advection–dispersion model at the Diyala–Tigris confluence to forecast organic matter and TDS, illustrating how tributary interactions reshape dispersion. For environments requiring higher turbulence resolution, Wang et al. (2023) implemented 3D models with k-ε closures, demonstrating that higher width-to-depth ratios enhance lateral mixing. These studies collectively suggest that contaminant fate is governed by the interplay of flow velocity, mixing intensity, and organic characteristics, which together dictate transport and retention (Zhang et al. 2024).

Water Quality Index (WQI) is a dimensionless indicator that aggregates multiple parameters into a single score to classify water quality (e.g., polluted to excellent) (Nasir et al. 2022). Recent developments have simplified sub-index calculations within the weighted-arithmetic approach while retaining the underlying sub-index structure that underlies WQI computation (Haghiabi et al. 2018). WQI outputs are commonly used either as continuous analysed targets or as categorical water-quality classes (Al-Asadi et al. 2022; Sidek et al. 2024), and several implementations remain robust when observations are incomplete (Asadollah et al. 2021).

In parallel, machine-learning (ML) methods are increasingly used to analysis and classify water quality. Decision Trees (DT) are valued for their interpretability and their ability to represent non-linear relationships in hydro-environmental data (Liu et al. 2023), while k-Nearest Neighbours (k-NN) offers locally adaptive classification that can be effective for spatially continuous datasets, although its performance is sensitive to feature scaling and standardization (Shams et al. 2024). Beyond these, ensemble learners such as Extra Trees have also shown strong performance for WQI prediction (Asadollah et al. 2021). Transfer-learning frameworks combining classifiers (including DT and k-NN) have been applied to map shifts among water-quality classes across basins; however, these approaches remain predominantly statistical and typically do not embed hydrodynamic context or flow-regime coupling (Niazkar and Piraei 2025). Similarly, ML-based drought and hydrological risk assessments quantify climatic and meteorological drivers at basin scales, but do not resolve reach-scale hydrodynamic and geometric controls that govern stagnation, pollutant retention, and localized degradation within river channels (Huang et al. 2025). Consequently, a persistent gap remains between catchment-scale ML forecasting and reach-scale flow–quality interactions in arid systems with strongly fluctuating discharge. While CFD–ML integration is increasingly applied in engineered water systems (e.g., clarifiers) to support process understanding and control (Li and Shatarah 2024), comparable integration for river water-quality vulnerability remains limited. A rare example is Seifi and Riahi-Madvar (2019), who coupled one-dimensional CFD with ML to improve dispersion modelling in open-channel flows, illustrating the potential of CFD–ML coupling for river pollution assessment.

Existing WQI and ML studies are predominantly station-oriented or statistics-driven. In contrast, traditional hydraulic and CFD research rarely quantifies the coupling between stagnation and WQI in a manner that is transferable or directly operational for water treatment plant (WTP) management in arid, regulated rivers. Unlike prior approaches, where WQI is analysed as a standalone indicator or ML classifiers are applied without explicit hydraulic context, this study develops a novel, integrated reach-scale framework specifically actionable for water extraction management. Through k–ε turbulence-based CFD simulations, flow connectivity is mapped and stagnation-prone zones delineated under variable discharge and morphological conditions, thereby providing a physical basis for identifying regions where low-velocity conditions persist, and self-recovery capacity is diminished. Using geometric and hydraulic data, DT and k-NN models classify these CFD patterns to quickly detect stagnation during routine operations. By coupling these stagnation diagnostics with observed WQI behaviour, the link between hydraulic vulnerability and treatment relevance is explicitly established. Accordingly, this research is directed toward mapping reach-scale stagnation under fluctuating discharge regimes, quantifying the linkage between stagnation extent and WQI variability, and developing interpretable ML classifiers to translate complex hydrodynamic outputs into operational stagnation classes for drought-time WTP decision support.

Methodology

Figure 1 illustrates the CFD-based Comprehensive Water Assessment Framework, integrating hydraulic data, water quality parameters, and river geometry into a unified diagnostic system. The workflow executes parallel tracks for WQI calculation and k–ε turbulence-based CFD modelling, ensuring high-resolution mapping of flow dynamics and pollutant transport. Validated outputs are categorized using the Decision Tree and k-NN algorithms to delineate stagnation zones. This coupling enables systemic spatial-temporal visualization and correlation analysis of mechanisms of water-quality deterioration.

Site description

The Tigris River flows through Iraq, ultimately reaching a length of 1418 km. The Tigris River basin, spanning around 340,500 km^2, primarily consists of sedimentary rocks, with igneous and metamorphic formations being confined to the northern and northeastern areas (Hussein et al. 2023). The climate of the catchment area ranges from arid to semi-arid; annual precipitation in the north exceeds 1000 mm, whilst in the south it is below 50 mm (Al-Ansari et al. 2019). Baghdad, the capital of Iraq with a population of over 5.4 million, relies solely on the Tigris River as its potable water source (Kibaroglu and Gürsoy 2015). In Baghdad, raw water is abstracted from the Tigris River through multiple intake stations; this study focuses on ten of them. These withdrawal points coincide with monitoring locations, ensuring that the measured physico-chemical parameters directly represent the quality of water entering the treatment facilities and supplied to the urban population. The network of plants spans from the East Tigris in the north to Al-Wahda in the south, which served as the sampling stations in this study (Fig. 2).

The water level data were retrieved from the Database for Hydrological Time Series of Inland Waters (DAHITI),

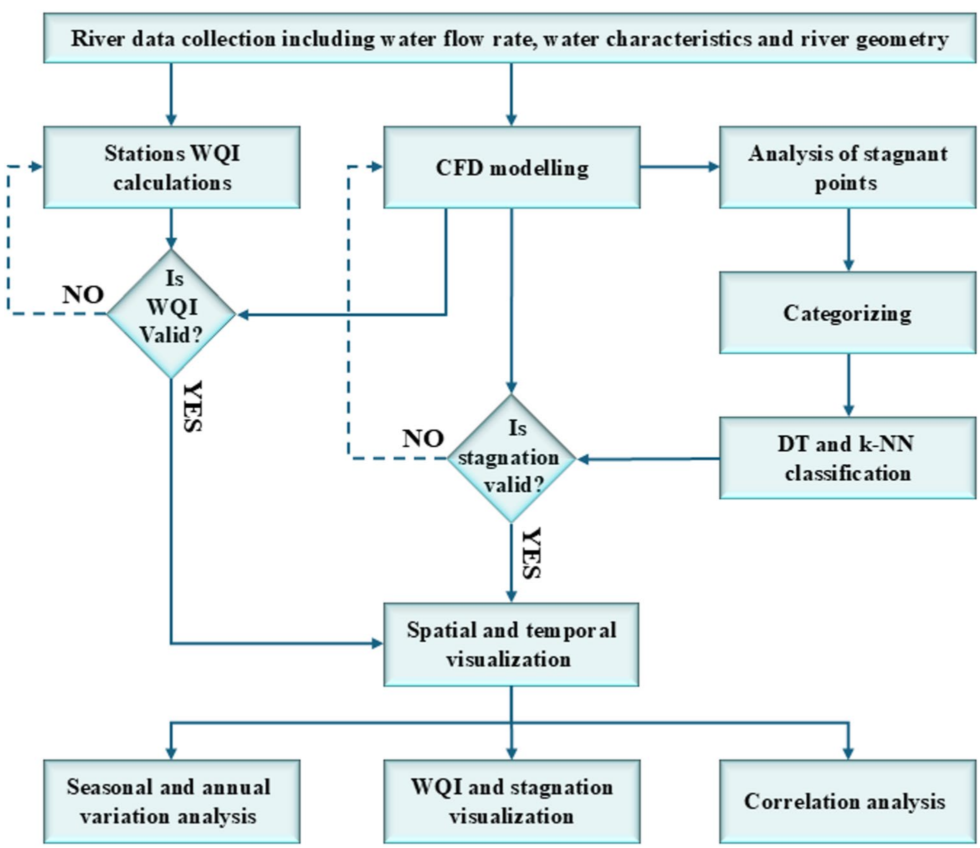


Fig. 1 Workflow of the CFD-based Water Assessment Framework for integrated analysis of flow, stagnation, and water quality in surface water bodies

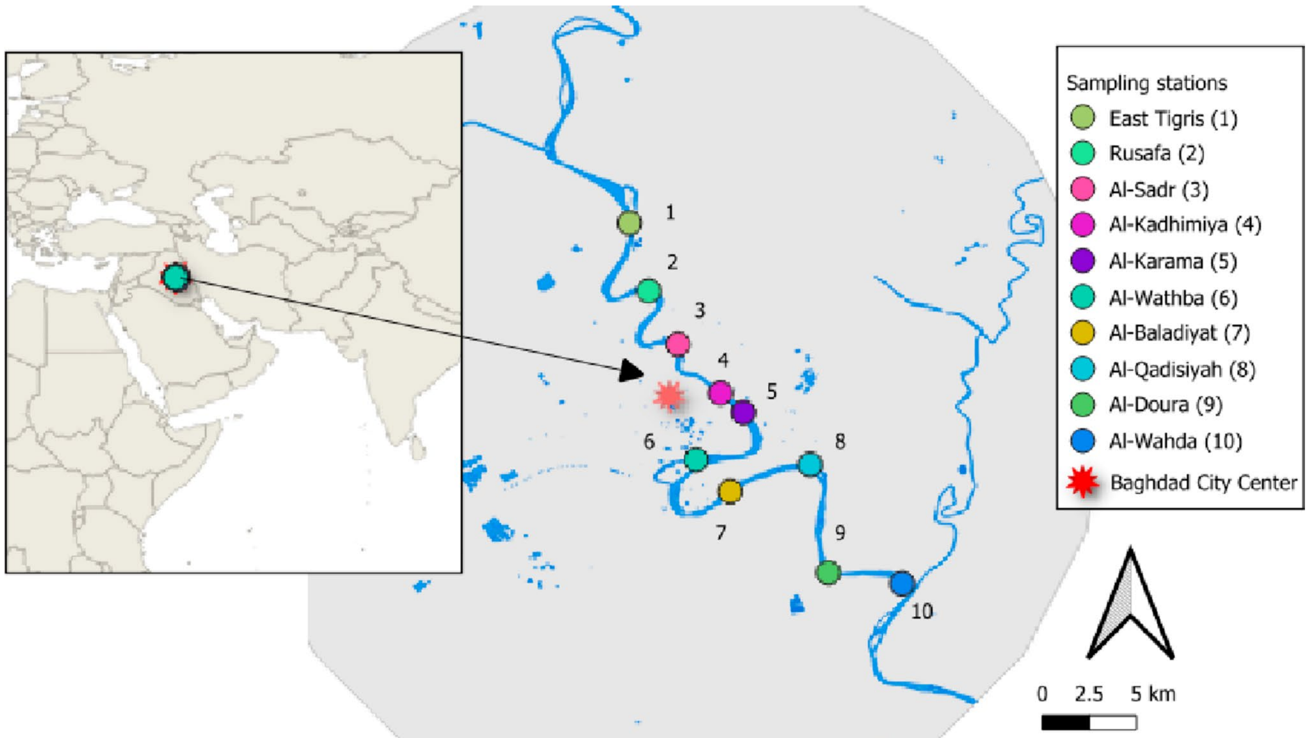


Fig. 2 Tigris River sampling stations in Baghdad. The monitoring sites coincide with municipal water delivery points

created in 2013 ("DAHITI Dataset" updated). DAHITI provides various hydrological data on diverse inland water bodies, including rivers, sourced from satellite photography. Using a correlation (Eq. 1) established with Baghdad Tigris River data from 1952 to 2019 (Hussein et al. 2023), relationships between river level and flow rate were derived to convert DAHITI water-elevation records into continuous flow-rate estimates.

$$Q = 81.263 \cdot (H - 25)^{1.614} \tag{1}$$

where: Q is the flow rate in m^3/s, and H is the water elevation in m. Figure 3 shows the time-series of Tigris River flow rate, precipitation, and air temperature in Baghdad (2014–2023). The upper panel compares river flow (m^3/s) with rainfall (mm/day), showing clear seasonal peaks during wet months and reductions in summer. The lower panel illustrates the corresponding air temperature cycle (°C). The data, extracted from the NASA POWER database ("NASA POWER | Data Access Viewer" updated) at coordinates 33.322°N, 44.3996°E, demonstrates the strong influence of seasonal rainfall patterns on river flow. To represent contrasting hydrological conditions, three reference years were selected: 2019, a moderate-flow year; 2022, a drought year; and 2023, an above-moderate-flow year.

