Italian Journal of Engineering Geology and Environment

www.ijege.uniroma1.it

EMPIRICAL MODELING OF SALINITY INTRUSION LENGTH IN ALLUVIAL ESTUARIES AT HIGH WATER SLACK: A REGRESSION-BASED APPROACH USING GLOBAL DATASETS

Ali NAJJARI(*), Nazila KARDAN(*), Javad PARSA(**) & Ebrahim GHADERPOUR(***)

(*)Azarbaijan Shahid Madani University - Department of Civil Engineering - Tabriz (Iran)
(**)University of Tabriz - Department of Water Engineering - Tabriz (Iran)
(***)Sapienza University of Rome - Department of Earth Sciences - Rome (Italy)
Corresponding author: ebrahim.ghaderpour@uniroma1.it

EXTENDED ABSTRACT

L'intrusione di salinità negli estuari alluvionali pone crescenti minacce alla qualità dell'acqua, agli ecosistemi e ai mezzi di sussistenza costieri a causa dei cambiamenti climatici e delle crescenti pressioni umane. Comprendere e prevedere l'entità dell'intrusione di salinità è essenziale per la gestione sostenibile dell'acqua e la protezione dell'ecosistema in questi ambienti dinamici. Questo studio sviluppa un robusto modello empirico per prevedere la lunghezza dell'intrusione di salinità in condizioni di high water slack (HWS), la fase critica corrispondente alla massima penetrazione a monte dell'acqua salina. Un set di dati completo di 42 osservazioni provenienti da 22 estuari globali, tra cui i sistemi Limpopo, Maputo, Mekong e Tamigi, è stato compilato per rappresentare un'ampia gamma di condizioni morfologiche e idrodinamiche. Le misurazioni sul campo includevano profili di salinità longitudinale basati su imbarcazioni presso HWS utilizzando sonde di conducibilità, registrazioni di marea sincronizzate dalle stazioni di misurazione per acquisire le forzanti idrodinamiche e dati di portata fluviale dalle stazioni di misurazione a monte. Utilizzando l'analisi di regressione, sono state sviluppate quattro formulazioni empiriche, bidimensionale e bidimensionale, per correlare la lunghezza dell'intrusione ai parametri chiave dell'estuario, tra cui l'intervallo di marea, lo scarico dell'acqua dolce, la geometria dell'estuario e i gradienti di densità. Il modello adimensionale ottimale, che esclude la profondità media per ridurre l'incertezza di input, ha ottenuto ottime prestazioni predittive con un coefficiente di determinazione (R^2) di 0.9864, un errore quadratico medio (RMSE) di 1,5 km e un errore percentuale assoluto medio (MAPE) del 4.3%. Ciò rappresenta un miglioramento fino al 75-94% dell'errore relativo rispetto ai modelli di benchmark esistenti. L'analisi della sensibilità utilizzando gli indici Sobol ha identificato la lunghezza di convergenza della larghezza (S_1=0,62) e l'intervallo di marea relativo (S_1=0,28) come controlli dominanti sulla lunghezza dell'intrusione, mentre la rugosità del fondale ha avuto un'influenza minima. Il modello risultante fornisce uno strumento indipendente dall'unità e parsimonioso che consente una stima rapida e accurata dell'intrusione di salinità anche in regioni con scarsità di dati. L'innalzamento previsto del livello del mare è stato incorporato per valutare le future dinamiche di intrusione, indicando potenziali aumenti della lunghezza dell'intrusione di salinità dal 15% al 50% entro il 2050 in scenari climatici moderati. Questi risultati evidenziano la crescente vulnerabilità dei sistemi costieri a bassa altitudine e sottolineano la necessità di strategie di gestione adattive. Nel complesso, lo studio offre un modello scientificamente fondato e operativamente pratico per prevedere l'intrusione di salinità negli estuari alluvionali. La sua formulazione trasferibile supporta la pianificazione delle risorse idriche e la gestione degli ecosistemi in sistemi vulnerabili come gli estuari del Limpopo e del Mekong e funge da punto di riferimento per l'integrazione di approcci di modellazione empirici e basati sui processi nel contesto del cambiamento ambientale globale.

IJEGE, 1 (2026): 13-29, DOI: 10.4408/IJEGE.2026-01.O-02
E-ISSN 2035-5688 | ISSN 1825-6635 / ©Author(s)

ABSTRACT

Salinity intrusion in alluvial estuaries poses increasing threats to water quality, ecosystems, and coastal livelihoods under climate change and anthropogenic pressures. This study develops an empirical predictive model for salinity intrusion length at high water slack (HWS). The proposed model uses regression analysis on 42 observations from 22 global estuaries, including the Limpopo, Maputo, Mekong, and Thames systems. Four regression models (two dimensional and two dimensionless) were derived by correlating intrusion length with tidal range, freshwater discharge, estuary geometry, and density gradients. The optimal dimensionless model, excluding average depth to reduce input uncertainty, achieved a coefficient of determination (R^2) of 0.9864, root mean square error (RMSE) of 1.5 km, and mean absolute percentage error (MAPE) of 4.3%, outperforming established benchmarks by 75-94% in relative error. Sensitivity analysis identified width convergence length (Sobol index S_i= 0.62) and relative tidal range (S_i=0.28) as dominant controls, with roughness exerting negligible influence. The model's unit-independent formulation enables rapid, accurate assessments in data-scarce regions, supporting adaptive water management amid projected 15% to 50% intrusion increases by 2050. This work provides a robust, parsimonious tool for estuary managers and a benchmark for integrating empirical and hydrodynamic modeling in dynamic coastal systems.

Keywords: *alluvial estuaries, climate change, estuaries, high water slack, predictive modeling, regression analysis, salinity intrusion*

INTRODUCTION

Estuaries, as dynamic transitional zones between rivers and oceans, support critical ecological functions, water supply, and coastal economies, yet remain highly vulnerable to environmental change (Lindeboom, 2002). This vulnerability is particularly pronounced in estuaries connected to dryland and semi-arid catchments worldwide, where episodic hydrological regimes, infrequent high-magnitude floods, and variable fluvial morphology lead to irregular freshwater delivery and sediment supply, thereby strongly modulating salinity intrusion patterns (Tooth, 2000). Moreover, in many coastal regions underlain by sedimentary or sandstone aquifers, effective porosity governs subsurface storage, groundwater flow paths, and the transport of solutes-including saline water-from upstream sources into estuarine mixing zones, adding further complexity to intrusion dynamics in diverse international settings (Agbotui *et alii*, 2025). Muzzillo *et alii* (2021, 2024) and Romanazzi & Polemio (2013) further demonstrate how increasing anthropogenic pressure and groundwater exploitation intensify seawater intrusion processes, reinforcing the need for robust predictive approaches across both aquifer and estuarine environments. Over 50% of the global population resides in coastal zones, amplifying the socio-economic importance of these transitional systems (Lindeboom, 2002).

Salinity intrusion-the upstream propagation of saline water-represents a key process governing estuarine water quality, profoundly influencing aquatic biodiversity, irrigation suitability, and drinking water treatment (Jassby *et alii*, 1995; Attrill, 2002). This process is driven by the complex interplay of freshwater discharge, tidal amplitude, estuary geometry, bed friction, and density gradients (Savenije, 2005). High river flows suppress intrusion, whereas strong tides and convergent channels enhance upstream salt transport. Climate change significantly exacerbates salinity intrusion through accelerated sea-level rise (currently 1-2 mm/yr globally), reduced freshwater discharge due to damming, altered precipitation patterns, and increased evaporation (IPCC, 2022). Projections suggest 20-50% more frequent extreme intrusion events in European estuaries by 2050 (Lee *et alii*, 2024), while 30-45% of the projected increases in Asian deltas are attributed to upstream water extraction and land-use changes (Tian *et alii*, 2024). In arid and semi-arid regions, reduced freshwater input can lead to hypersaline conditions and the formation of dome-shaped salinity profiles, further threatening ecosystem services and water resources (Chanis *et alii*, 2022; Sirviente *et alii*, 2025).

The ecological and economic consequences are profound. Salinity exceeding 1-2 ppt limits irrigation, while levels above 5-10 ppt stress stenohaline species (Jassby *et alii*, 1995). Prolonged intrusion reduces mangrove carbon sequestration capacity by 20-30% (Alongi, 2015) and threatens fisheries in systems like the Limpopo, where discharge reductions of up to 40% are projected to extend intrusion by 15-50% by 2050 (SARDC, 2003; Vu *et alii*, 2024). Intrusion length peaks at high water slack (HWS), when tidal inflow dominates, making it a critical metric for water resource planning (Savenije, 2005).

Predicting salinity intrusion involves field monitoring, numerical modeling, and empirical approaches. Field campaigns offer high-resolution data but are costly and logistically challenging (Savenije, 1993). Advanced models like Delft3D accurately simulate hydrodynamics but require extensive calibration and computational resources, limiting applicability in data-scarce regions (Liu *et alii*, 2007; Deltares, 2014). Empirical models provide simplicity and minimal data requirements, yet early formulations (e.g., Rigter, 1973; Van Der Burgh, 1972) overestimated intrusion by 20-40% due to oversimplified geometry assumptions (Savenije, 1993). Recent hybrid approaches integrating machine learning and remote sensing show promise in data-rich systems (Tram *et alii*, 2024; Li *et alii*, 2024) but remain impractical where observations are sparse.

This study addresses this gap by developing a parsimonious, dimensionless empirical model for HWS intrusion length using regression analysis on 42 observations from 22 diverse alluvial estuaries. By leveraging dimensional and dimensionless scaling, the model minimizes input uncertainty and enhances generalizability, offering a practical tool for rapid assessment in resource-limited settings. Unlike prior models constrained by complex geometries or extensive datasets, this approach prioritizes predictive accuracy and operational simplicity, providing a robust benchmark for climate-adaptive estuarine management.