Comprehensive water quality data for the Tigris River Baghdad reach (2014–2023) were sourced from the official monitoring program of the Baghdad Municipality. Monthly grab samples were collected at all active monitoring stations, yielding a decadal time series that captures critical seasonal and interannual variability under diverse hydrological regimes. Field sampling and laboratory analytical procedures were performed in strict adherence to the APHA Standard Methods for the Examination of Water and Wastewater. The 1998 edition was applied to samples collected prior to 2023, while the 2022 edition (24th ed.) was utilized specifically for the 2023 sampling phase (APHA 1998, 2022). Physical parameters, including electrical conductivity (EC) and turbidity, were measured in situ with calibrated sensors, and total dissolved solids (TDS) were estimated using standard EC–TDS conversion relationships. Total suspended solids (TSS) were determined gravimetrically, while total hardness (TH) and calcium (Ca^{2+}) were quantified through EDTA titration. Chloride (Cl^-) concentrations were determined via $AgNO_3$ titration, and nutrient levels (nitrate, NO_3^-; phosphate, PO_4^{3-}) were analyzed using photometric methods. Given that routine monitoring archives may lack exhaustive quality-assurance/quality-control (QA/QC) metadata—such as detailed instrument calibration logs or method detection limits (MDL/LOQ)—a structured data-screening protocol was implemented to ensure dataset integrity. This procedure included unit verification, plausibility/range checks, and internal consistency evaluations. Due to inconsistent reporting across the study period, dissolved oxygen (DO) was excluded from the Water Quality Index (WQI) aggregation to maintain a robust and continuous

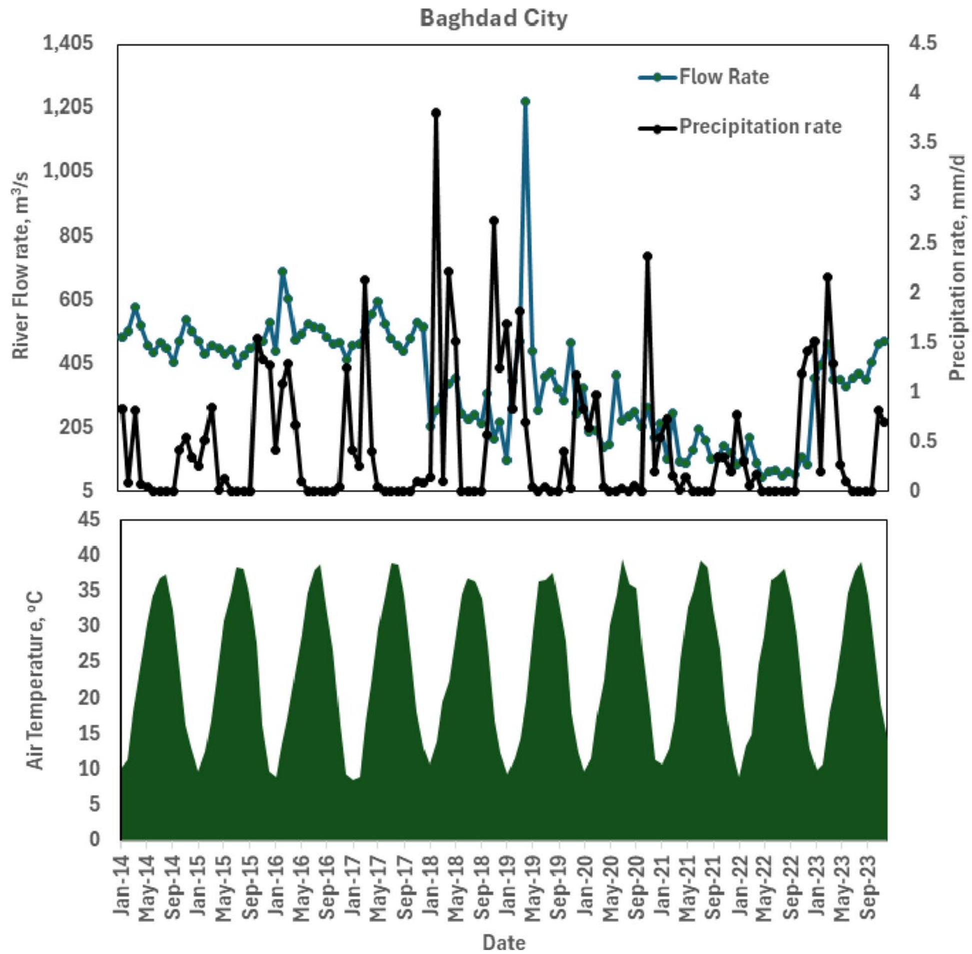


Fig. 3 Time series data of Tigris River flow rate, precipitation, and air temperature

analytical framework. The final parameter suite (pH, TH, TDS, TSS, turbidity, NO_3^-, PO_4^{3-}, Ca^{2+}, and Cl^-) was selected to represent both general hydrochemical characteristics and indicators of anthropogenic nutrient enrichment.

CFD simulation setup

Hydrodynamic modelling was conducted in COMSOL Multiphysics using the k–ε turbulence model coupled with advection–dispersion equations to simulate contaminant transport. This Multiphysics approach simultaneously resolves flow dynamics and pollutant dispersion, providing high-resolution concentration fields and enabling accurate analysis of contaminant fate in riverine ecosystems (Al-Mansori et al. 2020; Ubah et al. 2021). The equations provided embody the essential concepts regulating fluid dynamics and mass transfer. The mass continuity equation (Eq. 2) is the conservation of mass inside the fluid. It asserts that any variation in fluid density over time is attributable to the divergence of the mass flux , where U denotes the velocity field. This equation guarantees that mass is conserved, establishing the fundamental constraint for all subsequent flow computations (Bao et al. 2022).

$$\frac{\partial \rho}{\partial t} = -\nabla . (\rho \cdot U) \tag{2}$$

The momentum balance is solved in RANS form as:

$$\rho\left(\frac{\partial \mathrm{U}}{\partial t} + \mathrm{U} \cdot \nabla \mathrm{U}\right) = -\nabla p + \nabla \cdot \left[(\mu + \mu_t)\left(\nabla \mathrm{U} + (\nabla \mathrm{U})^T\right)\right] + \rho \mathrm{g} + \mathrm{F}_v \tag{3}$$

The term $\mathbf{F}_v$ represents additional distributed (volume) forces implemented to account for reach-scale resistance effects (e.g., bed-friction drag). In shallow-water-type applications, bed resistance can be represented via a Manning-type quadratic drag written as a momentum sink (Pu 2015):

$$F_v = -\rho C_f \mid U \mid U, \; C_f = \frac{gn^2}{R_h^{4/3}} \tag{3a}$$

Where n is Manning roughness, R_h is the hydraulic radius (the water level can be used for this variable), and g is gravitational acceleration. This flow specification, together with the turbulence closure provided by the *k–ε* model (Eq. 4), allowed the CFD simulations to resolve both the mean velocity distribution and turbulence effects under realistic hydrodynamic conditions (Ismail et al. 2022).

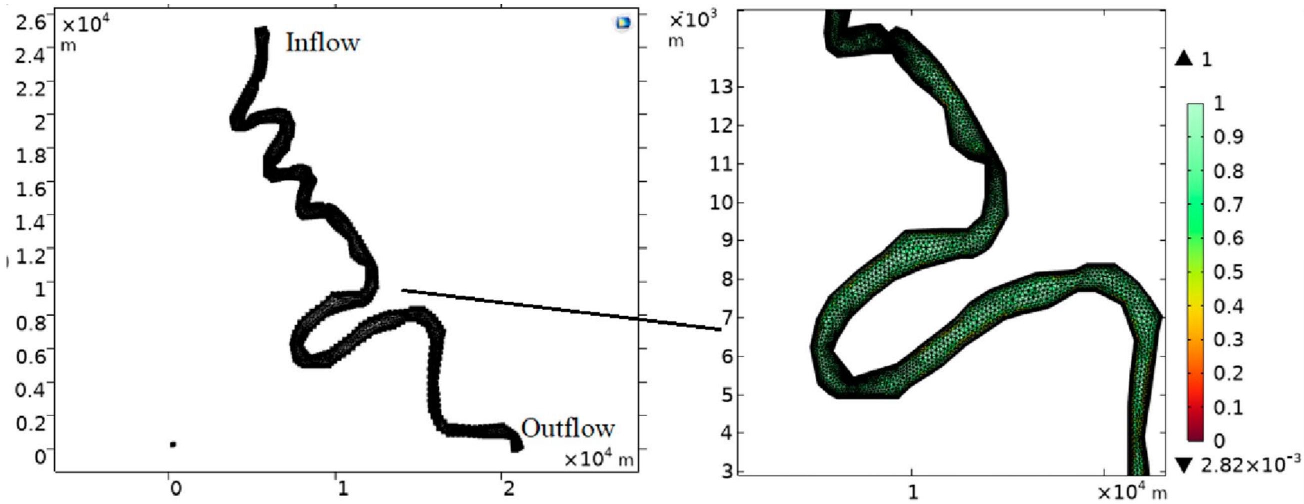


Fig. 4 Baghdad Tigris River geometry and mesh details

$$\frac{\partial\ (\rho\ .\ k)}{\partial\ t}+\nabla\ .(\rho\ \cdot\ k\cdot\ \ U)=\ P_k\ \ -\ \rho\ \cdot\ \ \ +\nabla\ .\left[\left(\mu\ \ +\frac{\mu_{\ t}}{\sigma_{\ k}}\right)+\ \nabla\ .k\right] \tag{4}$$

$$\frac{\partial\ (\rho\ \ )}{\partial\ t}+\nabla\ .(\rho\ \cdot\ \ \cdot U\ )=\frac{}{k}\cdot\ (\ C_1\cdot\ \ P_k-C_2\cdot\ \ )+\nabla\ .\left[\left(\mu\ \ +\frac{\mu_{\ t}}{\sigma}\right)+\ \nabla\ .\ \right] \tag{5}$$

The turbulence production term is computed from the mean strain-rate tensor (S):

$$P_k=2\mu_{\ t}\mathrm{S}:\mathrm{S}\ \text{with}\ \mathrm{S}=\frac{1}{2}\left(\nabla\ \mathrm{U}+(\nabla\ \mathrm{U})^{\mathrm{T}}\right) \tag{5a}$$

and eddy viscosity:

$$\mu_{\ t}=\rho\, C_{\mu}\,\frac{k^2}{\epsilon} \tag{5b}$$

The values of the model constants are taken from ref Ismail et al. 2022): $\sigma_k=1.0$, $\sigma_\varepsilon=1.3$, $C_1=1.44$, $C_2=1.92$.

Equation (6) represents the scalar transport equation for pollutant concentration $c$in water. In this study, it was applied to key physico-chemical parameters expressed in mg/L, namely TH, [$CaCO_3$-eq.], TDS, TSS, NO_3^-, PO_4^{3-}, Ca^{2+}, and Cl^-. For the CFD simulations, the observed concentrations of these parameters were prescribed as mass transport boundary conditions, with each sampling station assigned its own monthly input values for 2023. The inlet boundary condition was defined using river flow rates from Eq. (1) and time-series data (Fig. 3), with flow rates and water quality characteristics measured below the East Tigris station. The outlet was set to reference pressure (~0) to represent free downstream flow, while the remaining monitoring stations were treated as internal validation points, providing observed concentrations for model calibration. This station-wise and time-resolved specification enabled the model to capture both the spatial heterogeneity among stations and

Table 1 Mesh independence analysis for CFD simulation of the Tigris River

Case	Elements (N)	Error as compared to finest (%)
M1–Coarse	5,634	11.11
M2–Medium	7,645	8.08
M3–Fine	9,821	3.03
M4–Finest	11,871	0.00

the seasonal variability in pollutant loads along the Baghdad reach of the Tigris River. While D represents the diffusion coefficient in m^2/s, which dictates the velocity with which the chemical disperses through the fluid medium, especially in research concerning contaminant dispersion in river systems (Alfwzan et al. 2024).

$$\frac{\partial\ c}{\partial\ t}+U.\ \nabla\ c=\nabla\ \left(-D\cdot\ \nabla\ .c\right) \tag{6}$$

Figure 4 presents the geometric configuration and computational mesh of the Tigris River in Baghdad. The modelled reach spans ~45 km with variable width, reflecting the river's natural meandering. The CFD mesh was refined near riverbanks, intakes, and regions with steep velocity or concentration gradients to improve accuracy. The final mesh includes 8716 vertices and 11871 elements, with an average quality of 0.79 (average mesh skewness of 0.21, within the recommended range (<0.33) for reliable CFD simulations (Al-Farraji and Hathal 2022). Table 1 presents the mesh-independence analysis for the CFD model of the Tigris River using four increasingly refined meshes (M1–M4). Although M3 was acceptable, the final mesh M4 was adopted for robustness.

The CFD model was developed to resolve reach-scale hydrodynamic controls on flow connectivity and stagnation under seasonal and interannual discharge variability.

To ensure computational feasibility for decadal simulations while maintaining physical interpretability, several foundational assumptions were adopted. The fluid was modelled as a single-phase, incompressible medium with constant density and dynamic viscosity. While this approach omits temperature-induced viscosity fluctuations and potential density stratification (e.g., thermal or salinity gradients), the primary drivers of depth-averaged velocity in this system—discharge, bed friction, and reach morphology—are expected to exert a dominant influence on stagnation metrics compared to secondary fluid-property variability (Ismail and Muntasir 2018). For the simulation, representative constant values were adopted based on the prevailing temperature range (8–40 °C) and salinity levels (< 1 g/L), where density remains within ~995–1000 kg/m³ and dynamic viscosity ranges from ~0.65 to 1.4 mPa·s (Herrera-Granados 2009). The spatial and temporal distribution of stagnation was quantified by identifying areas where the resultant velocity fell below a critical threshold. As supported by literature, an initial threshold of 0.1 m/s was considered (Gualtieri 2008). The calibration of this threshold was performed using a sensitivity-based approach, as illustrated in Fig. 5. A monotonic increase in the stagnation area relative to the threshold velocity confirmed the physical consistency of the classification. The stability of the response curves across diverse months suggests that the threshold is temporally robust and independent of seasonal anomalies. A pronounced sensitivity zone was identified near 0.1 m/s, representing a transition regime between advective and quasi-stagnant flow conditions. Thresholds below this value tended to underestimate stagnant zones, while higher values led to saturation effects that diminished discriminatory power. Consequently, 0.1 m/s was selected as the optimal threshold for all subsequent analyses.

Suspended sediment transport and morphodynamic feedback were not explicitly simulated. Consequently, the model does not account for short-term turbidity pulses or bed evolution. Instead, sediment-related deterioration is implicitly accounted for through the stagnation indicator; persistent low-velocity zones are physically synonymous with increased susceptibility to fine-particle accumulation and reduced self-cleansing capacity (Murib 2018). Furthermore, a 2D depth-averaged approach was prioritized over 3D modeling to facilitate reach-scale mapping across a 10-year period. While vertical velocity gradients and secondary currents at meanders or confluences are not resolved, the 2D approach effectively characterizes reach-scale stagnation susceptibility and its temporal evolution (Ismail et al. 2022). The influence of bed topography and sediment-mediated hydraulic resistance was rigorously incorporated through a literature-validated range of Manning's n coefficients calibrated explicitly for the Baghdad reach. This procedure facilitated the derivation of a quantified uncertainty envelope for the reach-scale velocity field, thereby establishing a robust physical foundation for the delineation of stagnant zones. The sensitivity of the predicted stagnation extent to bed roughness was found to be highly regime dependent. Under drought flow conditions (Q ≈ 223 m³/s), the model exhibitedits highest sensitivity; an increase in n from 0.02 to 0.04 resulted in the stagnant fraction expanding from ~63% to ~74%. This pronounced response is attributed to the prevalence of low-flow velocities near the defined stagnation threshold (V < 0.1 m/s), where minor frictional perturbations are sufficient to reclassify significant portions of the domain. In contrast, bed roughness exerted a negligible influence during moderate (Q ≈ 500 m³/s) and high-flow (Q ≈ 1500 m³/s) regimes, with stagnation remaining consistently low at < 15% and < 5%, respectively. The adopted roughness range is consistent with reported Manning values for the Tigris River in Baghdad, where n values between 0.025 and 0.035 are recommended based on bed morphology and local elevation variability Asaad and Abed (2020).