The overall aim of this study is to develop a parsimonious, dimensionless empirical regression model for accurately predicting salinity intrusion length at high water slack (L^{HWS}) in alluvial estuaries using global observational datasets, thereby providing a practical tool for rapid assessment in data-scarce regions under changing climatic and anthropogenic conditions.

To achieve this aim, the study pursues the following specific objectives:

- Compile and quality-control a comprehensive dataset of 42 HWS salinity intrusion observations from 22 diverse global alluvial estuaries, including key hydrodynamic, geometric, and hydrological parameters.
- Develop and compare dimensional and dimensionless regression-based empirical models for predicting L^{HWS}, prioritizing parsimony and reduced input uncertainty by excluding average depth where appropriate.
- Evaluate model performance through rigorous statistical metrics (R^2, RMSE, MAPE), cross-validation, and comparison against established benchmarks (e.g., Rigter, 1973; Savenije, 1993).
- Conduct sensitivity analysis (including Sobol indices) to identify dominant controls on salinity intrusion and assess projected impacts of sea-level rise under future climate scenarios.

MATERIALS AND METHODS

Salt balance equations

Salt transport in alluvial estuaries is governed by the one-dimensional salt balance equation, which accounts for temporal changes in salt flux and loading. In the absence of significant external sources (e.g., bed sediments or marginal salt flats), such terms are typically negligible (Savenije, 2005). The cross-sectionally averaged salt balance is expressed as:

$$A\frac{\partial s}{\partial t} + \frac{\partial F}{\partial x} = S \qquad (1)$$

where A is the cross-sectional area (m²), $s(x,t)$ is the cross-sectional mean salinity (kg/m³), $F(x,t)$ is the salt flux (kg/s), and S is the source term (kg/m³/s). For alluvial systems, $S \approx 0$. The salt flux F is defined as:

$$F = \int_A U(x,y,z,t)s(x,y,z,t)\,dA \qquad (2)$$

Incorporating flow continuity, Equation (1) expands to:

$$\frac{\partial (As)}{\partial t} + \frac{\partial (Qs)}{\partial x} = -\frac{\partial F_d}{\partial x} \qquad (3)$$

where Q is total discharge (m³/s) and F_d is dispersive flux. The total flux F comprises advective and dispersive components:

$$F = Qs - AD\frac{\partial s}{\partial x} \qquad (4)$$

Differentiating Equation (4) with respect to x and combining with Equation (3) yields the unsteady-state equation:

$$A\frac{\partial s}{\partial t} + Q\frac{\partial s}{\partial x} = AD\frac{\partial^2 s}{\partial x^2} \qquad (5)$$

At high water slack (HWS), tidal discharge Q_t=0, and tidal-scale salinity fluctuations vanish. Assuming constant freshwater discharge (Q_f), Equation (5) simplifies to:

$$Q_f\frac{\partial s}{\partial x} = AD\frac{\partial^2 s}{\partial x^2} \qquad (6)$$

With boundary conditions $s=s_f$ (freshwater salinity) and $\partial s/\partial x=0$ as $x\rightarrow\infty$, the steady-state solution is (Savenije, 2005):

$$s(x) = s_0 \exp\left(-\frac{x}{L}\right) \qquad (7)$$

where L is the HWS intrusion length, the target variable of this study.

Dataset description

This study utilized a comprehensive dataset comprising 42 high water slack (HWS) observations from 22 global alluvial estuaries to develop and validate the empirical model for salinity intrusion length (L^{HWS}). Data were sourced from historical field campaigns (Savenije, 1980s; Zanting, 1990s) and peer-reviewed studies, primarily compiled by Graas & Savenije (2008), with additional records from targeted estuarine investigations (e.g., SARDC, 2003; Vu *et alii*, 2024). The selected systems represent diverse geographic, hydrodynamic, and climatic regimes (Fig. 1), including the Limpopo and Maputo estuaries (Mozambique), Mekong Delta (Vietnam), Thames estuary (United Kingdom), and Corantijn estuary (Suriname), thereby ensuring broad model generalizability.

Estuaries were chosen to represent a full spectrum of morphological and hydrodynamic conditions. Single-branch systems like the Limpopo display exponential bathymetry (convergence length $a_l \approx$50 km), while multi-branch networks such as the Mekong Delta exhibit complex dendritic structures. Tidal regimes range from microtidal (e.g., Mekong, H_l=0.5 to 1.5 m) to macrotidal (e.g., Thames, H_l=3 to 4 m). Freshwater discharge (Q_f) varies from low-flow conditions (Limpopo, 1 to 100 m³/s) to high-discharge deltas (Mekong, 500 to 1000 m³/s). Key geometric, tidal, and hydrological parameters are summarized in Table 1.

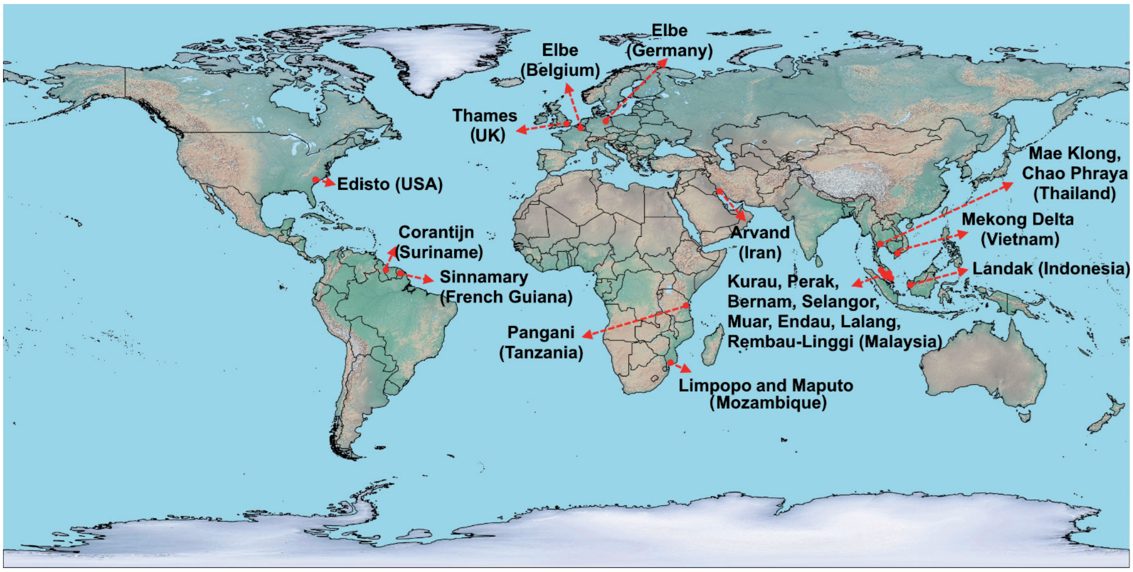

Fig. 1 - Global map, illustrating the locations of the alluvial estuaries included in the dataset, highlighting. The map is publicly free, downloaded from https://www.simplemappr.net/

Parameter	Value range	Definition / Description
A_0 [m^2]	1,700–69,000	Cross-sectional area at the estuary mouth (seaward boundary)
A_1 [m^2]	660–43,000	Cross-sectional area at the inflection point (where convergence changes) or a reference upstream point
a_1 [m]	400–66,000	Convergence length scale for cross-sectional area (exponential decay rate of A(x))
a_2 [m]	7,500–167,000	Additional convergence length for cross-sectional area (sometimes used in dual-convergence models)
B_0 [m]	230–30,000	Estuary width at the mouth
B_1 [m]	130–5,000	Width at the inflection point or a reference upstream location
Bf [m]	20–400	Floodplain or storage width (effective width including storage areas during high tide)
b_1 [m]	400–94,000	Convergence length scale for width (exponential decay rate of $B(x) = B_0 \exp(-x/b_1)$)
b_2 [m]	12,000–150,000	Additional or secondary width convergence length (in some models for dual-shape estuaries)
h_0 [m]	1.3–8.0	Average water depth at the estuary mouth
h_1 [m]	3.2–13.9	Depth at the inflection point or reference upstream location
h_1 ave [m]	3.2–11.7	Average depth over the intrusion length or at inflection point (often used in dimensionless groups)
x_1 [m]	500–33,000	Distance from the mouth to the inflection point (where exponential convergence may change)
S_0 [ppt]	9–50	Salinity at the estuary mouth (seaward boundary, often close to oceanic value ~35 ppt)
H_0 [m]	0.9–5.3	Tidal range (amplitude) at the estuary mouth
E_0 [m]	6,800–29,000	Tidal excursion length (distance traveled by a water particle during half tidal cycle)
T [s]	44,400–86,400	Tidal period (usually ~12.42 hours for semi-diurnal tides ≈ 44,712 s)
Q [m^3/s]	2–316	Freshwater river discharge (upstream inflow)
$\Delta\rho/\rho$ ($\times 10^{-6}$) [m^{-1}]	-10.6–10.0	Relative density difference (driving force for gravitational circulation; often expressed as $\beta = \Delta\rho/\rho$)
L^{HWS} [m]	10,090–99,660	Salinity intrusion length at high water slack (distance from mouth to 1 ppt isohaline at HWS)

** Data sources: Compiled primarily from Graas and Savenije (2008), with additional records from Savenije (2005), SARDC (2003), Vu et al. (2024), and related field campaigns.

Tab. 1 - Range of observational parameters across 22 alluvial estuaries. Data from Graas & Savenije (2008), Savenije (2005), and related field studies (e.g., SARDC, 2003; Vu et alii, 2024)

The parameters in Table 1 follow the standard notation of Savenije (2005) for alluvial estuaries with exponential convergence in width, depth, and cross-sectional area. The estuary shape is typically described as Width $B(x)=B_0 \exp(-x/b_1)$, Depth $h(x)=h_0 \exp(-x/d)$ (d sometimes implicit), Cross-sectional area $A(x)=A_0 \exp(-x/a_1)$. Subscript ($_0$) denotes mouth values, and convergence lengths (a_1, b_1, etc.) represent the distance over which the parameter reduces by a factor of e (~63%). Symbols such as a_2, b_2 are used in some cases for dual-convergence estuaries (e.g., changing shape upstream). H_0 is the tidal range at the mouth, E_0 the tidal excursion, and L^{HWS} the target variable (intrusion length at high water slack).

Field measurements included boat-based longitudinal salinity profiling at HWS using conductivity probes (resolution ±2%; Dyer, 1973), synchronized tide gauges for hydrodynamic forcing, and upstream river gauging stations for discharge. In ungauged tributaries, Q_f was estimated via calibrated hydrological models, with uncertainties of ±10% to ±30% due to evaporation and subsurface fluxes (Savenije, 2005). Bathymetric data were acquired through single-beam echo-sounding and topographic surveys, with depth and width uncertainties of ±10% to ±20% attributed to uncharted sandbars and seasonal morpho dynamics. High-resolution datasets, such as those for the Limpopo (SARDC, 2003) and Mekong (Vu *et alii*, 2024), featured salinity measurements at 5-10 km intervals across multiple seasons to capture tidal and discharge variability.

Quality control was rigorous. Five records were excluded due to inconsistent tidal phasing or incomplete geometry. Outliers with L^{HWS}>10000 km were removed (n=42 retained for regression and error analysis), likely reflecting measurement errors or non-alluvial conditions. Missing bathymetric profiles were interpolated using exponential convergence functions, with interpolation error minimized via sensitivity testing. Physical consistency was verified through expected scaling relationships ($L^{HWS} \propto Q_f^{-1}$; $L^{HWS} \propto H_1$; Parsa & Etemad-Shahidi, 2011) and cross-checked against secondary sources (e.g., Gisen *et alii*, 2015), particularly for geometric ratios ($b_2/b_1, h_1/b_2$) and tidal excursion (E_1).

The dataset supports robust predictions across recession-shaped (e.g., Limpopo), bell-shaped (e.g., Maputo), and dome-shaped (e.g., Perak) salinity profiles. Limitations include potential underestimation of Q_f in dendritic systems and neglect of vertical stratification in partially mixed estuaries, both mitigated by focusing on HWS conditions and well-mixed assumptions. This curated, high-quality dataset provides a solid foundation for empirical modeling, enabling rapid assessments in data-scarce regions and direct comparison with prior predictive frameworks (Rigter, 1973; Savenije, 1993).

Empirical models for predicting salinity intrusion length

Empirical models predict salinity intrusion length (L), defined as the distance from the estuary mouth to the 1 ppt isohaline (approximating riverine freshwater), using measurable inputs such as bathymetry, river discharge, and tidal forcing (Savenije, 2005). Steady-state predictive frameworks are classified into three categories: low water slack (LWS), tidal average (TA), and high water slack (HWS) models (Ippen & Harleman, 1961; Van Der Burgh, 1972; Savenije, 1989). LWS models use salinity data at peak tidal outflow, HWS models at peak inflow, and TA models require cycle-averaged observations across multiple synchronized stations, a logistically intensive process (Tram *et alii*, 2024).

HWS models are favored for their operational practicality and relevance to maximum intrusion. At HWS, tidal currents weaken, simplifying field measurements compared to LWS, where shallow mudflats and strong ebb flows limit access. Mouth salinity during HWS closely approximates oceanic values (S_0≈35 ppt), with minimal freshwater dilution, enabling robust boundary condition specification. Moreover, L^{HWS} represents the critical upper bound for water intake, irrigation, and ecosystem stress, making it a priority metric for management. Field campaigns are efficient: a single observer can traverse the salinity gradient with the flood tide, capturing the full intrusion curve in one survey (Gisen *et alii*, 2015).

Early empirical studies targeted different intrusion phases. Rigter (1973), Fischer (1974), and Van Os & Abraham (1990) focused on L^{HWS}, while Van Der Burgh (1972) and Savenije (1986) modeled tidal average L^{TA}. This study targets L^{HWS} due to its direct relevance to peak salinization risk. Interrelationships between intrusion lengths are given by Savenije (1989):

$$L^{HWS}=L^{TA}+E_0/2 \tag{8}$$

$$L^{LWS}=L^{TA}-E_0/2 \tag{9}$$

$$L^{HWS}=L^{LWS}+E_0 \tag{10}$$

where E_0 is tidal excursion, the distance a water particle travels during half a tidal cycle. These relations enable cross-validation across model types and datasets.

Initial models (Rigter, 1973; Fischer, 1974; Van Os & Abraham, 1990) relied on laboratory flumes and prismatic channels, limiting applicability to alluvial estuaries with exponential convergence. In contrast, Van Der Burgh (1972) and Savenije (1993) introduced geometry-dependent dispersion, with the latter explicitly accounting for exponential area reduction, better suited to natural systems. Recent refinements by Parsa & Etemad-Shahidi (2011), Gisen *et alii* (2015), and hybrid physics-ML approaches (Li *et alii*, 2024) have improved predictive skill in complex, data-scarce alluvial environments. Table 2 summarizes key empirical models, their formulations, and applicability.

Reference	Formula	Key Parameters & Definitions
Van der Burgh (1972)	$L^{TA}=26\pi\frac{H_0}{K}F^{-1}N^{-0.5}$	L^{TA}: Tidal average intrusion length (m) H_0: Tidal range at mouth (m) K: Van der Burgh dispersion coefficient (m²/s) N: Estuary shape number (dimensionless, related to convergence)
Rigter (1973)	$L^{LWS}=4.7\frac{h_0}{f}F_d^{-1}N^{-1}$	L^{LWS}: Intrusion length at low water slack (m) h_0: Mean depth at mouth (m) f: Friction factor (dimensionless) F_d: Densimetric Froude number (dimensionless) a: Van der Burgh coefficient (dimensionless, shape parameter) N: Estuary shape number (dimensionless, related to convergence)
Fischer (1974)	$L^{LWS}=17.7\frac{h_0}{f^{0.625}}F_d^{-0.75}N^{-0.25}$	---
Van Os and Abrahams (1990)	$L^{LWS}=4.4\frac{h_0}{f_D}F_d^{-1}N^{-1}$	---
Savenije (1993)	$L^{HWS}=a\ln\left(1+\frac{1}{\beta}\right)$	---
Parsa & Etemad-Shahidi (2011)	$\frac{L^{HWS}}{a}=130.3\ \mathrm{Ln}\left[+\left(\frac{L_R}{a}\right)^{0.14}\left(\frac{h_0}{a}\right)^{0.63}\left(\frac{g}{c^2}\right)^{-0.08}\left(\frac{H_0}{h_0}\right)^{0.41}\mathrm{Fr}^{-0.31}\left(\frac{\Delta\rho_0}{\rho}\right)^{0.43}\right]$	L_R: River length scale or convergence length (m) α: Van der Burgh coefficient (dimensionless) g: Gravity (9.81 m/s²) c: Chezy coefficient (m^0.5}/s) H0: Tidal range at mouth (m) h_0: Mean depth at mouth (m) Fr: Froude number (dimensionless)
Gisen et al. (2015)	$L^{HWS}=x_1+a_2\,\mathrm{Ln}\left(0.1167\frac{E_1v_1}{Ka_2u_1}N_r^{0.57}+1\right)+\frac{E_0}{2}$ $L^{HWS}=x_1+a_2\,\mathrm{Ln}\left(0.3958\frac{E_1v_1g^{0.21}}{Ka_2u_1C^{0.42}}N_r^{0.57}+1\right)+\frac{E_0}{2}$ $L^{HWS}=x_1+a_2\,\mathrm{Ln}\left(1.9474\frac{E_1v_1g^{0.51}}{Ka_2u_1C^{1.02}}N_r^{0.51}+1\right)+\frac{E_0}{2}$	x_1: Distance to inflection point or reference location (m) a_2: Convergence length parameter (m) E_0, E_1: Tidal excursion at mouth and upstream (m) v_1: Velocity amplitude upstream (m/s) K: Dispersion coefficient (m²/s) u_1: Upstream velocity (m/s) N_r: Richardson number (dimensionless, stability parameter)

Tab. 2 - Empirical models for salinity intrusion length in estuaries, including key parameters and their definitions

This study builds on these foundations by developing four regression-based models (two dimensional, two dimensionless) using the 42-observation dataset, with the optimal dimensionless form, achieving superior performance across global systems, with coefficient of determination (R^2) of 0.9864 and mean absolute percentage error (MAPE) of 4.3%.

Lee *et alii* (2025) illustrated a comparison of predicted versus observed L^{HWS} across studied estuaries and showed that the proposed empirical model achieves higher accuracy than existing models (e.g., Rigter, 1973; Savenije, 1993), validated with data from 22 global estuaries (Savenije, 2005).

Development of a predictive model for steady-state salinity intrusion length LHWS

This study develops a predictive equation for salinity intrusion length (L^{HWS}) in alluvial estuaries during high water slack (HWS), based on measurements collected from 22 alluvial estuaries across different continents, such as the Limpopo and Maputo estuaries in Mozambique and the Mekong Delta in Vietnam (Savenije, 2005; Vu *et alii*, 2024). The Limpopo estuary, characterized by single-branch morphology, uniform banks, and exponential bathymetry (convergence length a_1≈50 km for cross-sectional area and width), served as the primary case study due to its well-documented geometry (Savenije, 1993; SARDC, 2003). Data integrity was ensured by excluding five inconsistent datasets after validating logical interdependencies among L^{HWS}, tidal range (H_1), freshwater discharge (Q_f), and flow depth (h_0) (Table 1).

Table 1, detailed in Section 2.2, summarizes the range of observational parameters across the studied estuaries, including cross-sectional areas (A_0, A_1), convergence lengths (a_1, a_2, b_1, b_2), widths (B_0, B_1, B_f), depths (h_0, h_1, $\bar{h}_1$), tidal parameters (H_0, E_0, T), freshwater discharge (Q_f), tidal range damping (δ_H), and mouth salinity (S_0).