The bends and shallow cross-sections of the river channel significantly alter local flow dynamics. To account for

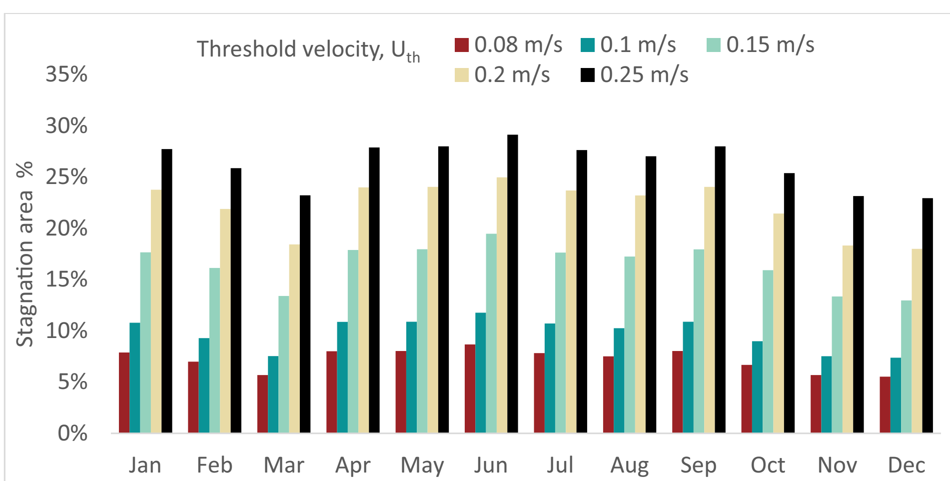


Fig. 5 Stagnation area calibration for choosing threshold velocity based on 2023 seasonal results

Table 2 Rivers ecological system standard values and proposed weights

Parameters	Standard values (S_i)	Weight (w_i) [e]
pH	8.5[a, b, c]	0.147316
TH, CaCO3 Equivalent (mg/L)	500[b]	0.018114
TDS (mg/L)	1000[b]	0.018551
TSS (mg/L)	25[b]	0.02117
Turbidity (NTU)	45[b]	0.319948
NO_3^- (mg/L)	40[b]	0.154736
PO_4^{3-} (mg/L)	5.5[b, d]	0.227848
Ca^{+2} (mg/L)	150[b]	0.084679
Cl^- (mg/L)	350[b]	0.007639
		sum = 1

The parameter standards and relative weights are synthesized from various authorities as follows: [(a)]APHA (2022); [(b)]Iraqi national standards as cited by Mahdi et al. (2021) [(c)]WHO (2022) guidelines; [(d)] Mensoor et al. (2022); and [(e)] weighting factors adapted from Uddin et al. (2021)

the influence of channel geometry and evaluate the intensity of secondary circulation across varying flow regimes, the dimensionless Dean number (De) was employed (Ma et al. 2026). The Dean number quantifies the relative magnitude of centrifugal and viscous forces and is expressed as:

$$De = Re\sqrt{\frac{2d}{R}} \tag{7}$$

where Re is the Reynolds number based on the average flow velocity ($Re = \frac{\rho\ d\,U}{\mu}$), d represents the localized hydraulic diameter of the channel section, and R is the radius of curvature of the channel's centerline.

Water quality index (WQI)

The monitoring dataset (2014–2023) was analysed through two complementary approaches to capture multi-scale variability. Temporal averaging across the decadal record was performed to assess spatial convergence and long-term station comparisons, while monthly spatial averaging across all monitoring stations was conducted to evaluate seasonal dynamics. Particular emphasis was placed on the contrasting hydrological conditions of the 2022 drought and the subsequent 2023 recovery period. The WQI aggregation was conducted using the weighted arithmetic (Brown) approach, as shown in Eq. (7). The parameter weights (relative weights) were adopted from Uddin et al. (2021), ensuring that each variable's contribution to the composite index was standardized. Within this framework, a higher weight was assigned to turbidity to reflect its significance as a rapid-response indicator of particulate loading and river impairment in urban reaches, where discharge-driven sediment mobilization frequently dictates short-term water quality variability.

Table 3 Water quality evaluation table (Ewaid et al. 2020)

Quality class	WQI range
Excellent	91–100
Good	76–90
Poor (fair)	51–75
Bad	26–50
Polluted	0–25

Iraqi and international standards/guideline limits are taken from Table 2: Rivers ecological system standard values and proposed weights (Mahdi et al. 2021; Majeed 2024), as shown in Table 2.

$$WQI = \sum w_i \cdot Q_i \tag{8}$$

For each parameter, the observed value (*vi*) was converted into a quality rating (*Qi*) relative to its standard permissible value (*si*) and an ideal value (v_o), which is mentioned in Eq. 8.

$$Q_i = 100 \cdot \left(\frac{S_i - v_i}{S_i - v_o}\right) \tag{8a}$$

Where v_o was taken as zero for all parameters except pH, which has an ideal value of 7. Each rating was then multiplied by its assigned weight (*wi*), reflecting the parameter's relative importance to aquatic health based on WHO and Iraqi river water standards. The final WQI was calculated as the weighted summation of these scores, providing a single quantitative value of river water quality. Classification thresholds followed Ewaid et al. (2020) as shown in Table 3, which divide WQI into categories of Excellent, Good, Poor, Bad, and Polluted. The standard values and suggested parameter weights utilized in the calculations are comprehensively summarized in Table 2. These values were synthesized from a diverse range of authoritative sources and international organizations, including WHO (2022), APHA (2022), and Iraqi national standards, ensuring both methodological rigor and consistency with established regional studies. The standard value (Si) for pH was assigned as 8.5, representing the upper permissible limit for drinking water as stipulated by the World Health Organization (WHO 2022) and the American Public Health Association (APHA 2022). As detailed in Eq. 8, the calculation of the quality rating (Qi) for pH incorporates both the standard limit (Si = 8.5) and the ideal value (vo = 7.0). This dual-reference approach ensures that the index accurately reflects the acid-base equilibrium of the Tigris River, which is governed by the carbon dioxide-bicarbonate-carbonate system (APHA 2022). By utilizing the 8.5 threshold as the standard permissible limit, the model maintains consistency with regional water quality protocols and international regulatory benchmarks.

Statistical analysis was conducted using a combination of MINITAB and Python to ensure robust inference and cross-validation of relationships among variables. MINITAB was employed to perform response-surface regression, analysis of variance (ANOVA), multicollinearity diagnostics (Variance Inflation Factor, VIF), and model selection based on statistical significance and predictive performance. In parallel, a Pearson correlation analysis was implemented in Python to quantify linear associations between stagnation metrics, water-quality parameters, and WQI, and to verify the consistency of trends identified by the regression-based framework. The combined use of ANOVA-based inference and correlation analysis enabled the identification of dominant drivers from redundant or collinear variables, thereby strengthening the physical interpretability of the final model.

Stagnant area statistics using decision tree and k-nearest neighbours classification

CFD-derived datasets comprising monthly stagnation percentages and spatial velocity profiles were used as supervisory inputs for machine learning (ML) to analyze stagnation dynamics. Two specific algorithms, DT and k-NN, were implemented. DT was selected for its transparent, threshold-based architecture, which is ideal for identifying hydrodynamic regime boundaries and communicating rule-based decision logic (Raviteja et al. 2023). Conversely, k-NN was employed as a non-parametric complement to yield locally adaptive boundaries, facilitating the representation of spatially continuous river features and gradual transitions between stagnant and non-stagnant zones (Shams et al. 2024). These models were intentionally prioritized over more complex "black-box" learners to ensure operational interpretability and robustness within water management workflows. To provide unbiased performance estimation, a leave-one-month-out (LOMO) protocol was executed using MATLAB's Classification Learner. For each month of the 2023 dataset, models were trained on the remaining 11 months, with training further refined via internal 5-fold cross-validation. Hyperparameters—including DT complexity controls (maximum splits, minimum leaf size, and pruning) and k-NN settings (k-value, distance metrics, and weighting)—were optimized within each LOMO training set using a nested design to prevent optimistic bias. Feature standardization was restricted to training folds to eliminate information leakage into the held-out test month. Following the 2023 LOMO cycle, final models were retrained on the complete 2023 dataset and externally validated against 2019 and 2022 data to assess generalization across contrasting hydrological regimes, including extreme drought. The classification features integrated geometric and hydraulic descriptors, including signed local curvature (κ)—estimated via quadratic fitting in a PCA-aligned tangent–standard frame (Ratliff 2019)—bend intensity, cumulative bend count, local orientation, and proximity to the bank or thalweg. These were synthesized with CFD-derived velocity distribution features. The interpretability analysis reveals that the flow rate acts as the root node for the primary split, as it establishes the system's total kinetic energy. This initial split reflects the high sensitivity of the local velocity distribution to global discharge changes. Once the flow rate is defined, the Decision Tree incorporates Local Curvature (κ) and Geometry-derived features to pinpoint where the velocity is most likely to drop below the stagnant threshold (0.1 m/s). This logic ensures that the model captures the physical process where the coupling between flow rate and geometry dictates the velocity field, rather than just correlating variables, thereby identifying the precise boundaries of stagnation zones.

Model performance was quantified using a suite of class-aware metrics, including precision, recall (sensitivity), and the F1-score, as presented in Eqs. 9–12 (Shen and Cheng 2022). These scores are used to ensure that the classification of the stagnant class (velocity < 0.1 m/s) was robust despite shifts in class prevalence, particularly during the high-stagnation periods of 2022. The ML classifiers (DT and k-NN) were trained on high-resolution spatial velocity data. For each monthly simulation run, velocity vectors were extracted from 7788 discrete mesh nodes distributed across the reach. To create a robust binary classification target, a hydraulic threshold of 0.1 m/s was applied: nodes with velocities below this threshold were encoded as "Stagnant" (1), while those above were encoded as "Flowing" (0). To ensure the model captures seasonal and inter-annual variability, a Leave-One-Month-Out (LOMO) cross-validation protocol was implemented. For the 2023 baseline year 85,668 instances (11 months) were used for training and 7788 instances (1 month) for testing in each fold. To assess generalization across extreme hydrological shifts, the full 2023 dataset (93,456 instances) was used to train the classifiers, which were then tested against independent monthly data from 2019 to 2022. The confusion matrices for training (Table D.1) and testing across independent years (2019, 2022, and 2023) (Table D.2) demonstrate that the classifiers learn local hydraulic patterns (e.g., area-velocity relationships) rather than fixed stagnation percentages. This explains the model's high generalizability: by learning the local mechanics of flow separation at the node level, the ML framework successfully predicts high-stagnation regimes (e.g., 72% in 2022) even when trained on lower-stagnation years. To further validate the selection of DT and k-NN, they were benchmarked against alternative classifiers (QDA, Naïve Bayes, and SVM) under identical data splits. This rigorous framework ensures that the ML layer functions as a reliable, interpretable diagnostic tool that translates complex hydrodynamic fields into repeatable stagnation classes for operational decision support.

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}} \tag{9}$$

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \tag{10}$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \tag{11}$$

$$\text{F1} = \frac{2 \cdot \text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}} = \frac{2\text{TP}}{2\text{TP} + \text{FP} + \text{FN}} \tag{12}$$

Results

The primary findings of this research are presented with a focus on the hydrodynamic controls of the Tigris River and their systemicimplications for water quality. Stagnant zones are characterized by seasonal and interannual flow variability, followed by a spatiotemporal analysis of the Water Quality Index (WQI) across the monitoring network. Furthermore, the outputs of the CFD, DT, and k-NN models are validated against empirical data, and the mechanistic correlation between reach-scale stagnation and chemical degradation is established.

Stagnation area

As shown in Fig. 6, the extent of stagnation is primarily governed by hydrological forcing and its associated seasonal and interannual variability. A distinct inverse relationship was observed between river discharge and stagnant-area coverage. High-flow periods, characteristic of wet-season conditions, maintained low and stable stagnation fractions, indicating robust hydraulic connectivity and enhanced flushing that suppresses low-velocity zones. Conversely, declining discharge during dry seasons resulted in a marked expansion of stagnant areas. This shift reflects reduced momentum and weakened mixing, promoting flow separation and recirculation within planform-controlled zones such as inner meander bends and widened reaches. This response was significantly amplified at a multi-year scale.

The extreme low-flow phase of 2022 was accompanied by peak stagnation values of ~49–72%, confirming that drought-regime discharge levels trigger a non-linear expansion of stagnation hotspots. Following the partial recovery of flow in 2023, stagnation levels decreased toward baseline conditions, reinforcing the conclusion that dynamics in the Baghdad reach are predominantly discharge-driven, with extreme drought acting as a critical modulator. The historical progression for the driest month of each year (2014–2023) is illustrated in Fig. 7. Stagnation remained relatively constrained between 2014 and 2018 (7–21%), followed by a progressive escalation to 34% in 2019 and 36% in 2021, eventually reaching a zenith of 72% in May 2022. Conditions ameliorated in 2023, with the stagnant fraction contracting to approximately 12%. These trends align closely with hydro-meteorological data, in which reduced discharge and precipitation directly correlate with stagnant expansion. Spatially, the 2022 stagnation was extensive and contiguous, particularly in segments downstream of Rusafa and Al-Kadhimiya, and throughout the Al-Qadisiyah–Al-Doura–Al-Wahda reach. In these zones, tight meanders and shallow bathymetry promoted flow deceleration and large-scale recirculation. By 2023, these clusters had fragmented, leading to improved longitudinal connectivity, with residual pockets concentrated near Al-Kadhimiya, Al-Karama, and Al-Sadr.