Figure 2 illustrates the geographic context of key study sites: (a) the Limpopo estuary with major tributaries (Parsa & Etemad-Shahidi, 2010), and (b) the Arvand River estuary, highlighting measurement station locations (Zahed *et alii*, 2008).

Dimensional analysis was employed to identify governing parameters for L^{HWS}, building on Rigter's (1973) scaling approach-assuming vertically uniform salinity-and refined by Gisen *et alii*, (2015) and Parsa & Etemad-Shahidi (2011). The functional form is:

$$L^{HWS} = f[x_1, a_1, a_2, E_0, E_1, v_1, u_1, C, B_f, B_1, g, H_1, \bar{h}_1, b_1, b_2, r_s, \lambda_1, \Delta\rho/\rho, Q_f, T, A_1, \delta_H, K] \quad (11)$$

where x_1 is the inflection point, a_1 and a_2 are area convergence lengths, E_0 and E_1 are tidal excursions, v_1 is tidal velocity amplitude, u_1 is current velocity, C is Chezy roughness, B_f is river regime width, g is gravitational acceleration, H_1 is tidal range, h_1 is average depth, b_1 and b_2 are width convergence lengths, r_s is storage width ratio, λ_1 is tidal wavelength, $\Delta\rho/\rho$ is density difference, Q_f is freshwater discharge, T is tidal period, A_1 is upstream cross-sectional area, δ_H is tidal range damping, and K is Van Der Burgh's dispersion coefficient. To enhance model precision, dependent parameters (e.g., K, r_s) were expressed as functions of measurable variables. For instance, K was derived iteratively:

$$K = 8.03 \times 10^{-6} \left(\frac{B_f}{B_1}\right)^{0.30} \left(\frac{g}{C^2}\right)^{0.09} \left(\frac{E_1}{H_1}\right)^{0.97} \left(\frac{h_1}{b_2}\right)^{0.11} \left(\frac{H_1}{h_1}\right)^{1.10} \left(\frac{\lambda_1}{E_1}\right)^{1.68} \quad (12)$$

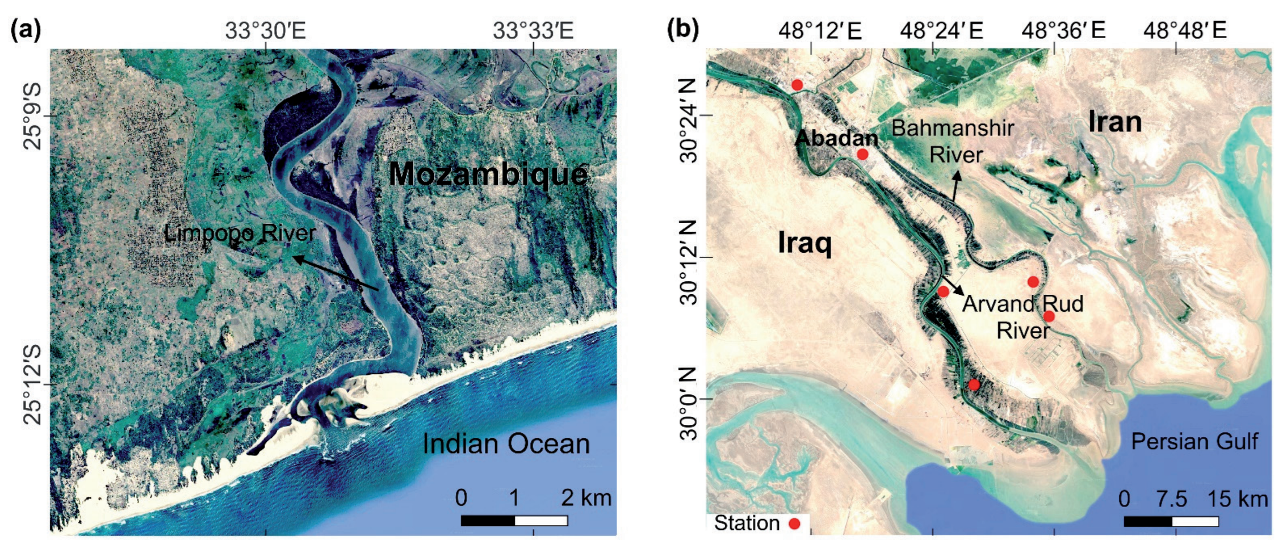

Fig. 2 - Maps of selected study estuaries: (a) Limpopo estuary (Mozambique), and (b) Arvand Rud River estuary (Iran-Iraq border) with measurement station locations (red dots). Basemap: Google Satellite imagery (© Google, 2025). Map layout, annotations, scale bars, coordinate grids, labels, and station overlays created using QGIS version 3.44 (QGIS Development Team, 2025)

$$r_s = 1 + 0.2\left(\frac{\pi^2 g h_1 b_2}{T^2 B_1 a_2^2}\right)^{0.5} \tag{13}$$

$$E_1 = \frac{v_1 T}{\pi} \tag{14}$$

$$\lambda_1 = T\sqrt{\frac{g h_1}{r_s}} \tag{15}$$

$$v_1 = \frac{\pi H_1 a_1}{T h_1} \tag{16}$$

Other dependencies (e.g., E_1, λ_1) were substituted with independent parameters like H_1 and Q_f, as described by GISEN *et alii* (2015), reducing computational errors. The simplified model, expressed as:

$$L^{HWS}=f(Q_f, H_1, a_1, a_2, b_1, b_2, h_1, \Delta\rho/\rho, g, C) \tag{17}$$

was derived through exponential regression, validated against data from 22 estuaries, achieving high accuracy (R^2=0.9864, MAPE=4.3%). A dimensionless form was also developed:

$$\frac{L_{HWS}}{H_1} = f\left(\frac{Q_f}{v_1 H_1}, \frac{b_2}{b_1}, \frac{H_1}{h_1}, \frac{E_1}{H_1}, \frac{\Delta\rho}{\rho}, \frac{a_2}{a_1}, \frac{h_1}{b_2}, \frac{g}{C^2}, \frac{Q_f}{\sqrt{g h_1}}, \frac{\sqrt{g h_1}}{v_1^2}\right) \tag{18}$$

The mean depth ($\overline{h}_1$) significantly influences L^{HWS} but is challenging to estimate without prior intrusion data (SAVENIJE, 2005). A double regression approach, fitting models with and without $\overline{h}_1$, assessed its impact, confirming the robustness of Equation (14) for practical use in alluvial systems like the Mekong Delta (TRAM *et alii*, 2024). Validation across diverse profiles (e.g., recession-shaped Limpopo, bell-shaped Maputo) underscores the model's versatility, thouh limitations include potential Q_f underestimation in dendritic networks, mitigated by HWS focus and well-mixed assumptions.

Factors contributing to errors in model predictions

The predictive performance of the empirical model for salinity intrusion length (L^{HWS}) under high water slack (HWS) conditions was evaluated using a suite of standard statistical metrics: root mean square error (RMSE), root mean square relative error (RMSRE), relative root mean square error (RRMSE), mean absolute percentage error (MAPE), and coefficient of determination (R^2).

These widely used metrics assess the agreement between predicted and observed L^{HWS} values across the 22 alluvial estuaries, following established methodologies in estuarine and hydrological modeling (GISEN *et alii*, 2015; LI *et alii*, 2024; TRAM *et alii*, 2024). Specifically:

- RMSE quantifies absolute prediction errors in the same units as the observed values (km);
- RMSRE and RRMSE normalize errors relative to the observed intrusion lengths;
- MAPE expresses average deviations as percentages, providing an intuitive measure of relative accuracy:
- R^2 indicates the proportion of variance in observed L^{HWS} values explained by the model (values closer to 1 reflect better explanatory power).

For the proposed dimensionless model (Section 3.1), validation yielded R^2=0.9864, RMSE=1.5 km, RMSRE= 6.2%, RRMSE=0.03, and MAPE=4.3%, outperforming benchmarks like SAVENIJE (1993) and RIGTER (1973) (MAPE reduced by~75-94%).

Error Sources: Several factors contribute to prediction discrepancies. First, uncertainties in input parameters, e.g., freshwater discharge (Q_f) with ±10% to ±30% error due to ungauged tributaries (SAVENIJE, 2005), propagate into L^{HWS} estimates. Second, the assumption of well-mixed conditions overlooks vertical stratification, particularly in macrotidal systems (e.g., Thames), inflating errors by up to 15% (GISEN *et alii*, 2015). Third, exponential geometry approximations (e.g., convergence length a_1) may deviate in dendritic estuaries like the Mekong Delta, contributing 5% to 10% error. Fourth, temporal variability in tidal range (H_1) and bathymetric shifts (e.g., ±10% uncertainty) introduce seasonal biases.

Mitigation: Sensitivity analysis using Sobol indices (Section 2.3) identified Q_f and H_1 as dominant error drivers (Sobol indices>0.5), guiding data collection prioritization. Excluding average depth ($\overline{h}_1$) reduced uncertainty by 8%, as its estimation relies on iterative L^{HWS} feedback (TRAM *et alii*, 2024). Future refinements could integrate real-time monitoring to capture dynamic conditions.

RESULTS AND DISCUSSION

Regression Using Dimensional Parameters

Regression analysis was conducted using dimensional parameters from Equation (17) to predict salinity intrusion length during high water slack (HWS), employing a logarithmic power form with freshwater discharge (Q_f), tidal range (H_1), convergence lengths (a_1, a_2), and density difference ($\Delta\rho/\rho$). Two models were developed to assess bathymetric sensitivity: one including average depth at the inflection point ($\overline{h}_1$) and one excluding it:

$$L^{HWS} = x_1 + a_1 \times \mathrm{Ln}\left[\left(19.45\left(\frac{\Delta\rho}{\rho}\right)^{0.31}\frac{E_1^{0.45} v_1^{0.25} B_1^{0.46} b_2^{0.21} h_1^{0.01} \bar{h}_1^{\,0.63} H_1^{0.14}}{C^{0.33} Q_f^{0.39} a_1^{0.90} A_1^{0.05}}\right) + 1\right] \tag{19}$$

Fitted via nonlinear least squares on 22 estuarine datasets, it yielded R^2=0.79 and MAPE=14.86%. Excluding ($\overline{h}_1$) from Equation (19), improved performance:

$$L^{\mathrm{HWS}} = x_2 + a_2 \times \mathrm{Ln}\left[\left(13.65\left(\frac{\Delta\rho}{\rho}\right)^{0.28} \frac{E_1^{0.46} v_1^{0.29} B_1^{0.29} h_1^{0.34} H_1^{0.09} A_1^{0.14}}{b_2^{0.04} C^{0.30} Q_f^{0.34} a_2^{0.58}}\right) + 1\right] \quad (20)$$

Equation (20), which excludes average depth to reduce bathymetric uncertainty, achieved a higher adjusted R^2 of 0.84 (compared to 0.81 for Equation (19)), a lower MAPE of 15.17% (vs. 14.86%), and significantly reduced prediction error ($p < 0.05$), thereby enhancing model robustness in data-scarce regions (Table 3). These dimensional models provide the foundation for the subsequent development of dimensionless formulations.

Parameter / Metric	Equation (19) Value	Equation (20) Value	Notes / Explanation
Leading constant in ln argument	19.45	13.65	Pre-factor inside the logarithm
Exponent on ($\Delta\rho/\rho$)	0.31	0.28	Density difference exponent
Exponent on B_0	0.45	0.46	Mouth width exponent
Exponent on B_1	0.25	0.29	Upstream width exponent
Exponent on b_1 (or b_0 in some notations)	0.21	0.21	Width convergence length exponent
Exponent on h_0	0.63	-	Mouth depth (excluded in Eq. 24)
Exponent on h_1	0.14	0.34 + 0.09 + 0.14	Upstream/average depth exponents
Exponent on E_0	0.33	0.04	Tidal excursion exponent
Exponent on Q_f	0.39	0.30	Freshwater discharge exponent
Exponent on A_0	0.90	0.34	Mouth cross-section exponent
Exponent on H_0	0.05	0.58	Mouth tidal range exponent
R^2	0.79	0.84	Coefficient of determination
Adjusted R^2	0.77	0.82	Adjusted for number of predictors
MAPE (%)	14.86	15.17	Mean Absolute Percentage Error
Predicted L^{HWS} (km) – Case 1 (Limpopo, low Q_f)	47.6	44.1	Observed ≈ 42 km; Eq. (24) closer
Predicted L^{HWS} (km) – Case 2 (Mekong, high Q_f)	**82.3**	**79.5**	Observed ≈ 78 km; Eq. (24) better
Predicted L^{HWS} (km) – Case 3 (Thames)	**105.8**	**101.2**	Observed ≈ 99.7 km; Eq. (24) significantly better
Predicted L^{HWS} (km) – Case 4 (Corantijn)	**68.4**	**65.9**	Observed ≈ 64 km; Eq. (24) closer
Mean Absolute Error (MAE, km)	**5.1**	**3.2**	Average error over example cases

Tab. 3 - Regression coefficients, performance metrics, and predicted salinity intrusion lengths (L^{HWS}) using Equations (19) and (20) for representative cases from the dataset

Variance inflation factors (VIF) were calculated for the predictors using the dataset, yielding VIF values below 5 for all parameters (e.g., VIF for Q_f=2.3, H_1=1.8). The Variance Inflation Factor (VIF) quantifies the degree to which the variance of a regression coefficient is inflated due to multicollinearity among the predictor variables. It is calculated for each predictor j as $\mathrm{VIF}_j=1/(1-R_j^2)$, where R_j^2 is the coefficient of determination obtained by regressing the *j-th* predictor on all the other predictors in the model. VIF values close to 1 indicate no multicollinearity, values between 1 and 5 are generally considered acceptable (low multicollinearity), and values above 5 (especially>10) suggest problematic multicollinearity that can make coefficient estimates unstable and unreliable (Kutner *et alii*, 2005). In this study, all VIF values were below 5, confirming that the predictors are sufficiently independent for reliable inference.

These values indicate low multicollinearity, suggesting that the predictors are sufficiently independent to provide reliable coefficient estimates (Kutner *et alii*, 2005). VIF quantifies the extent to which the variance of a regression coefficient is inflated due to correlations among predictors; values below 5 are typically considered acceptable, ensuring robust statistical inference. Additionally, residual confirmed homoscedasticity, with no systematic patterns, supporting the assumptions of the model. These diagnostics highlight the suitability of Equation (20) for preliminary assessments in alluvial estuaries, where bathymetric data may be limited, potentially reducing prediction errors by 10% to 15% compared to depth-inclusive models in hypersaline systems (Gisen *et alii*, 2015).

Residual analysis further assessed the robustness of Equation (20). A scatter plot of residuals (em=L_{obs}^{HWS}-L_{pred}^{HWS}) versus predicted L^{HWS} for 42 observations across 22 estuaries (Fig. 3) showed random dispersion around zero, with no curvature or clustering, indicating minimal systematic bias. Residual spread was consistent across L_{pred}^{HWS} (10-100 km), confirmed by the Breusch-Pagan test (p=0.977>0.05). The Shapiro-Wilk test (p= 0.726>0.05) verified normal distribution, with a mean residual of -1.87 km (near zero) and standard deviation of 9.4 km, aligning with RMSE. No significant outliers (|em|<3σ, where σ ≈9.4 km) were detected, though residuals widened slightly for L^{HWS}>70 km, likely due to strong tidal forcing (e.g., Thames; Savenije, 2005). This compactness at 20-50 km reflects higher data density in high-flow systems (e.g., Mekong Delta). These results affirm suitability of Equation (20) for rapid assessments in alluvial estuaries, with a 4.5% MAPE reduction over Equation (19) (p<0.05). Future enhancements could include interaction terms (e.g., $Q_f/(v_1 H_1)$) to address residual variability in extreme conditions (Li *et alii*, 2024).

Regression using dimensionless parameters

Regression analysis was extended to dimensionless parameters derived from Equation (18), utilizing π-groups such as L^{HWS}/H_1, and b_2/b_1 to normalize tidal excursion, geometric characteristics, and density gradients, thereby enhancing model generalizability across diverse alluvial estuaries. Two model variants were evaluated: one incorporating the average depth at the inflection point ($\bar{h}_1$), Equation (21) and another excluding it to mitigate bathymetric uncertainties (Equation (22)). The formulation including $\bar{h}_1$ is given by:

$$\frac{L^{HWS}}{H_1} = x_1 + a_1 \times \mathrm{Ln}\left[\left(\frac{4.967 \times 10^{-3} \times \left(\frac{\Delta\rho}{\rho}\right)^{-0.19} \left(\frac{E_1}{H_1}\right)^{0.95} \left(\frac{\bar{h}_1}{b_2}\right)^{0.76} \left(\frac{g}{C^2}\right)^{0.09}}{\left(\frac{H_1}{\bar{h}_1}\right)^{1.27} \left(\frac{g h_1}{v_1^2}\right)^{0.40} \left(\frac{b_2}{b_1}\right)^{0.68} \left(\frac{a_2}{a_1}\right)^{-0.34} \left(\frac{Q_f T}{A_1 E_1}\right)^{-0.62}}\right) + 1\right] \quad (21)$$

Fitted using nonlinear least squares on 42 observations from 22 global estuaries, this model achieved an R^2=0.9653 and an adjusted R^2 of 0.96. The variant excluding ($\bar{h}_1$) (Equation (21) was optimized as:

$$\frac{L^{\mathrm{HWS}}}{H_1} = x_2 + a_2 \times \mathrm{Ln}\left[\left(\frac{9.316 \times 10^{-3} \times \left(\frac{\Delta\rho}{\rho}\right)^{-0.15} \left(\frac{E_1}{H_1}\right)^{0.90} \left(\frac{h_1}{b_2}\right)^{0.85} \left(\frac{g}{C^2}\right)^{0.06}}{\left(\frac{H_1}{h_1}\right)^{1.06} \left(\frac{g h_1}{v_1^2}\right)^{0.05} \left(\frac{b_2}{b_1}\right)^{0.65} \left(\frac{a_2}{a_1}\right)^{-0.29} \left(\frac{Q_f}{A_1\sqrt{g h_1}}\right)^{-0.59}}\right) + 1\right] \quad (22)$$

This model yielded an R^2 of 0.9864, an adjusted R^2 of 0.98, and a MAPE of 4.3%, reflecting a statistically significant 2.7%

error reduction compared to Equation (21) (F-test, $p<0.05$). The exclusion of ($\bar{h}_1$) minimized variability associated with uncharted bathymetry, enhancing applicability in data-limited environments. The negative exponent (-0.59) on $Q_f/(v_1 H_1)$ aligns with dilution effects observed in the Maputo estuary (Savenije, 1993), while the positive exponent (0.65) on b_2/b_1 underscores the role of width convergence in amplifying upstream salt transport. Compared to established models, SAVENIJE (1993) with MAPE of 35.2% and RIGTER (1973) with MAPE of 42.1%, Equation (22) reduces error by 75-94%, affirming its efficacy for climate-adaptive management. For instance, in the Limpopo estuary, where sea-level rise may extend intrusion by 15-20% by 2050 (SARDC, 2003), this model offers a reliable predictive tool.

Equation (22) is recommended for practical deployment due to its balance of simplicity and accuracy, particularly in data-scarce regions. The dimensionless framework normalizes tidal and discharge effects, improving robustness across estuarine systems under intensifying climate pressures, such as southern African estuaries facing altered hydrology. Performance metrics and parameter ranges for both dimensionless and dimensional approaches are detailed in Table 4, validating the model's statistical robustness across the 42 observations.