Spatial and temporal analysis of WQI

Analysis of Water Quality Index (WQI) dynamics reveals a persistent longitudinal decline coupled with pronounced seasonality, establishing a deterministic link between spatial

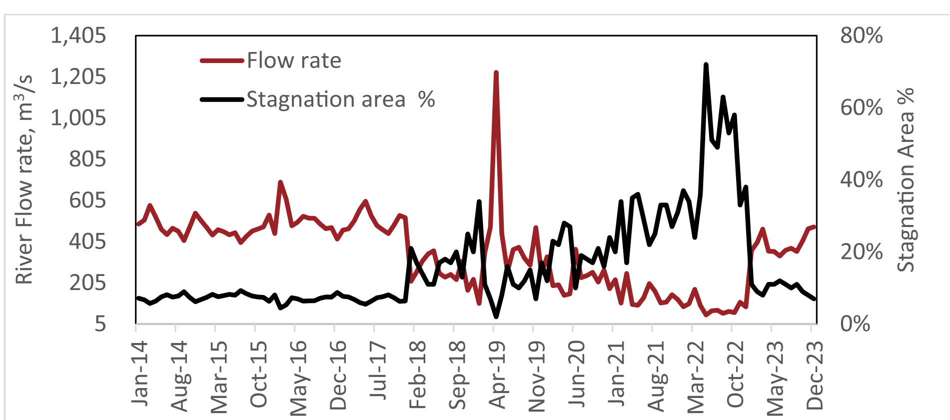
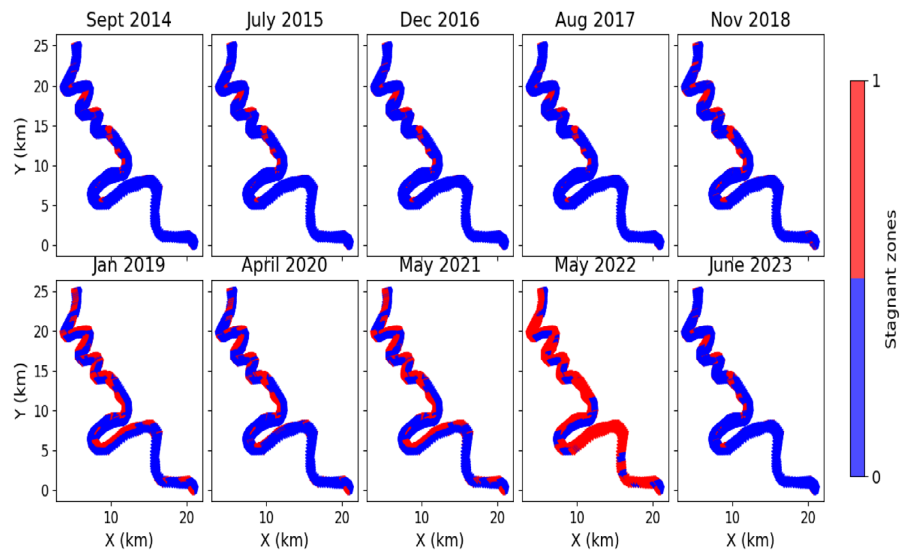
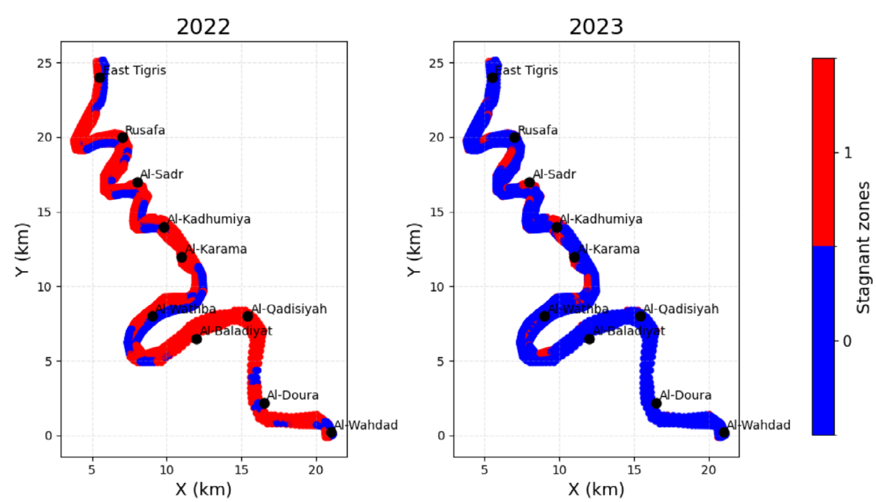


Fig. 6 Stagnation area % for various flow conditions of Baghdad Tigris River based on CFD solution

Fig. 7 Stagnant area profile of the Baghdad Tigris River downstream of the maximum drought conditions between 2014 and 2023



(a) stagnation zones during drought conditions.



(b) comparing between 2022 and 2023 drought months.

gradients and hydrological variability. As illustrated in Fig. 8, decadal averages (2014–2023) demonstrate a progressive downstream degradation gradient. WQI values decrease from ~65 at East Tigris (upstream) to ~60.6 at Al-Wahda (downstream). Upstream reaches are characterized by lower concentrations of TDS (~464–586 mg/L), TH (~284–320 mg/L), and turbidity (~25 NTU). Conversely, mid-city stations exhibit significantly higher turbidity (~96 NTU) and TDS (600–631 mg/L), resulting in WQI values in the moderate-to-poor range (~60–63). Temporally, as shown in Fig. 9, WQI values peak during winter and early spring (66–68) before declining to seasonal minima (51–55) between July and September. A partial recovery in late autumn (61–66) coincides with seasonal precipitation. The high congruence between simulated and observed data (1–6% error) validates the model's ability to capture the complex spatiotemporal dynamics of the Tigris system.

The monthly reach-scale analysis presented in Fig. 10 identifies hydrodynamic stagnation as a primary indicator of hydraulic vulnerability, particularly regarding turbidity-driven deterioration. Degraded water quality is non-uniformly concentrated within high-stagnation regimes (>40–50%), where prolonged low-connectivity hydraulic states are inherently associated with lower WQI values. Among the examined parameters, turbidity showed the most pronounced dependence on stagnation, with larger stagnant areas coinciding with elevated levels due to weakened self-cleansing and increased fine-particle retention. In contrast,

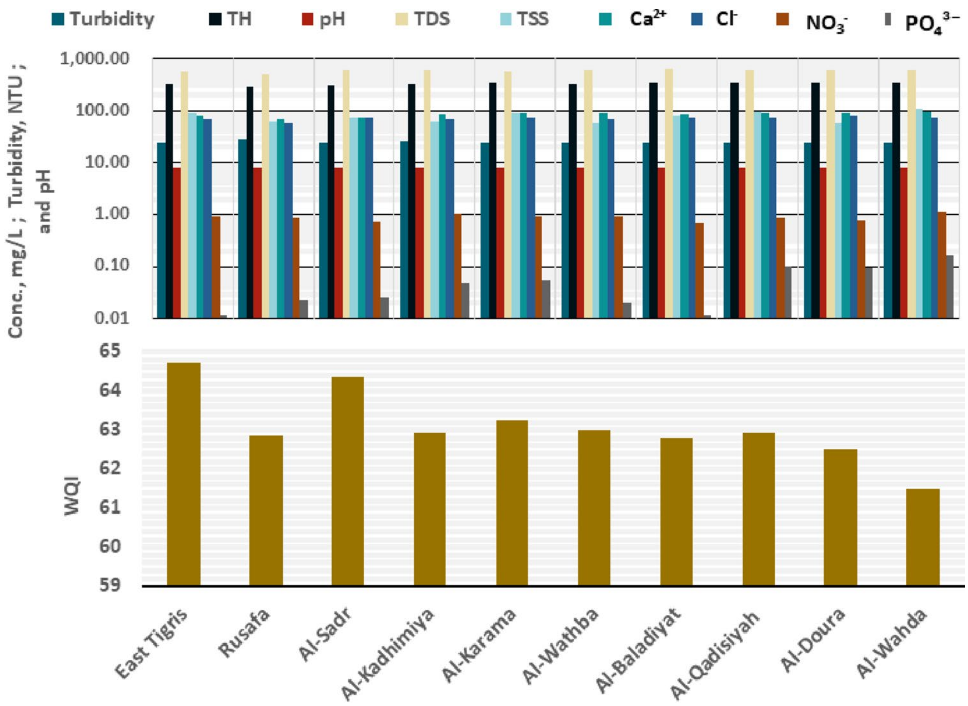


Fig. 8 Average water characteristics and WQI of various stations on the Tigris River (2014–2023)

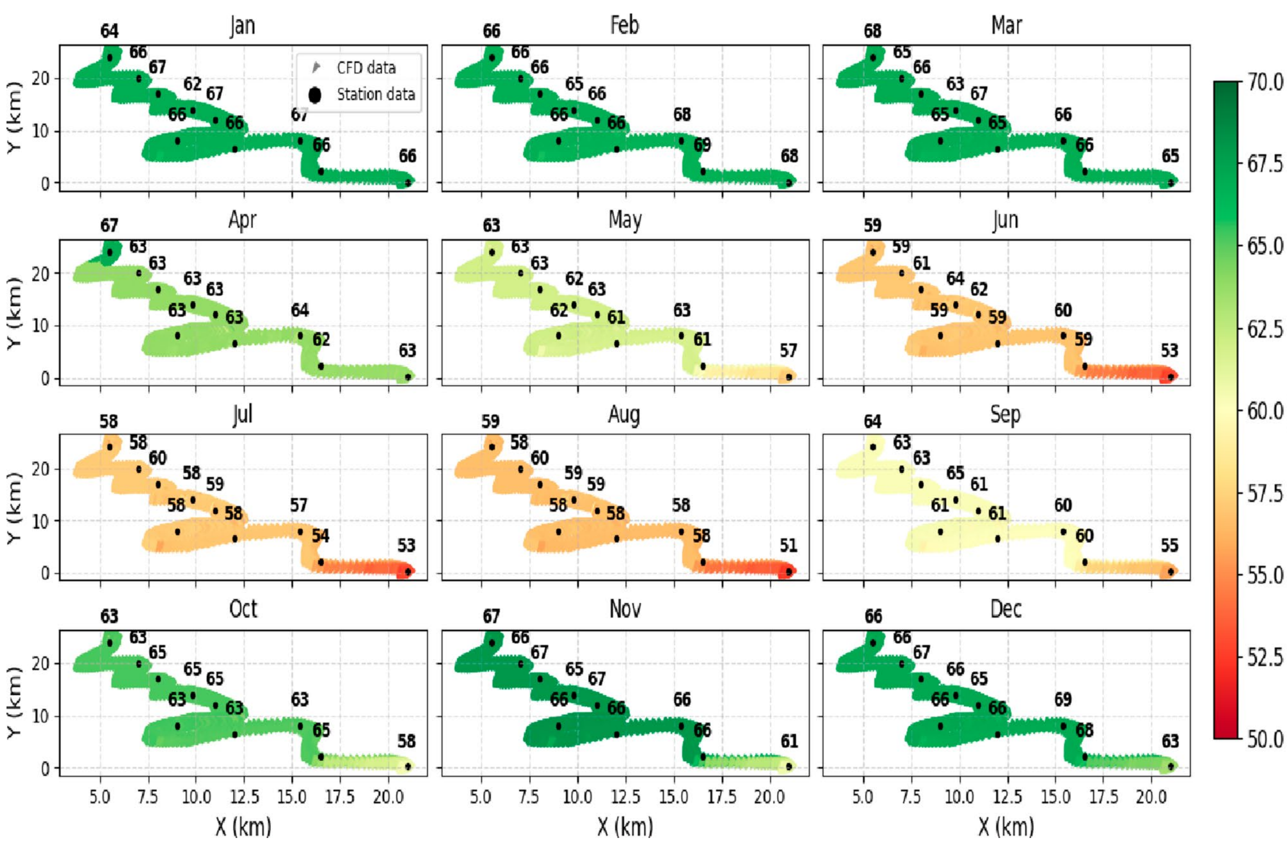


Fig. 9 WQI spatial distribution in the Tigris River, and comparison between CFD and stations data (2023)

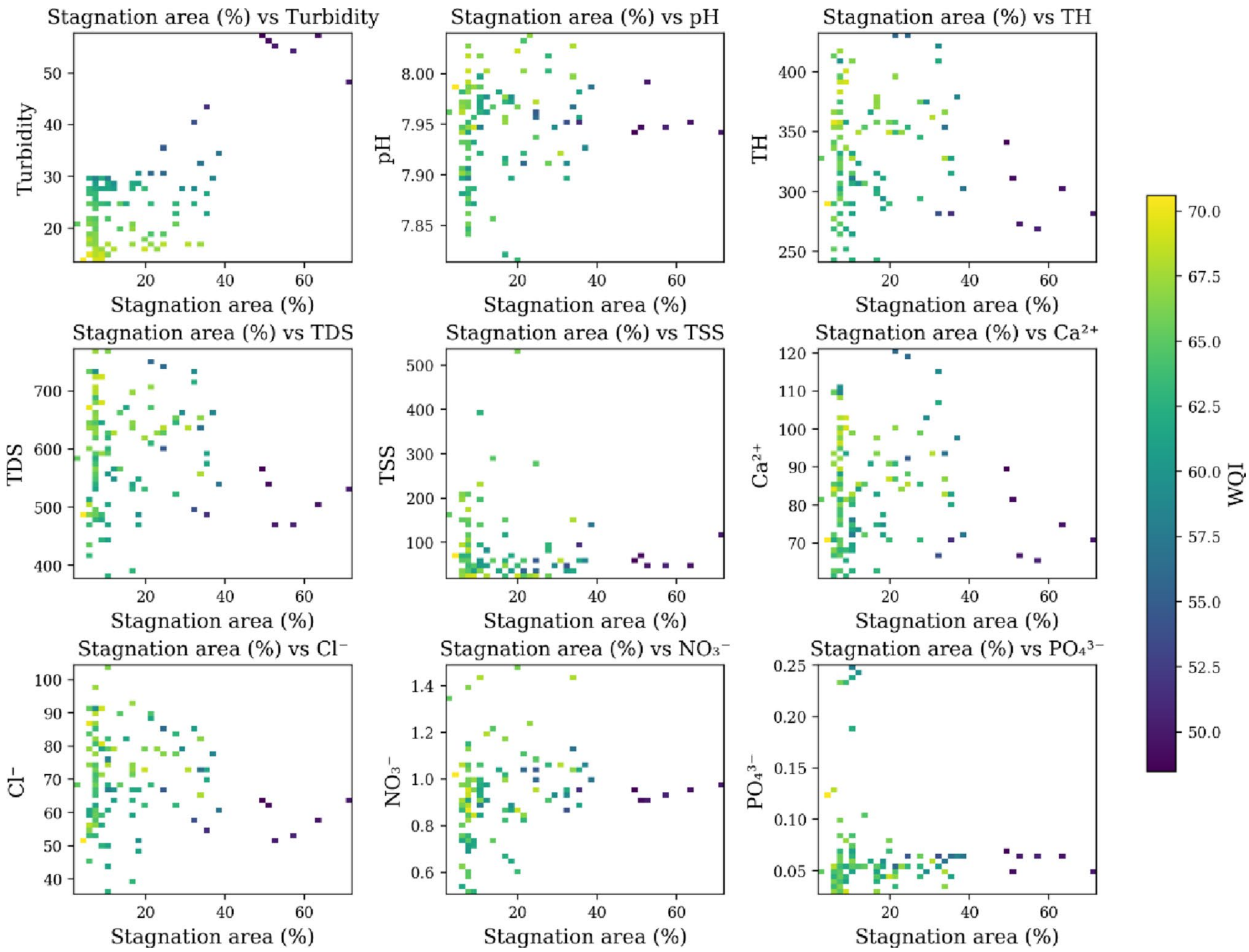


Fig. 10 Pairwise relationships between reach-scale stagnation extent and physico-chemical water-quality parameters (2014–2023), with point colour representing WQI

pH remained relatively buffered and independent of reach-scale hydraulics. Ionic variables—including TH, TDS, Ca^{2+}, and Cl^-—exhibited moderate scatter, reflecting a complex interplay of geochemistry and anthropogenic inputs rather than stagnation alone. Similarly, nutrient indicators (NO_3^- and PO_4^{3-}) showed limited dependence on the extent of stagnation, remaining primarily governed by upstream loading and point-source emissions. Comprehensive spatiotemporal datasets, including monthly spatial distributions and reach-scale aggregated data, are detailed in Supplementary Tables A and B.