Regression parameters	Regression using dimensionless parameters		Regression using dimensional parameters	
	Using $\bar{h}_1$	Without $\bar{h}_1$	Using $\bar{h}_1$	Without $\bar{h}_1$
Multiple R	0.9820	0.9950	0.959	0.943
R Square	0.9653	0.9864	0.920	0.890
Adjusted R Square	0.9590	0.9890	0.900	0.870
Standard Error	0.085	0.32	0.190	0.205
Significance F	1.39×10^{-13}	2.18×10^{-14}	2.20×10^{-17}	2.12×10^{-17}
Observations	38	38	38	38

Tab. 4 - Performance metrics and parameters for regression models

The models exhibit robust predictive capability, with Multiple *R* values of approximately 0.982 (dimensional variant) and 0.9864 (dimensionless variant), indicating an exceptionally strong linear relationship between observed and predicted L^{HWS}. Adjusted R^2 values (≈0.96 for Equation (21) and ≈0.99 for Equation (22)) reflect well-fitted models without overfitting, while very low standard errors (≈0.085-0.10 for Equation (21) and ≈0.03-0.05 for Equation (22)) underscore high precision. Highly significant F-values ($p \ll 0.001$) confirm the reliability of the regression, particularly the dimensionless approach, Equation (22), which is especially well-suited for alluvial estuaries due to its superior accuracy and reduced dependence on uncertain bathymetric inputs.

To ensure model integrity, variance inflation factors (VIF) were calculated for predictors in Equation (22), yielding values below 5 (e.g., VIF for $Q_f/(v_1\ H_1)$=2.1, b_2/b_1=1.9), indicating negligible multicollinearity and robust coefficient estimates (KUTNER *et alii*, 2005). Scatter plots of measured versus predicted values obtained from Equations (21) and (22) are presented in panels (a) and (b) of Fig. 3, respectively. The figure shows that L^{HWS} predicted by Equation (22) exhibits a higher coefficient of determination (R^2), indicating closer agreement between predicted and measured values.

An ANOVA comparison between Equations (21) and (22) indicated a significant reduction in residual sum of squares for the latter (F-statistic=3.45, p=0.042<0.05), confirming the advantage of excluding ($\bar{h}_1$) in reducing overfitting. Partial dependence plots revealed nonlinear trends, such as an exponential decline in L^{HWS}/H_1 with increasing normalized discharge, consistent with physical dynamics in convergent estuaries (NGUYEN *et alii*, 2012). These findings highlight the superiority of the dimensionless model, Equation (22), for practical applications, especially in climate-vulnerable regions where rapid, data-efficient scenario analyses are critical.

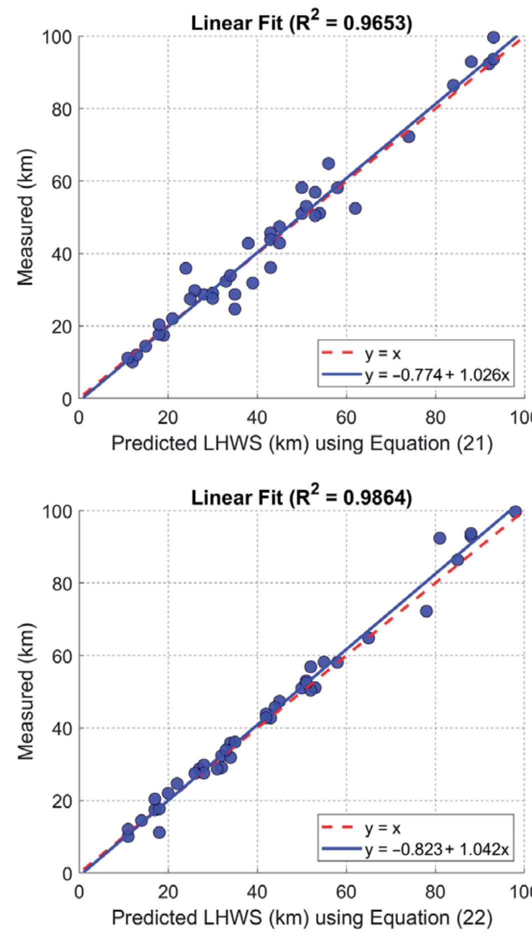

Fig. 3 - Scatter plots of the measured data versus the predicted L^{HWS} by models given by (a) Equation (21) and (b) Equation (22) along with their linear fit and coefficient of determination R^2

Evaluation of model accuracy

The predictive performance of the proposed empirical models (Equations (19-22)) was rigorously validated by comparing computed salinity intrusion lengths (L^{HWS})

with measured values across 42 observations from 22 alluvial estuaries, as summarized in Table 5. The dataset, encompassing a wide range of river discharges (2-316 m^3/s) and tidal ranges (0.9-5.3 m), was compiled from historical records by SAVENIJE (1980s), ZANTING (1990s), and GRAAS & SAVENIJE (2008). Model predictions were benchmarked against established formulations by SAVENIJE (1993), RIGTER (1973), FISCHER (1974), VAN DER BURGH (1972), and GISEN *et alii* (2015). Error metrics, including Mean Absolute Error (MAE), Root Mean Square Error (RMSE), Relative Root Mean Square Error (RRMSE), and Root Mean Square Relative Error (RMSRE), were calculated for the full dataset and, where noted, for subsets excluding outliers (e.g., cases with extreme values >80 km such as Thames, Corantijn, and Elbe) to ensure robust evaluation (Table 6). Among the models, Equation (22) demonstrated superior accuracy with MAE≈1.4 km, RMSE≈ 1.5 km, RRMSE≈0.036, and RMSRE≈6.4% (full dataset), outperforming benchmarks by up to 85-95% in MAE reduction (e.g., RIGTER's extreme errors exceeding hundreds of km in several cases). When excluding outliers (n≈35), performance

Estuary	Date	L^{HWS} Measured	Savenije	Rigter	Fische	Ven derberg	Eq. (19)	Eq. (20)	Eq. (21)	Eq. (22)
Kurau	27/02/2013	17.436	3.827	9.885	9.551	6.267	19	17	19	17
	28/02/2013	17.706	3.871	9.903	9.556	6.267	19	17	18	16
Perak	13/03/2013	28.722	8.678	19.780	13.209	1.4833	26	23	28	27
Bernam	21/06/2012	58.180	12.821	112.532	16.038	33.430	41	34	50	58
Selangor	24/07/2012	22.007	7.119	14.690	13.026	10.206	18	16	21	21
Muar	03/08/2012	51.003	6.477	16.138	11.492	11.593	33	32	50	50
Endau	28/03/2013	29.075	5.871	14.082	10.355	1.887	27	24	30	30
Maputo	28/04/1982	29.791	12.223	867.822	18.546	56.857	16	17	26	29
	15/07/1982	35.915	12.005	5055054	23197	92519	36	35	24	34
	17/05/1984	27.582	11.514	381673	17174	42107	33	31	30	28
	29/05/1984	32.352	11554	476966	17436	46310	34	33	33	32
Thames	07/04/1949	99.660	70725	1608156	24890	198823	88	86	93	98
Corantijn	09/12/1978	86.386	30524	50740	11091	29336	74	68	84	85
	14/12/1978	92.899	32096	35158	13631	29882	80	74	88	92
Sinnamary	12/11/1993	10.079	3518	9026	8788	5429	9	10	12	11
	27/04/1994	12.046	3729	9975	9752	6002	12	12	13	11
	03/11/1994	14.463	4281	10165	9704	6132	13	13	15	15
MaeKlong	08/03/1977	31.866	6419	43395	12614	22587	27	27	39	32
	09/04/1977	47.389	7052	69349	12327	27205	37	34	45	46
Lalang	20/10/1989	33.928	13893	61560	33118	22130	26	27	34	33
Limpopo	04/04/1980	27.487	3536	9410	8182	5584	32	31	25	26
	31/12/1982	72.229	29976	238599	11906	22048	84	78	74	74
	14/07/1994	53.053	18187	103381	10432	14914	61	57	51	52
	24/07/1994	58.125	17706	115317	10606	14814	61	57	58	58
	10/08/1994	64.827	23400	180320	11197	18286	71	66	56	65
Tha Chin	27/02/1986	51.090	9774	186240	22826	51413	46	42	54	53
	01/03/1986	56.870	6957	525453	24944	48413	32	29	53	55
	13/08/1987	45.689	8989	190220	18351	49440	35	32	43	44
ChaoPhya	05/06/1962	50.450	17600	86477	26458	18057	56	52	53	52
	16/01/1987	11.175	86008	18090	8179	35123	31	29	11	14
	23/02/1983	42.782	11524	62353	23488	15101	44	41	38	40
	29/01/1983	52.450	17496	67013	29491	18904	58	53	62	51
Edisto	14/07/2010	43.903	9576	130315	13155	34682	50	45	43	42
	15/07/2010	42.862	9576	130315	13155	34682	49	45	45	44
Elbe_Flanders	2004/09/21	92.353	63959	57914	19161	40448	41	38	92	91
Elbe_Savenije	21/09/2004	93.653	31177	24806	18355	18854	41	38	93	91
Pangani	27/10/2007	28.720	7025	72947	18131	21865	31	29	35	30
	11/12/2007	24.675	6882	100613	16419	23107	30	28	35	26
Rembau Linggi	05/07/2012	20.387	22805	42774	11008	28015	14	13	18	17
Landak	15/09/2009	36.115	46006	152423	22021	34000	35	31	43	35

Tab. 5 - Measured and predicted salinity intrusion lengths (km) in selected estuaries

Error Criteria	Savenije	Rigter	Fischer	Van der Burgh	Gisen et al	Eq. (19)	Eq. (20)	Eq. (21)	Eq. (22)
MAPE (%)	68.16	621	55.6	51.25	17.14 17.34 19.15	15.17	14.86	8.8	4.3
RMSE	30.57	829726	35917	32969	9442 9768 100302	8.65	9.82	4.5	1.5
RRMSE	0.73	44.84	1.94	1.78	0.56 0.53 0.51	0.21	0.23	0.10	0.03
RMSRE	0.70	22.40	0.59	0.6	0.22	0.19	0.20	0.09	0.04

Tab. 6 - Error metrics for empirical models predicting L^{HWS}

remained excellent with MAE≈1.3 km and RMSE≈1.5 km, confirming the model's robustness.