Validation with previous work

To benchmark the hydrodynamic realism of the CFD configuration and address concerns associated with the 2D depth-averaged assumption, the simulated depth-averaged velocity response was compared with published hydrodynamic results for the Baghdad reach (Al-Asadi et al. 2022). As shown in Fig. 11, the present model reproduces the expected discharge–velocity scaling across representative flow rates (≈289–870 m^3/s) and consistently captures lower mean velocities in meander segments relative to straight reaches, reflecting planform-induced deceleration. This cross-study agreement supports that the adopted turbulence closure and roughness treatment yield a physically consistent reach-scale momentum balance, and it directly validates the key diagnostic variable used in this work to delineate stagnation hotspots (depth-averaged velocity and a threshold-based stagnation criterion). Although the 2D framework cannot explicitly resolve vertical shear and secondary circulation, the close agreement in the discharge–velocity trend across both planform classes indicates suitability for reach-scale stagnation susceptibility mapping under multi-yeardischarge variability, with an overall normalized RMSE of ~ 10% across the benchmark points.

The monthly comparison between CFD-derived Water Quality Index (WQI) values and station-observed data for

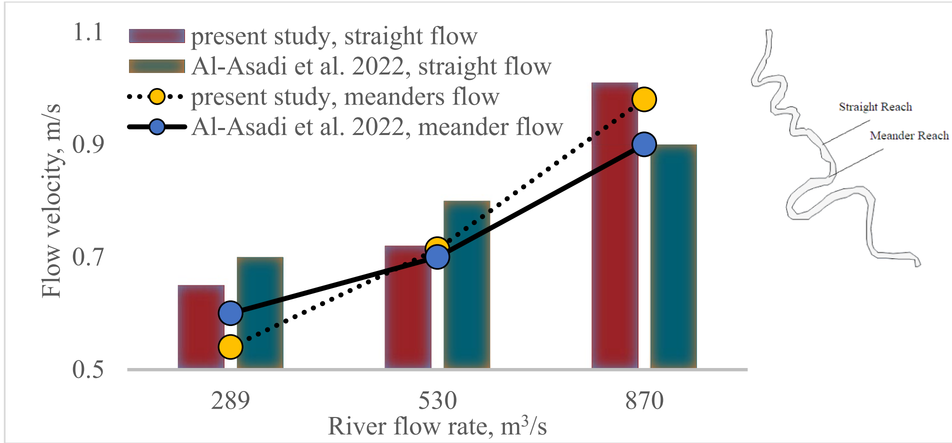


Fig. 11 Benchmarking of CFD-predicted depth-averaged velocities against published hydrodynamic results for the Baghdad Tigris reach

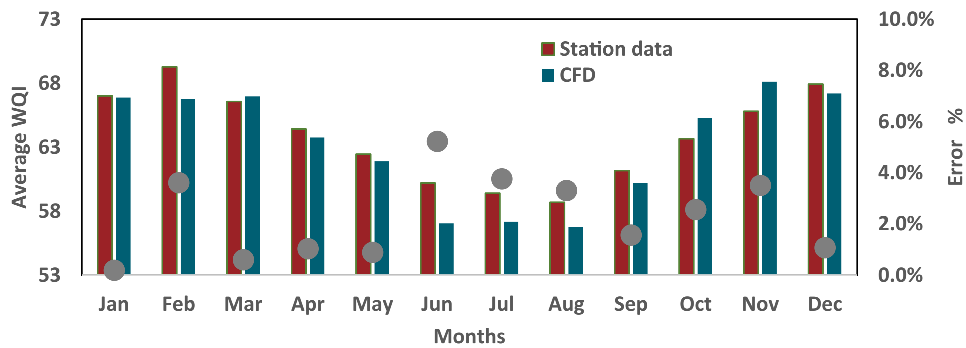


Fig. 12 Comparison between CFD and station data WQI with error % for 2023

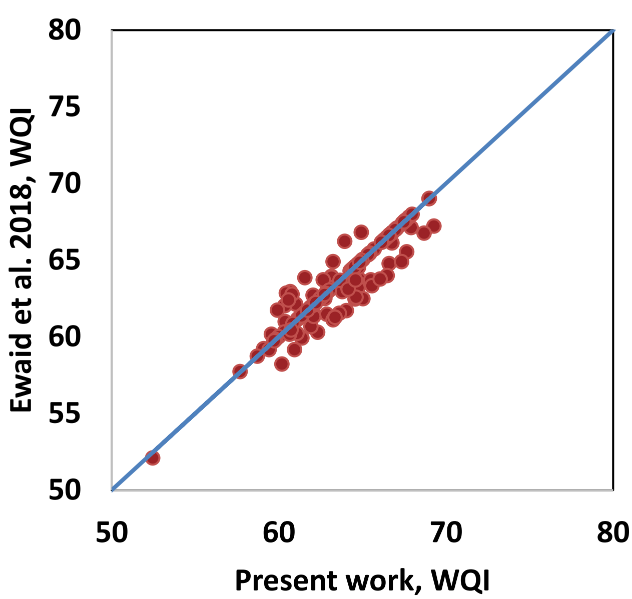


Fig. 13 Validation of present data with (Ewaid et al. 2018)

2023 is shown in Fig. 12, along with the associated error percentages. Both analytical approaches exhibited highly congruent seasonal trends, with CFD outputs aligning closely with empirical values. Error rates were consistently low throughout the year, typically ranging from 1% to 4%, with a single peak of approximately 6% in June. This high degree of convergence confirms the model's capacity to effectively capture the complex seasonal fluctuations inherent to the Tigris River system. Further validation of the current model was conducted against the benchmark dataset of Ewaid et al. (2018), a widely recognized standard for water quality assessments in Iraq (Fig. 13). Correlation analysis revealed a robust linear relationship (Pearson $r=0.99$; $R^2 = 0.98$), with data points clustering tightly around the 45° line of agreement. When the intercept was constrained to zero, a regression slope of 0.996 was obtained—virtually reaching unity (test vs. 1: $p=0.81$)—indicating negligible proportional deviation.

DT and k-NN models

The credibility of the proposed machine learning (ML) layer was evaluated by benchmarking Decision Tree (DT) and k-Nearest Neighbour (k-NN) against alternative learners, including Quadratic Discriminant Analysis (QDA), Naïve Bayes (NB), and Support Vector Machine (SVM). Under the leave-one-month-out (LOMO) framework for 2023 (Fig. 14a), both DT and k-NN achieved near-unity classification scores across all performance metrics. This high accuracy reflects the strong separability of flow regimes when physics-labeled CFD targets ($U<0.1$ m/s) are integrated with hydrodynamically meaningful predictors, such as reach geometry and velocity-distribution descriptors. The superiority of DT and k-NN was further confirmed through the notably weaker performance of QDA and NB, particularly in precision and F1-scores. This disparity suggests that distribution-based classifiers struggle to capture the non-linear decision boundaries inherent to reach-scale stagnation dynamics. While the SVM baseline demonstrated high general performance, DT and k-NN were prioritized as the primary models due to their transparent decision logic and operational interpretability—qualities essential for drought-time decision support in water treatment plant (WTP) management.

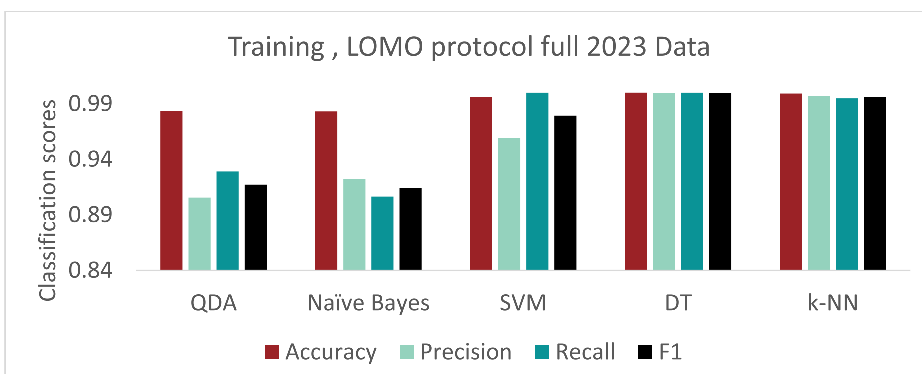
External validation during the May 2022 extreme drought (Fig. 14b) confirms that the DT and k-NN models maintain superior generalization capacity. Their robustness against overfitting is evidenced by stable performance despite significant hydrological regime shifts, with classification results anchored in the physical consistency of CFD-labeled stagnation definitions. Under 2023 LOMO testing (Fig. 15), stagnation area fluctuated between 6% and 13%, with analytical errors generally remaining below 2%. Critically, external validation on the 2022 dataset (Fig. 16) demonstrated strong agreement with CFD reference outputs, even as stagnation prevalence escalated to 20–72% due to severe drought. Analytical errors during this extreme regime remained predominantly under 2%, further validating the model's cross-year stability. Comprehensive performance metrics, including confusion matrices and secondary classification scores, are detailed in Tables D1–D2 and E1–E4 of the Supplementary Material.

The generalization capability of the ML classifiers from the 2023 training regime (6–13% stagnation) to the extreme conditions of 2022 (20–72% stagnation) is rooted in the node-level feature space. Rather than learning global seasonal fractions, the models were trained on 93,456 individual hydraulic instances (7788 nodes × 12 months). Even in low-stagnation months, the 2D CFD mesh provides thousands of 'stagnant' samples (velocity < 0.1 m/s) near boundaries, bars, and meanders. This ensures that the geometric and hydraulic 'fingerprint' of flow separation is well-represented in the training data. Consequently, when discharge dropped in 2022, the model did not face a 'new' physical phenomenon, but rather a spatial expansion of the same hydraulic susceptibility it had already mastered at the node level. The high accuracy across 2022 confirms that the classifiers learned the fundamental relationship between channel geometry and velocity thresholds, which remains robust across varying hydrological scales.

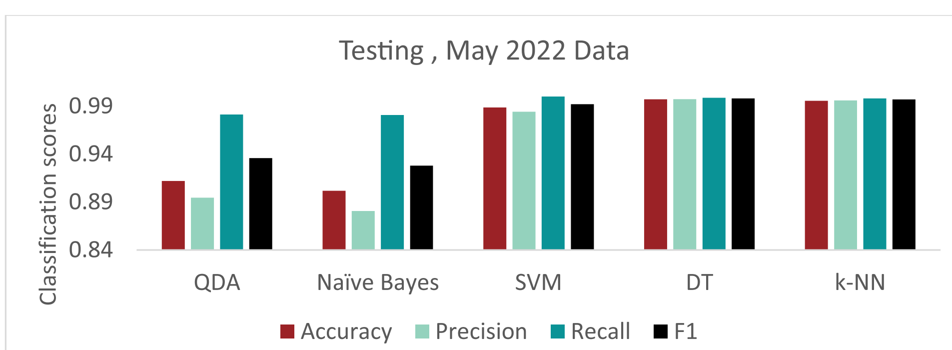
Flow–stagnation and WQI–stagnation correlations

Figure 17 highlights the strong non-linear influence of river flow on stagnation extent ($R^2 = 0.98$). When flows dropped below 200 m^3/s, the stagnant area expanded dramatically

Fig. 14 Comparative performance of ML classifiers for CFD-based stagnation detection: **a** Training performance under the LOMO protocol using full 2023 data; **b** External testing performance for May 2022 (extreme-drought regime)



a. Training classification performance.



b. Testing score performance

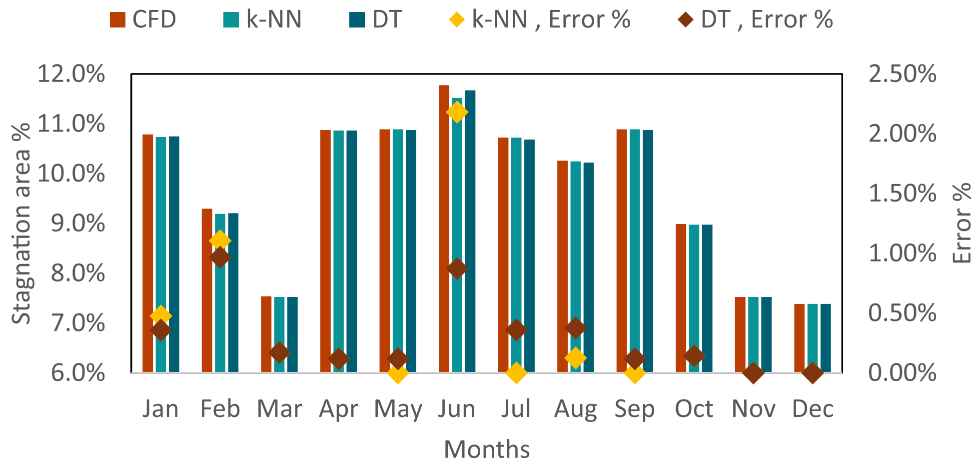


Fig. 15 Monthly stagnation area (%) in 2023 CFD, DT, and k-NN analysing under LOMO testing; right axis shows absolute error (%)

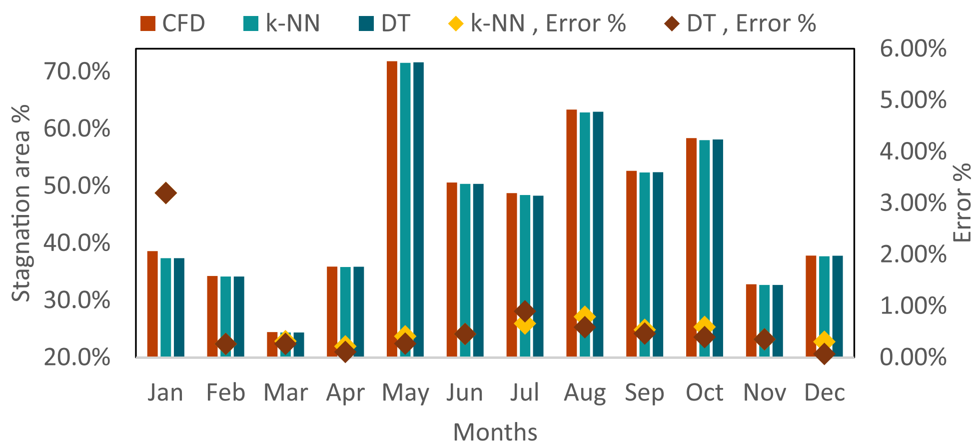


Fig. 16 External validation monthly stagnation area (%) with CFD, DT, and k-NN analysing (models trained on 2023: testing on 2022); right axis shows absolute error (%)

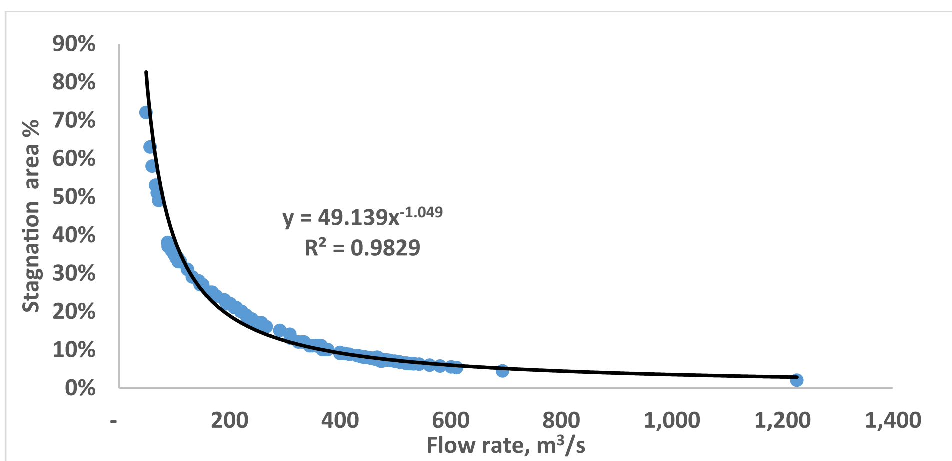
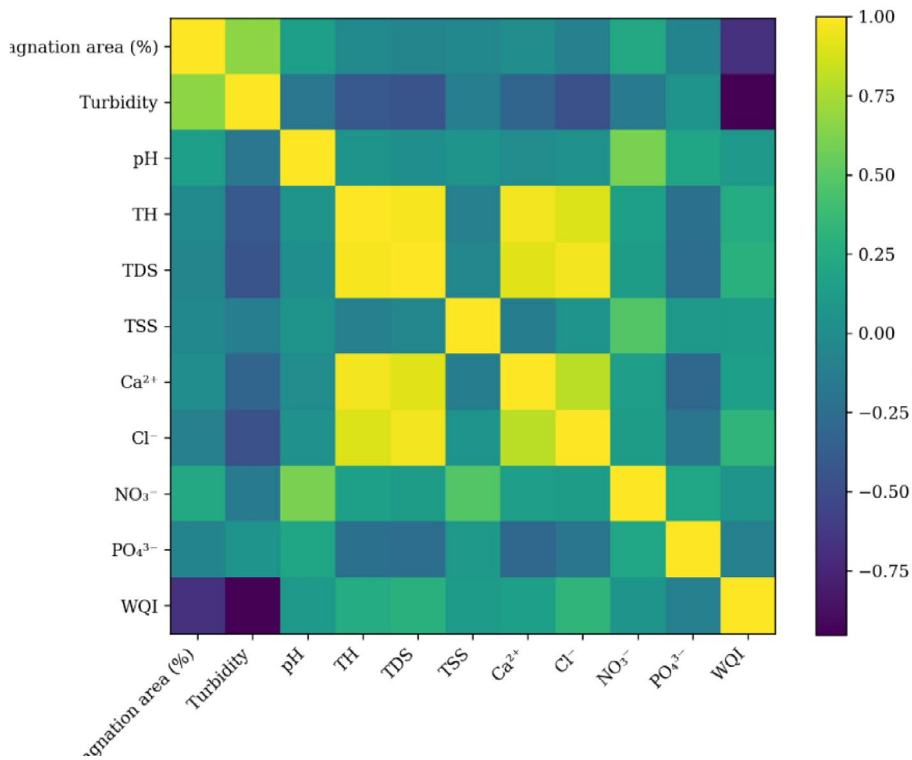


Fig. 17 Stagnation area dependence on river flow rate

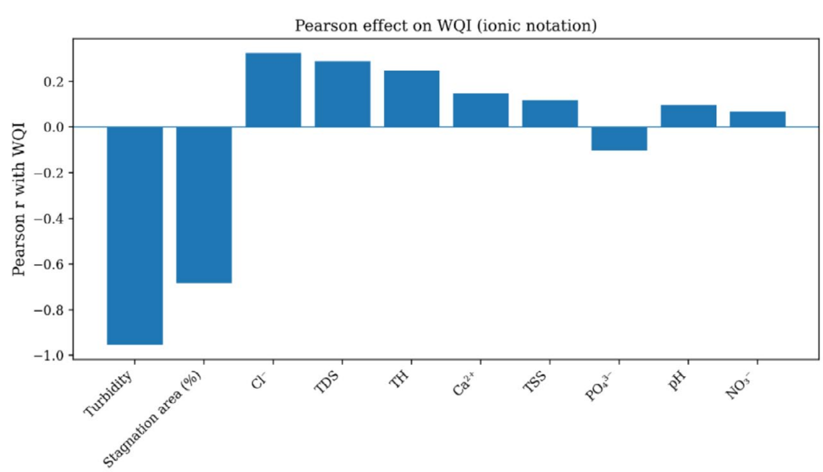
(exponentially), often exceeding 30–72%, while flows above 600 m³/s restricted stagnation to less than 15%. This finding yields a threshold that explains the critical deterioration observed in 2022 (Fig. 14), when drought-driven flow reductions coincided with a sharp decline in the Water Quality Index (WQI) to ~54. By contrast, flows above 600 m³/s were associated with reduced stagnation (< 15%) and higher WQI values near 66, confirming the close link between flow, stagnation, and water quality. Figure 17 data are presented in Table C in the supplementary materials.

The bivariate correlation structure, summarized in Fig. 18, indicates that the Water Quality Index (WQI) is most strongly associated with turbidity (representing the largest negative Pearson *r*). The stagnation area (%) also demonstrates a significant relationship with WQI and several co-varying water quality descriptors. Due to the intrinsic interdependence of physico-chemical parameters in river systems, predictor selection was governed by the Variance Inflation Factor (VIF) to ensure a robust multivariate model. Candidate main effects and interaction terms were systematically evaluated, and any term that induced severe multicollinearity (VIF > 100) or lacked statistical support ($p > 0.05$) was removed to prevent coefficient instability and inflated standard errors. Specifically, TH (x_4), TSS (x_6), Cl^- (x_8), and NO_3^- (x_9) were excluded from the final response-surface model. These variables exhibited excessive variance inflation (VIF ≫ 100) alongside non-significant p-values, indicating that their effects were redundant. Correlation analysis (Supplementary Tables F1–F2) confirmed that the excluded variables are highly correlated with the retained predictors—notably turbidity (x_2), TDS (x_5), stagnation area (x_1), and phosphate (x_{10}), suggesting that their influence on WQI is captured indirectly. The resulting Minimal model (Table 4) comprises stagnation area (x_1), turbidity (x_2), pH (x_3), TDS (x_5), Ca^{2+} (x_7), PO_4^{3-} (x_{10}), and a significant

Fig. 18 Correlative analysis of the effect of stagnation area % and physico-chemical properties on WQI



a. Full Pearson effect Matrix



b. Pearson effect on WQI.

stagnation × turbidity interaction ($x_1 \times x_2$). All retained terms demonstrated statistical significance ($p < 0.05$) and acceptable multicollinearity. This VIF-controlled pruning yielded a model that is both statistically stable and physically consistent, providing high explanatory power while minimizing parameter redundancy.

The predictor selection is corroborated by the correlation structures shown in Fig. 18, where the retained variables exhibit the most significant and interpretable relationships with the Water Quality Index (WQI). The alignment between Pearson patterns and the VIF-constrained ANOVA (Table 4) confirms that the fitted relationship is governed by a stable set of physically plausible drivers. WQI variability was found to be dominated by a minimal set of factors. Based on the predictive regression equation (Eq. 13), the primary drivers include stagnation area (x_1; $\beta_1 = -12.87$), pH (x_3; $\beta_3 = -11.44$), and turbidity (x_2; $\beta_2 = -0.7242$). These variables collectively account for most of the explained variance, reflecting the coupled effects of particulate loading and hydraulic retention during low-flow conditions. Notably, the positive interaction term between stagnation and turbidity ($x_1 \times x_2$; $\beta_{1,2} = 0.4576$) confirms that the impact of turbidity on WQI is significantly amplified under stagnant conditions. The exceptional predictive performance ($R^2 = 0.97$, Adjusted $R^2 = 0.97$, $S = 0.79$) confirms that these parameters represent the dominant mechanisms governing the water-quality response.

$$WQI = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_5 x_5 + \beta_7 x_7 + \beta_{10} x_{10} + \beta_{1,2} x_1 \times x_2 \quad (13)$$

Discussion

A critical quantitative link between hydrodynamic alterations, localized stagnation, and water quality degradation along the Tigris River's urban corridor is established by the findings of this study. Through the integration of high-resolution computational fluid dynamics (CFD) with machine learning and hydrochemical assessments, the physical mechanisms driving water-quality shifts are clarified, moving beyond a sole reliance on empiricalcorrelations.

Table 4 ANOVA table of selective model

Term	Coefficients (β)	T-Value	*P*-Value	VIF
Constant	180	230.00	0.000	
x_1	− 12.87	3.71	0.000	2.82
x_2	− 0.7242	− 37.80	0.000	3.35
x_3	− 11.44	− 6.33	0.000	1.41
x_5	− 0.00825	− 3.40	0.001	8.41
x_7	− 0.0327	− 2.27	0.025	7.22
x_{10}	− 7.8	− 4.72	0.000	1.20
$x_1 \times x_2$	0.4576	9.87	0.000	1.38

In the subsequent sections, spatiotemporal regime shifts in stagnation are contextualized, the robustness of the predictive models is validated, and the governing hydrochemical dynamics are unpacked. Finally, these insights are translated into actionable operational thresholds for adaptive water management.

Hydrological regime shifts and spatiotemporal stagnation

A distinct regime shift in stagnation dynamics was observed over the decadal study period. From 2014 to 2019, peak stagnation (up to 34%) occurred during winter, driven by upstream dam regulation and snowfall storage in the upper basin (Mason et al. 2023). Post-2020, however, the critical stagnation window shifted to late spring and early summer, culminating in a historical maximum of 72% in May 2022. This expansion reflects the compound effects of extended drought, elevated temperatures, and depressed discharge (<200 m^3/s). Conversely, the 2023 recovery year saw discharge restored to >200 m^3/s, reducing stagnant zones to ~15%. Spatially, the downstream stations of Al-Sadr, Al-Kadhumiya, and Al-Karama were identified as persistent hotspots. Their convoluted planform and shallow cross-sections exacerbate hydrodynamic stress during drought, aligning with observations in analogous regulated rivers where complex morphology and diminished flow create enduring stagnant zones (Abessi et al. 2012). These patterns were independently corroborated by Sentinel-2 imagery ("Sentinel-2 | Copernicus Data Space Ecosystem" updated) as seen in Fig. 19; the spatial correspondence between CFD-derived stagnant segments and remotely observed low-energy surface expressions (e.g., inner bends and sheltered zones) confirms that channel geometry rather than numerical artifacts fundamentally govern stagnation. The Sentinel-2 imagery utilized in this study provides high-resolution, multi-spectral optical data that captures the persistent surface expressions of low-energy hydraulics and stagnation-prone zones.