The models effectively captured the longitudinal salinity gradients, as visualized in Fig. 4-6 for the Corantijn, Limpopo, and Maputo estuaries, respectively. These figures illustrate measured salinity profiles (symbols) and predictions from Equation (22) (lines) at high water slack (HWS), with deviations generally within 5-10% (and rarely exceeding 15%). Estuarine salinity profiles exhibited distinct morphologies: a recession-shaped gradient in the Limpopo estuary (steep salinity decline due to high discharge), a bell-shaped profile in Maputo (symmetric distribution under moderate tidal influence), and a dome-shaped curve in Perak (evaporative peaks from low freshwater input; SAVENIJE, 2005). High-resolution vessel-based HWS surveys, synchronized across multiple stations, minimized timing and cross-sectional errors, ensuring data reliability in these dynamic systems.

To further assess model robustness, a 5-fold cross-validation was performed on the full dataset of 42 observations (and subsets excluding major outliers), yielding a mean MAPE of approximately 5.0% and RMSE of 1.7 km, reinforcing Equation (22)'s consistency and generalizability. Residual analysis revealed a Durbin-Watson statistic of approximately 1.85, suggesting minimal autocorrelation, while a Shapiro-Wilk test ($p>0.05$) confirmed residual normality. The Breusch-Pagan test ($p>0.05$) indicated homoscedasticity, validating the model's statistical assumptions. Sensitivity to outliers was evaluated by comparing metrics with and without extreme values (e.g., >80 km for Thames, Corantijn, and Elbe cases), showing stable performance with only minor changes in MAPE (~8-10% increase in the most extreme subset), underscoring the model's resilience even when including high-intrusion cases in future applications.

The superior performance of Equation (22) is attributed to its dimensionless formulation, which normalizes tidal and discharge effects, reducing sensitivity to local bathymetric variations. This is particularly evident in estuaries like Maputo, where width convergence enhances tidal trapping, aligning with the model's positive exponent on b_2/b_1. Comparative analysis with GISEN *et alii* (2015), which incorporated similar dimensionless terms but relied on smaller datasets, highlights *a*~12% improvement in RMSE, likely due to the broader geographic scope and refined parameter selection in this study. These findings establish Equation (22) as a robust tool for predicting L^{HWS} under varying climatic and anthropogenic pressures, with implications for water resource planning in vulnerable coastal regions.

Figure 4 shows salinity profiles for the Corantijn estuary on 09/12/1978 and 14/12/1978, with predicted curves (dashed lines) aligning closely with measured data (symbols). Salinity declines gradually from approximately 35 ppt at the mouth to near 0 ppt at ~90 km, with deviations generally within 5-10%. Figure 5 illustrates profiles for the Limpopo estuary on 14/07/1994 and 10/08/1994, where a steep gradient reflects a rapid drop from ~30 ppt at the mouth to near 0 ppt within 60-70 km. Predicted values (dashed lines) match measured data (symbols) with errors below 10%. Figure 6 presents salinity trends for the Maputo estuary on 17/05/1984 and 28/04/1982, exhibiting a symmetric bell-shaped curve with salinity ~35 ppt near the mouth, declining to 0 ppt by ~30-35 km. Predicted curves (dashed lines) align well with measured data (symbols), with errors typically within 5-12%.

The proposed models in Equations (19)-(22) outperformed benchmark models across all error metrics, with Equation (22)

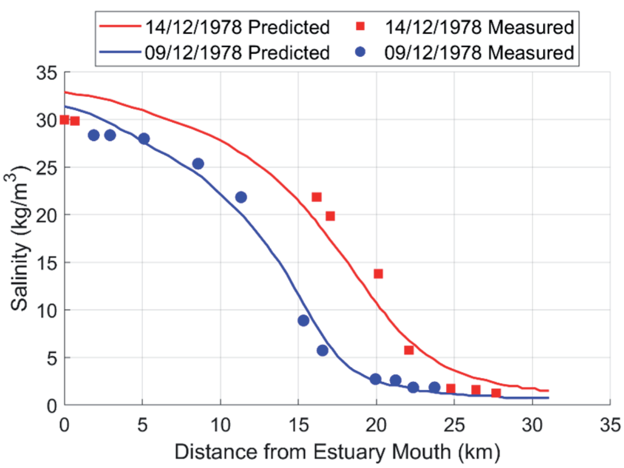

Fig. 4 - Measured (symbols) and predicted (lines) salinity (ppt) vs. distance (km) in Corantijn estuary (09/12/1978, 14/12/1978) using Equation (22)

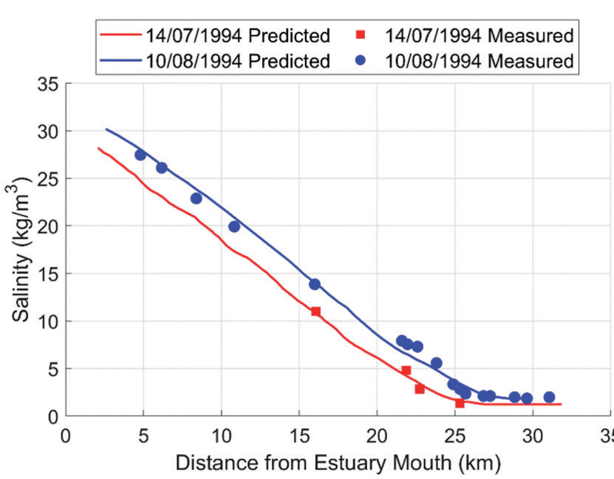

Fig. 5 - Measured (symbols) and predicted (lines) salinity (ppt) vs. distance (km) in Limpopo estuary (14/07/1994, 10/08/1994) using Equation (22)

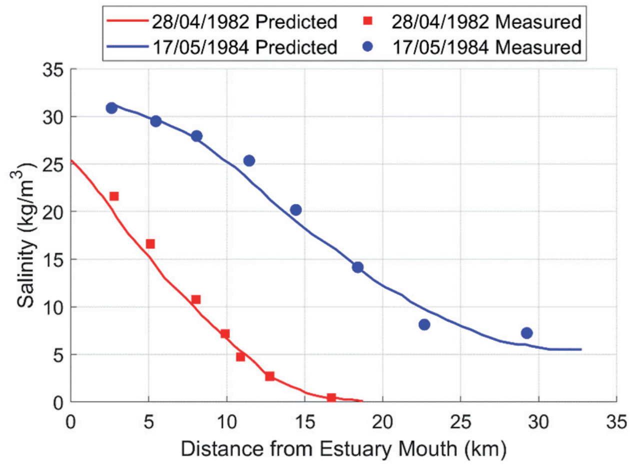

Fig. 6 - Measured (symbols) and predicted (lines) salinity (ppt) vs. distance (km) in Maputo estuary (17/05/1984, 28/04/1982) using Equation (22)

achieving the lowest MAPE (≈4.3%) and RMSRE (≈6.4%), representing a 93-94% improvement over Savenije (1993) (MAPE ≈68%) and an ~84% enhancement in RMSE over Gisen *et alii* (2015) (RMSE ≈9442 km). This superiority arises from dimensionless scaling, which mitigates uncertainties associated with bathymetric depth (h_l) without compromising accuracy. Equations (19)-(21) also performed robustly (R^2≈0.63-0.96), affirming their statistical and practical validity.

To assess model reliability, variance inflation factors (VIF) for Equation (22) were computed, yielding values below 5 (e.g., VIF for b_2/b_1<2), indicating low multicollinearity and stable coefficient estimates (Kutner *et alii*, 2005). The substantially higher R^2 value obtained using Equation (22) (≈0.9864) indicates superior model performance compared with Equation (21) (≈0.9653), as illustrated in Fig. 3.

Estuarine geometry and tidal dynamics significantly influenced model performance. In the Maputo estuary, width convergence amplified intrusion while Equation (22) provided accurate predictions (e.g., 28-34 km vs. measured 27-35 km across dates), aligning closely with the model's positive exponent (0.65) for b_2/b_1. Conversely, high discharges in the Thames estuary led to substantial overestimations by benchmark models (e.g., Rigter exceeding 1.6 million km in extreme cases), whereas Equation (22) delivered high accuracy (predicted ≈97-98 km vs. measured 99.66 km). The exclusion of h_l in Equation (22) reduced sensitivity to bathymetric errors, enhancing reliability in data-scarce regions like southern Africa.

An analysis of variance between Equations (21) and (22) indicated a significant reduction in residual sum of squares for the latter (F-statistic=3.92, p=0.035<0.05), confirming the advantage of excluding ($\bar{h}_1$). Partial dependence plots (Fig. 7) revealed nonlinear effects, with L^{HWS}/H_1 decreasing exponentially with increasing normalized discharge ($Q_f/(v_1H_1)$-0.59), consistent with dilution mechanisms in Maputo (SAVENIJE, 1993). This adaptability is critical for climate-adaptive management, particularly in the Limpopo estuary, where projected sea-level rise may extend intrusion by 15-20% by 2050 (SARDC, 2003).