Spatially, the downstream stations of Al-Qadesiyah, and Al-Karama were identified as persistent hotspots. As summarized in Table 5 (derived from the benchmarking in Fig. 11), the influence of secondary currents in meandering segments is not negligible, with De values ranging from 4.9×10^4 to 1.2×10^5. In such "shallow-water" meanders, the hydrodynamic behavior is predominantly governed by bed-generated turbulence and frictional resistance (Manning's n) rather than 3D transverse mixing. Crucially, the depth-to-radius ratio (d/R) at the sharpest bends is approximately 0.006. Such a low ratio indicates that the river is hydraulically "shallow," where vertical and transverse momentum exchange is secondary to longitudinal momentum and bed-generated turbulence. Consequently,

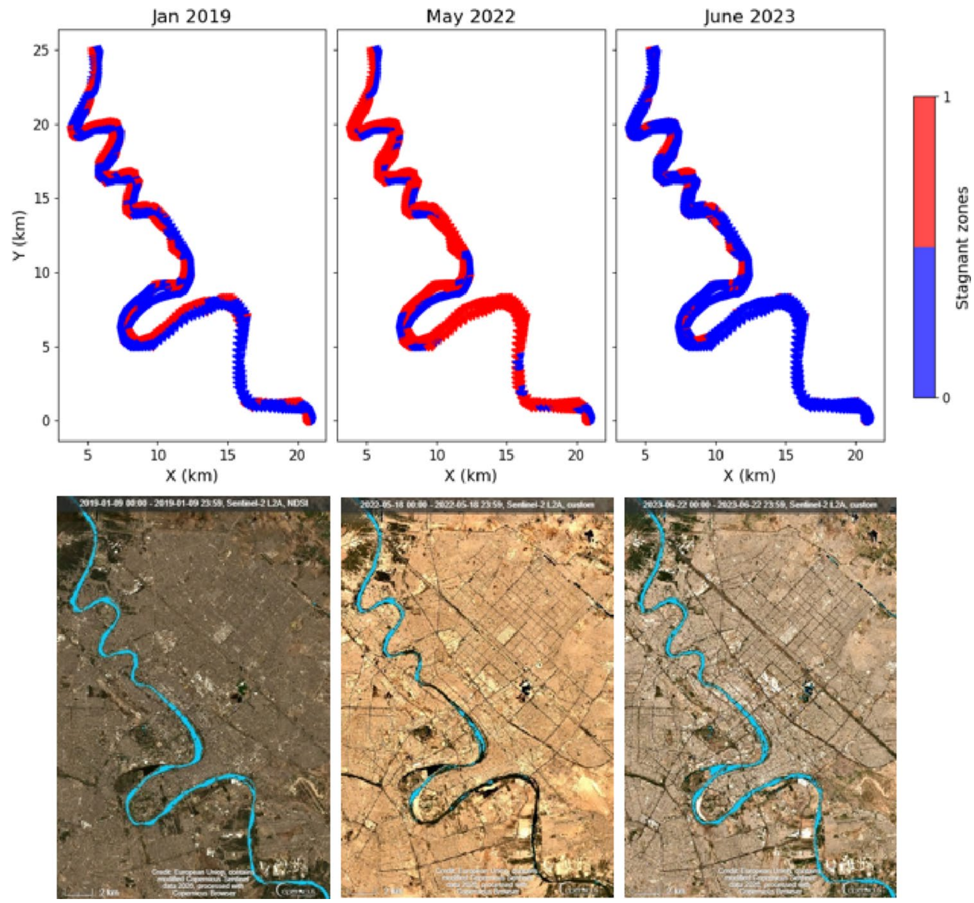


Fig. 19 Spatial consistency between modelled stagnation areas and Sentinel-2 imagery

Table 5 Calculated Dean numbers (De) across representative discharge levels for straight and meandering reaches ($d/R \approx 0.006$)

Q (m3/s)	Velocity v (m/s)	De (Meandering reach)
289 (Low flow)	0.18	$4.9 \times 10_4$
400	0.25	$6.8 \times 10_4$
870 (High flow)	0.45	$1.2 \times 10_5$

Table 6 The stagnation area % sensitivity of various manning factors (n) and flow rates

Manning factor (n)	Drought, Q = 49 m^3/s	Moderate, Q = 475 m^3/s	High flow, Q = 1228 m^3/s
0	63%	8%	2%
0.01	72%	7%	2%
0.032	74%	4%	2%
0.04	73%	4%	1%

the likely bias of the 2D formulation is a slight overestimation of stagnation (conservative bias) due to the absence of transverse mixing. However, this uncertainty is bounded by the Manning sensitivity already reported in Table 6, which demonstrates that frictional resistance exerts a more dominant control over the stagnant fraction than 3D secondary circulation in this specific reach. This is confirmed by the Manning sensitivity analysis (Table 6) and the high spatial correspondence with Sentinel-2 imagery (Fig. 19), where remotely observed low-energy zones align perfectly with CFD-derived stagnation hotspots. This confirms that channel geometry, captured robustly by the depth-averaged k–omega framework, fundamentally governs the reach-scale stagnation.

Machine learning robustness and generalization

The reliability of the CFD–ML integration was demonstrated through a "train-once, test-across-years" protocol. Models trained on 2023 data were applied to 2019 (moderate flow) and 2022 (Fig. 20, extreme drought) without retraining, achieving high fidelity with errors $\leq$2%. This cross-year stability—maintained despite significant hydrological regime shifts—confirms that the model captures physically separable flow characteristics rather than overfitted seasonal signatures. The exceptionally high DT and k-NN scores do not indicate overfitting, trivial classification, or circularity, as the model's robustness was demonstrated across multiple leakage-controlled and regime-shift conditions. First, performance was evaluated using leave-one-month-out (LOMO), where each model was trained on 11 months and tested on an entirely unseen month, with

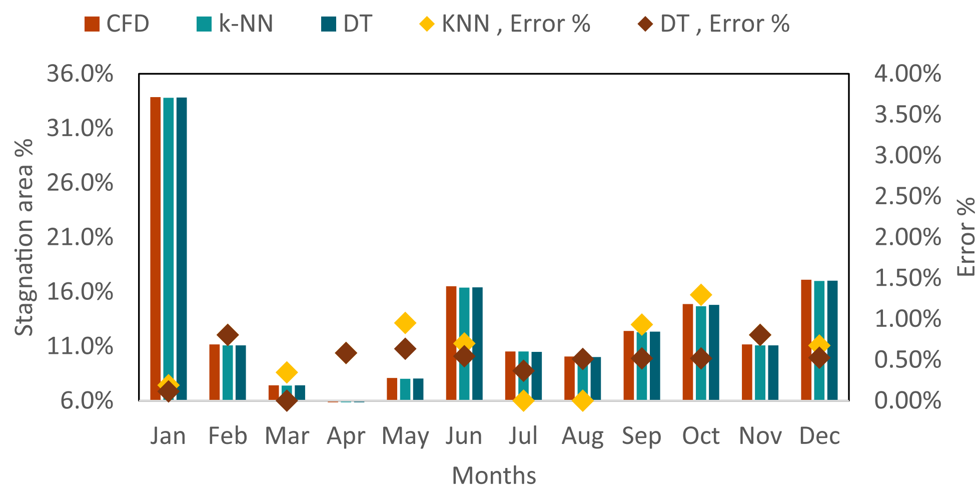


Fig. 20 External validation monthly stagnation area (%) with CFD, DT, and k-NN analysing (models trained on 2023; testing on 2019); right axis shows absolute error (%)

standardization and hyperparameter tuning confined to the training folds to prevent information leakage. Second, the models trained in 2023 were applied unchanged to 2019 and the extreme-drought year 2022, maintaining consistently high precision/recall/F1 despite significant changes in stagnation prevalence, which would not occur under an overfit or majority-class strategy. Third, complete confusion matrices confirm low FP/FN rates, proving that accuracy is not driven by class imbalance. Finally, benchmarking against QDA, Naïve Bayes, and SVM on the same split shows lower performance for the baselines, supporting the idea that DT/k-NNsuccess reflects a physically separable, physics-labelled stagnation problem, not circular learning or trivial rule matching. This protocol is similar to the previous one, which focused on event-based flood depth transfer (e.g., Pakdehi et al. 2024). This study targets the extent of stagnation and examines cross-year performance in a large, natural arid-region river.

Hydrochemical dynamics and process-based predictor selection

The longitudinal decline in WQI from ~65 upstream to ~60 downstream is consistent with the cumulative pollutant loading observed within the Baghdad urban corridor. This spatial gradient is characterized by increases in TDS, turbidity, and total hardness (TH), reflecting declining water quality as the river traverses the metropolitan reach. Seasonal variability further highlights the influence of hydrological conditions; higher WQI values in winter and early spring coincide with elevated flows that maximize dilution capacity, whereas summer declines (51–55) align with flow contraction and the subsequent accumulation of pollutants. These patterns correspond with findings by Varol et al. (2012), who noted that eutrophication risks in the Tigris are highest during summer low-flow periods—characterized by temperatures exceeding 25 °C and PO_4^{3-} levels surpassing 0.10 mg/L. In the Tigris River, statistical collinearity among physicochemical variables is rooted in shared hydrogeochemical and hydraulic controls. The strong covariation between TH, Ca^{2+}, and TDS is a product of the basin's carbonate-dominated mineralogy. Hardness is primarily governed by the dissolution of limestone and dolomite formations, during which calcium ions are released in stoichiometric proportion to total mineralization (Shakir et al. 2017). Similarly, the close correspondence between Cl^- and TDS reflects chloride's conservative behavior as a tracer for both natural mineralization and anthropogenic wastewater inputs, particularly under low-flow conditions. Furthermore, TSS and turbidity are tightly coupled, as both respond to the sediment entrainment and resuspension processes that intensify during discharge fluctuations (Kadhem 2013).

Consequently, the elevated Variance Inflation Factors (VIF > 100) reported for these variables are interpreted as indicators of genuine process redundancy rather than statistical artifacts. Their exclusion from the final multivariate model was necessary to avoid multicollinearity-driven instability while ensuring the model coefficients remain identifiable and physically interpretable. The resulting model structure retains representative predictors—such as TDS, Ca^{2+}, and turbidity—to capture the dominant signals of mineralization, carbonate equilibrium, and particulate dynamics, respectively. This selection is entirely consistent with established hydrochemical characterizations of the Tigris (Kadhem 2013)and represents a physically meaningful aggregation of correlated processes.

It is important to acknowledge a limitation in the current WQI framework: the exclusion of **Dissolved Oxygen (DO)**. While DO is the parameter most directly linked to the biological consequences of flow separation—such as anoxia and increased sediment oxygen demand—it was excluded from this study due to inconsistent reporting in the historical monitoring records. Consequently, the calculated WQI primarily reflects physical and mineral loading (e.g., turbidity, Total hardness and TDS) rather than immediate aerobic stress dynamics. While this limits the direct interpretation of oxygen-related biological impacts, the strong spatial correlation between stagnation and high-weight parameters like turbidity and phosphates indicates that the model still effectively identifies zones at high risk of environmental and chemical degradation.

The application of a 2D depth-averaged model is robust for long-term reach-scale simulations; however, it is essential to discuss the implications of vertical flow structures in stagnant zones. In the low-flow regimes of the Tigris, particularly during summer months, the reduction in shear-induced mixing can trigger localized thermal stratification and vertical velocity gradients. As highlighted by Higashino (2011), stagnant water bodies in arid environments act as settling basins where the downward flux of particulate-bound nutrients (such as phosphates) exceeds vertical transport. This settling process intensifies Sediment Oxygen Demand (SOD) and creates a decoupling between surface water quality and near-bed concentrations. While present 2D model identifies these zones based on lateral velocity thresholds, the biological degradation within these 'hotspots' is likely exacerbated by these 3D vertical dynamics, which favour anoxic conditions and internal nutrient loading from the riverbed. Crucially, the 2D identification of stagnation hotspots serves as the spatial precursor to these complex 3D processes. Stagnation is the primary hydraulic driver; once the horizontal velocity threshold (0.1 m/s) is breached, the environment transitions into a state where vertical stratification and nutrient settling become physically inevitable. Therefore, current depth-averaged framework maintains a robust mechanistic link with water quality dynamics by accurately mapping the spatial extent where these three-dimensional degradation risks are concentrated, ensuring that the hydrodynamic-water quality nexus remains valid despite the vertical simplification.

Operational thresholds for water management

The identification of a safe-flow threshold of ≈200 m^3/s is critical for mitigating raw-water risks in urban arid reaches. Below this limit, stagnation expands non-linearly to 30–72% of the reach, precipitously diminishing hydraulic connectivity and self-cleansing capacity. Given that Baghdad's total extraction is approximately 47.5 m^3/s (Iraqi Ministry of planning 2024), withdrawals during low-flow periods can materially exacerbate these stagnant regimes. Water-quality degradation is a deterministic process concentrated within these hydraulically disconnected states, supported by robust correlations between discharge and stagnation extent (R^2 = 0.98) and between discharge and the Water Quality Index (R^2 = 0.97). Practically, stagnation serves as a robust hydraulic vulnerability indicator for adaptive management. Utilizing discharge as a real-time early-warning trigger allows operators to redistribute withdrawals away from high-vulnerability corridors—such as Al-Kadhimiya, Al-Doura, and Al-Wahda—toward high-connectivity segments, such as East Tigris and Al-Wathba. For plants within persistent hotspots, intensified protocols, including optimized chemical dosing and enhanced filtration, are essential to manage the elevated pollutant loads (e.g., PO_4^{3-} and turbidity) inherent to low-energy hydraulic states.

Conclusions

This study identifies hydrodynamic stagnation as a robust reach-scale vulnerability indicator for water-quality deterioration in the Baghdad reach of the Tigris River. Numerical simulations reveal that stagnation extent is primarily discharge-controlled, exhibiting a non-linear expansion below a critical threshold of ≈200 m^3/s. During the extreme drought of 2022, hydraulic connectivity loss reached ~49–72% of the reach, whereas discharges exceeding 600 m^3/s effectively suppressed stagnant zones. This flow-driven transition is mirrored in the Water Quality Index (WQI), where degraded conditions are non-uniformly concentrated within high-stagnation regimes. The framework's analytical robustness is validated by aligning depth-averaged velocities with discharge–velocity scaling (NRMSE ≈ 10%) and by confirming that WQI statistics are equivalent to regional benchmarks within ±2 units. Furthermore, the interpretable machine learning layers (DT and k-NN) demonstrated strong generalization capacity through "train-once, test-across-years" performance, successfully mapping unseen hydrological regimes without retraining. Multivariate analysis confirms that WQI variability is governed by a Minimal set of drivers—primarily turbidity, pH, and stagnation extent—with a significant stagnation × turbidity interaction revealing that particulate loading impacts are amplified under stagnant conditions. Operationally, these findings support a discharge-based early-warning system for water treatment plants (WTPs), enabling adaptive abstraction and proactive treatment adjustments during low-flow periods. While limitations in depth-averaging remain, this research provides a scalable blueprint for enhancing water-quality resilience in regulated arid-region rivers. Future research should prioritize integrating real-time monitoring and plant-specific datasets to extend these findings to other vulnerable aquatic systems formally. To adapt this framework as a scalable blueprint for other regulated arid rivers with differentmorphometric characteristics, three primary requirements must be met: (1) high-resolution bathymetric surveys to accurately represent site-specific channel geometry in the CFD environment, (2) consistent historical records of flow rates and water quality parameters to calibrate the WQI reference targets, and (3) re-training of the ML classifiers on the specific hydraulic feature space of the new river to account for unique curvature and roughness variations.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s13201-026-02883-1.

Acknowledgements The research was carried out within the framework of the National Laboratory for Water Science and Water Security RRF-2.3.1-21-2022-00008. The authors acknowledge the data provided by the Baghdad Municipality. Viktor Sebestyén's contribution to this paper was supported by the János Bolyai Research Scholarship of the Hungarian Academy of Sciences (BO/00889/24).

Author contributions Mustafa Hathal: Conceptualization, CFD modelling, machine learning implementation, formal analysis, data curation, and original draft preparation. Viktor Sebestyén: Supervision, machine learning guidance, writing – review and editing. Stanisław Anweiler: Statistical analysis, validation (ANOVA and regression), and language editing. Viola Somogyi: Supervision, conceptual support, project administration, financial support, and writing – review and editing.

Funding Open access funding provided by University of Pannonia. This research was supported by the National Laboratory for Water Science and Water Security (Project No. RRF-2.3.1-21-2022-00008). Viktor Sebestyén's contribution was supported by the János Bolyai Research Scholarship of the Hungarian Academy of Sciences (BO/00889/24).

Data availability The datasets generated during and/or analyzed during the current study are available from the corresponding author on reasonable request. Some data derived from public sources (NASA POWER, Sentinel-2, DAHITI) are available through their respective web portals.

Declarations

Conflict of interest The author declares that there are no known financial or personal relationships that could have appeared to influence the work reported in this paper.

AI use During the preparation of this work, the authors used AI-based tools for language polishing and stylistic refinement of the manuscript. After using these tools, the authors reviewed and edited the content as needed and take full responsibility for the content of the publication.

References

Abessi O, Saeedi M, Bleninger T, Davidson M (2012) Surface discharge of negatively buoyant effluent in unstratified stagnant water. J Hydro-environment Res 6:181–193. https://doi.org/10.1016/J.JHER.2012.05.004

Ahmed W, Mohammed S, El-Shazly A, Morsy S (2023) Tigris River water surface quality monitoring using remote sensing data and GIS techniques. Egypt J Remote Sens Sp Sci 26:816–825. https://doi.org/10.1016/j.ejrs.2023.09.001

Al-Ansari N, Adamo N, Sissakian VK (2019) Water shortages and its environmental consequences within Tigris and Euphrates rivers. J Earth Sci Geotech Eng 9:1792–9660

Al-Asadi S, Hussain TA, Al-Mukhtar M (2022) Modeling flow patterns in Tigris River-Baghdad using the CCHE2D model. IOP Conf Ser Earth Environ Sci https://doi.org/10.1088/1755-1315/1120/1/012012

Al-Farraji A, Hathal MM (2022) The numerical simulation and thermal analysis of nanoparticles blends within a stirred batch reactor for biodiesel production from waste cooking oil. AIP Conf Proc 2660:020139. https://doi.org/10.1063/5.0107690

Al-Mansori NJ, Al-Kizwini RS, Al-Husseini FK (2020) Modeling and simulation of pollutants dispersion in natural rivers using comsol multiphysics. J Eng Sci Technol 15:1167–1185

Alfwzan WF, Selim MM, Almalki AS, Alharbi IS (2024) Water quality assessment using Bi-LSTM and computational fluid dynamics (CFD) techniques. Alexandria Eng J 97:346–359. https://doi.org/10.1016/j.aej.2024.04.030

APHA (Americans Public Health Association) (1998) Standard Methods for Examination of Water and Wastewater, 20th Editi. Washington DC

APHA (2022) Standard Methods for the Examination of Water and Wastewater. 24th Edition, American Public Health Association

Asaad BI, Abed B (2020) Flow characteristics of Tigris River within Baghdad City during drought. J Eng 26:77–92

Asadollah SBHS, Sharafati A, Motta D, Yaseen ZM (2021) River water quality index prediction and uncertainty analysis: a comparative study of machine learning models. J Environ Chem Eng 9:104599. https://doi.org/10.1016/j.jece.2020.104599

Bao J, Chen Y, Fang Y et al (2022) Modeling framework for evaluating the impacts of hydrodynamic pressure on hydrologic exchange fluxes and residence time for a large-scale river section over a long-term period. Environ Model Softw 148:105277. https://doi.org/10.1016/j.envsoft.2021.105277

Chen S, Huang J, Wang P et al (2024) A coupled model to improve river water quality prediction towards addressing non-stationarity and data limitation. Water Res https://doi.org/10.1016/j.watres.2023.120895

Coupling Computational Fluid Dynamics with Decision Tree and k-Nearest Neighbours Classifications for Stagnation and Water Quality Assessment in Arid Rivers: Evidence from the Tigris

Dataset:München DAHITI (2025) Deutsches Geodätisches Forschungsinstitut der Technischen Universität. https://dahiti.dgfi.tum.de/en/ NASA POWER | Data Access Viewer (DAV). https://power.larc.nasa.gov/data-access-viewer/. Accessed 16 Aug 2025

Ewaid SH, Abed SA, Kadhum SA (2018) Predicting the Tigris River water quality within Baghdad, Iraq by using water quality index and regression analysis. Environ Technol Innov 11:390–398. https://doi.org/10.1016/J.ETI.2018.06.013

Ewaid SH, Abed SA, Al-Ansari N, Salih RM (2020) Development and evaluation of a water quality index for the Iraqi rivers. Hydrology 7:1–14. https://doi.org/10.3390/HYDROLOGY7030067

Gualtieri C (2008) Numerical Simulation of Flow Patterns and Mass Exchange Processes in Dead Zones. In: 4th International Congress on Environmental Modelling and Software - Barcelona, Catalonia, Spain - July. https://scholarsarchive.byu.edu/cgi/viewcontent.cgi?article=2895&context=iemssconference

Haghiabi AH, Nasrolahi AH, Parsaie A (2018) Water quality prediction using machine learning methods. Water Qual Res J 53:3–13. https://doi.org/10.2166/wqrj.2018.025

Herrera-Granados O (2009) Computational Fluid Dynamics (CFD) in river engineering: a general overview. In: Conference: II International Interdisciplinary Technical Conference of Young Scientists, InterTech. Poznań, Poland

Higashino M (2011) Oxygen consumption by a sediment bed for stagnant water: comparison to SOD with fluid flow. Water Res 45:4381–4389. https://doi.org/10.1016/J.WATRES.2011.04.051

Huang C, She D, Liu X et al (2025) Predicting future meteorological drought risk in mainland China using random forest model. J Hydrol Reg Stud 61:102633. https://doi.org/10.1016/J.EJRH.2025.102633

Hussein HA, Alshami AH, Tuama Al-Awadi A, Ibrahim MA (2023) Hydrological characteristics of the Tigris River at the Baghdad Sarai station. Ain Shams Eng J 14:101846. https://doi.org/10.1016/j.asej.2022.101846

Iraqi Ministry of planning C (2024) Environmental Statistics. Iraq

Ismail AH, Muntasir AH (2018) Estimation of river Tigris dispersivities using a steady-state numerical model. Appl Water Sci https://doi.org/10.1007/s13201-018-0749-z

Ismail A, Shareef M, Hussein H et al (2022) Two-dimensional simulation of pollutant dispersion at the confluence of Diyala river with Tigris River using CFD technique. J Water Resour Geosci 1

Kadhem AJ (2013) Assessment of water quality in Tigris River-Iraq by using GIS mapping. Nat Resour 04:441–448. https://doi.org/10.4236/nr.2013.46054

Kibaroglu A, Gürsoy SI (2015) Water-energy-food nexus in a transboundary context: the Euphrates-Tigris River basin as a case study. Water Int 40:824–838. https://doi.org/10.1080/02508060.2015.1078577

Li H, Shatarah M (2024) Operator learning for urban water clarification hydrodynamics and particulate matter transport with physics-informed neural networks. Water Res 251:121123. https://doi.org/10.1016/j.watres.2024.121123

Liu Y, Zhu Y, Li D et al (2023) Computational simulation of mass transfer in membranes using hybrid machine learning models and computational fluid dynamics. Case Stud Therm Eng 47:103086. https://doi.org/10.1016/j.csite.2023.103086

Ma C, Gu Y, Cheng J et al (2026) Analysis of dean vortices and sediment distribution in inlet passage of nearshore pump stations under solid-liquid two-phase conditions. Appl Ocean Res 170:105043. https://doi.org/10.1016/j.apor.2026.105043

Mahdi BA, Moyel MS, Jaafar RS (2021) Adopting the water quality index to assess the validity of groundwater in Al-Zubair city southern Iraq for drinking and human consumption. Ecol Environ Conserv 27:73–79

Majeed OS (2024) Evaluation of Tigris River water quality within Baghdad City using the CCME index for various purposes. Int J Aquat Biol 12:360–373. https://doi.org/10.22034/ijab.v12i4.2274

Mannina G, Viviani G (2010) River water quality assessment: a hydrodynamic water quality model for propagation of pollutants. In: 7ème Conférence internationale sur les techniques et stratégies durables pour la gestion des eaux urbaines par temps de pluie/7th International Conference on sustainable techniques and strategies for urban water management. Lyon, France

Mason M, Akıncı S, Bilgen A et al (2023) owards Hydro-Transparency on the Euphrates-Tigris Basin: Mapping Surface Water Changes in Iraq, 1984–2015 (PeaceRep Policy Brief). London, UK

Mensoor MK, Said A, Mansoor ZK (2022) Monitoring Pollution of the Tigris River in Baghdad by Studying Physico-chemical Characteristics. J Agric Hortic Res https://doi.org/10.33140/jahr.05.02.08

Murib M, Al (2018) Hydrodynamic and water quality modeling of the Tigris River system in Iraq using CE-QUAL-W2. Portland State University

Najah Ahmed A, Binti Othman F, Abdulmohsin Afan H et al (2019) Machine learning methods for better water quality prediction. J Hydrol https://doi.org/10.1016/j.jhydrol.2019.124084

Nasir N, Kansal A, Alshaltone O et al (2022) Water quality classification using machine learning algorithms. J Water Process Eng 48:102920. https://doi.org/10.1016/j.jwpe.2022.102920

Niazkar M, Piraei R (2025) Enhancing estimation of water quality index using stacking machine learning techniques: The case of Southern Bug River. Sci Total Environ 1003:180744. https://doi.org/10.1016/J.SCITOTENV.2025.180744

Pakdehi M, Ahmadisharaf E, Nazari B, Cho E (2024) Transferability of machine-learning-based modeling frameworks across flood events for hindcasting maximum river water depths in coastal watersheds. Nat Hazards Earth Syst Sci 24:3537–3559. https://doi.org/10.5194/NHESS-24-3537-2024

Pu JH (2015) Turbulence modelling of shallow water flows using Kolmogorov approach. Comput Fluids 115:66–74. https://doi.org/10.1016/j.compfluid.2015.03.010

Qi Q, Wang Z, Yin H et al (2024) Construction and application of a seasonal river health evaluation system in arid and semi-arid areas. Water (Switzerland) 16:1–13. https://doi.org/10.3390/w16050691

Rahman IMM, Islam MM, Hossain MM et al (2011) Stagnant surface water bodies (SSWBs) as an alternative water resource for the Chittagong metropolitan area of Bangladesh: physicochemical characterization in terms of water quality indices. Environ Monit Assess 173:669–684. https://doi.org/10.1007/S10661-010-1414-7/METRICS

Ratliff HN (2019) Cartesian formulas for curvature, circumradius, and circumcenter for any three two-dimensional points. Zenodo 1:2–6

Raviteja N, Saiteja N, Sreenu N et al (2023) Water quality classification using machine learning. Tuijin Jishu/Journal Propuls Technol 44:5862–5873. https://doi.org/10.1109/ICRAIE59459.2023.10468171

Seifi A, Riahi-Madvar H (2019) Improving one-dimensional pollution dispersion modeling in rivers using ANFIS and ANN-based GA optimized models. Environ Sci Pollut Res 26:867–885. https://doi.org/10.1007/s11356-018-3613-7

Sentinel-2 | Copernicus Data Space Ecosystem (2026b) https://dataspace.copernicus.eu/data-collections/copernicus-sentinel-data/sentinel-2. Accessed 22 Jan

Seo J, Won J, Lee H, Kim S (2024) Probabilistic monitoring of meteorological drought impacts on water quality of major rivers in South Korea using copula models. Water Res 251:121175. https://doi.org/10.1016/J.WATRES.2024.121175

Shakir E, Zahraw Z, Al-obaidy AHMJ (2017) Environmental and health risks associated with reuse of wastewater for irrigation. Egypt J Pet 26:95–102. https://doi.org/10.1016/j.ejpe.2016.01.003

Shams MY, Elshewey AM, El-kenawy ESM et al (2024) Water quality prediction using machine learning models based on grid search method. Multimed Tools Appl 83:35307–35334. https://doi.org/10.1007/s11042-023-16737-4

Shen Y, Cheng W (2022) A Tree-Based Machine Learning Method for Pipeline Leakage Detection. Water (Switzerland) https://doi.org/10.3390/w14182833

Sidek LM, Mohiyaden HA, Marufuzzaman M et al (2024) Developing an ensembled machine learning model for predicting water quality index in Johor River Basin. Environ Sci Eur https://doi.org/10.1186/s12302-024-00897-7
Torabi Haghighi A, Akbari M, Noori R et al (2023) The impact of Turkey's water resources development on the flow regime of the Tigris River in Iraq. J Hydrol Reg Stud 48:101454. https://doi.org/10.1016/J.EJRH.2023.101454
Tyagi M, Singh H, Thakur DA, Mohanty MP (2024) Compound risks of floods and droughts over multi-hazard catchments: revealing association through hydrodynamic-cum-statistical modelling and novel bivariate risk classifier. Sci Total Environ 957:177689. https://doi.org/10.1016/J.SCITOTENV.2024.177689
Ubah JI, O LC, O NM, K.N. O (2021) River sediment transport modeling of ele river using computational fluid dynamics model of comsol multiphysics program. ASM Sci J 16:1–7. https://doi.org/10.32802/asmscj.2021.562
Uddin MG, Nash S, Olbert AI (2021) A review of water quality index models and their use for assessing surface water quality. Ecol Indic 122:107218. https://doi.org/10.1016/j.ecolind.2020.107218
Varol M, Gökot B, Bekleyen A, Şen B (2012) Water quality assessment and apportionment of pollution sources of Tigris River (Turkey) using multivariate statistical techniques-a case study. River Res Appl 28:1428–1438. https://doi.org/10.1002/rra.1533
Wang X, Yang J, Wang F et al (2023) Numerical Modeling of the Dispersion Characteristics of Pollutants in the Confluence Area of an Asymmetrical River. Water (Switzerland) https://doi.org/10.3390/w15213766
WHO (2022) Guidelines for drinking-water quality: fourth edition incorporating the first and second addenda
Zhang L, Wu Y, Li J et al (2024) Hydrodynamics and dissolved organic matter components shaped the fate of dissolved heavy metals in an intensely anthropogenically disturbed estuary. Sci Total Environ 934:173293. https://doi.org/10.1016/J.SCITOTENV.2024.173293
Zheng Y, Zhang Q, Zhang X et al (2025) A spatial-temporal trend-aware neural network model for accurate water quality prediction in river. Water Res 287:124389. https://doi.org/10.1016/J.WATRES.2025.124389