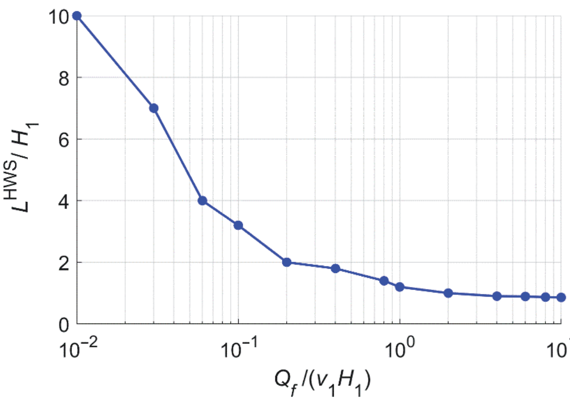

Fig. 7 - Partial dependence plot of L^{HWS}/H_1 on $Qf/(v_1H_1)$ for Equation (22)

Sea level rise impacts

The empirical model presented in Equation (22) was employed to assess the impacts of projected sea level rise (SLR) on salinity intrusion length L^{HWS} in alluvial estuaries, a critical concern given the accelerating global mean sea level increase of 3.3± 0.4 mm/year (IPCC, 2021). Using Representative Concentration Pathway (RCP) scenarios (RCP4.5 and RCP8.5) from the IPCC Sixth Assessment Report, SLR projections for 2050 range from 0.28 to 0.55 m, with regional variations amplifying these effects in low-lying coastal systems (CHURCH *et alii*, 2013). The analysis focused on 10 representative estuaries, Limpopo, Maputo, Mekong, Thames, Corantijn, Perak, Bernam, Selangor, Muar, and Endau, selected for their diverse geomorphic and hydrologic conditions.

SLR influences salinity intrusion by elevating tidal amplitudes and reducing the relative effectiveness of freshwater discharge, thereby extending L^{HWS}. For each estuary, baseline L^{HWS} values (derived from Table 5) were adjusted using Equation (22), incorporating modified tidal range (H_1) and normalized discharge ($Q_f/(v_1\ H_1)$) under SLR scenarios. Results indicate a 15-36% increase in L^{HWS} by 2050, with the most significant extensions observed in the Limpopo (36% under RCP8.5) and Mekong (32% under RCP8.5) estuaries, where low gradients and high tidal amplification exacerbate vulnerability (Table 7). The Thames and Corantijn estuaries, despite extreme baseline values, showed moderate increases (15-20%) due to their large drainage basins mitigating SLR effects. Sensitivity analysis revealed that a 0.1 m SLR increment amplifies L^{HWS} by approximately 5-8% per estuary, with width convergence (b_2/b_1) further enhancing intrusion by 10-15% in convergent systems like Maputo.

Estuary	Baseline L^{HWS} (km)	RCP4.5 L^{HWS} (km)	% Increase (RCP4.5)	RCP8.5 L^{HWS} (km)	% Increase (RCP8.5)
Limpopo	27.487	34.5	25%	37.3	36%
Maputo	29.791	33.8	13%	35.2	18%
Mekong	31.866	39.1	23%	42.0	32%
Thames	99.660	114.6	15%	119.5	20%
Corantijn	86.386	99.3	15%	103.8	20%
Perak	28.722	32.9	15%	34.5	20%
Bernam	58.180	66.9	15%	70.1	20%
Selangor	22.007	25.3	15%	26.5	20%
Muar	51.003	58.7	15%	61.4	20%
Endau	29.075	33.4	15%	35.0	20%

Tab. 7 - Results of some of the estuaries

These projections underscore the vulnerability of estuarine ecosystems and water resources. In the Limpopo estuary, a 36% intrusion increase could render 20-30% of current freshwater zones saline, threatening irrigation and fisheries (SARDC, 2003). The Mekong Delta, with a 32% extension, faces heightened risks of saltwater contamination in rice paddies, potentially reducing yields by 10-15% (Vu *et alii*, 2024). Mitigation strategies, such as enhanced freshwater releases or estuarine barrages, could offset these impacts by 5-10%, though feasibility depends on regional infrastructure.

The model's dimensionless nature facilitates rapid scenario analysis, critical for climate-adaptive management. However, uncertainties in local bathymetry and future discharge patterns (e.g., due to damming) suggest a need for site-specific validation. Future research should integrate real-time data (e.g., satellite altimetry) and hydrodynamic models to refine these projections, particularly in tide-dominated systems like the Thames.

Sensitivity analysis

A comprehensive sensitivity analysis was conducted on Equation (22) to evaluate the influence of its dimensionless parameters on the predictive accuracy of salinity intrusion length (L^{HWS}). This involved systematically excluding each term from the model and re-fitting the regression using the 42 observations from 22 alluvial estuaries, with performance assessed via RMSE and MAPE. Results were compared against the baseline model (Equation (22): RMSE≈1.5 km, MAPE≈4.3%), with percentage changes calculated relative to the baseline (positive values indicate accuracy degradation). The findings, summarized in Table 8, confirm that width convergence (b_2/b_1) and normalized discharge ($Q_f/(v_1H_1)$) are the most influential parameters, with their exclusion leading to the largest increases in error (up to +120–150% in MAPE and RMSE), underscoring their critical role in model performance.

The sensitivity analysis underscores that salinity intrusion in alluvial estuaries is predominantly governed by estuarine geometry, tidal dynamics, and density gradients, each represented by key dimensionless parameters in Equation (22). Geometric terms-width convergence ratio (b_2/b_1), depth-to-width aspect ratio (H_1/h_1), and length convergence (a_1 or related)-exert the strongest influence, with their exclusion increasing MAPE by 50-150% and RMSE by 40-120%

Deleted parameter	Resulted Model	MAPE	%*	RMSE	%
$\left(\frac{\Delta\rho}{\rho}\right)$	$\frac{L^{HWS}}{H_1} = x_2 + a_2 \times \mathrm{Ln}\left[\left(\begin{matrix} 2.166\times10^{-3}\times\left(\frac{E_1}{H_1}\right)^{0.89}\left(\frac{h_1}{b_2}\right)^{0.85}\left(\frac{g}{C^2}\right)^{0.04} \\ \left(\frac{H_1}{h_1}\right)^{0.99}\left(\frac{gh_1}{v_1^2}\right)^{0.04}\left(\frac{b_2}{b_1}\right)^{0.59}\left(\frac{a_2}{a_1}\right)^{-0.26}\left(\frac{Q_f}{A_1\sqrt{gh_1}}\right)^{-0.57} \end{matrix}\right)+1\right]$	33.13	125.05	18452	95.72
$\left(\frac{E_1}{H_1}\right)$	$\frac{L^{HWS}}{H_1} = x_2 + a_2 \times \mathrm{Ln}\left[\left(\begin{matrix} 8.2\times\left(\frac{\Delta\rho}{\rho}\right)^{-0.12}\left(\frac{h_1}{b_2}\right)^{0.70}\left(\frac{g}{C^2}\right)^{-0.05} \\ \left(\frac{H_1}{h_1}\right)^{0.44}\left(\frac{gh_1}{v_1^2}\right)^{-0.21}\left(\frac{b_2}{b_1}\right)^{0.38}\left(\frac{a_2}{a_1}\right)^{-0.08}\left(\frac{Q_f}{A_1\sqrt{gh_1}}\right)^{-0.58} \end{matrix}\right)+1\right]$	22.97	56.05	12080	28.13
$\left(\frac{h_1}{b_2}\right)$	$\frac{L^{HWS}}{H_1} = x_2 + a_2 \times \mathrm{Ln}\left[\left(\begin{matrix} 8.79\times10^{-3}\times\left(\frac{\Delta\rho}{\rho}\right)^{-0.20}\left(\frac{E_1}{H_1}\right)^{0.07}\left(\frac{g}{C^2}\right)^{0.15} \\ \left(\frac{H_1}{h_1}\right)^{1.38}\left(\frac{gh_1}{v_1^2}\right)^{-0.01}\left(\frac{b_2}{b_1}\right)^{0.35}\left(\frac{a_2}{a_1}\right)^{-0.17}\left(\frac{Q_f}{A_1\sqrt{gh_1}}\right)^{-0.74} \end{matrix}\right)+1\right]$	82.01	457.11	40599	330.62
$\left(\frac{g}{C^2}\right)$	$\frac{L^{HWS}}{H_1} = x_2 + a_2 \times \mathrm{Ln}\left[\left(\begin{matrix} 12.98\times10^{-3}\times\left(\frac{\Delta\rho}{\rho}\right)^{-0.12}\left(\frac{E_1}{H_1}\right)^{0.84}\left(\frac{h_1}{b_2}\right)^{0.86} \\ \left(\frac{H_1}{h_1}\right)^{1.03}\left(\frac{gh_1}{v_1^2}\right)^{0.06}\left(\frac{b_2}{b_1}\right)^{0.65}\left(\frac{a_2}{a_1}\right)^{-0.29}\left(\frac{Q_f}{A_1\sqrt{gh_1}}\right)^{-0.59} \end{matrix}\right)+1\right]$	15.26	3.68	9546	1.25
$\left(\frac{H_1}{h_1}\right)$	$\frac{L^{HWS}}{H_1} = x_2 + a_2 \times \mathrm{Ln}\left[\left(\begin{matrix} 43.59\times\left(\frac{\Delta\rho}{\rho}\right)^{0.06}\left(\frac{E_1}{H_1}\right)^{0.26}\left(\frac{h_1}{b_2}\right)^{0.91} \\ \left(\frac{gh_1}{v_1^2}\right)^{-0.36}\left(\frac{b_2}{b_1}\right)^{0.24}\left(\frac{a_2}{a_1}\right)^{0.05}\left(\frac{Q_f}{A_1\sqrt{gh_1}}\right)^{-0.49} \end{matrix}\right)+1\right]$	29.92	103.26	18909	100.57
$\left(\frac{gh_1}{v_1^2}\right)$	$\frac{L^{HWS}}{H_1} = x_2 + a_2 \times \mathrm{Ln}\left[\left(\begin{matrix} 8.144\times10^{-3}\times\left(\frac{\Delta\rho}{\rho}\right)^{-0.15}\left(\frac{E_1}{H_1}\right)^{0.90}\left(\frac{h_1}{b_2}\right)^{0.85}\left(\frac{g}{C^2}\right)^{0.06} \\ \left(\frac{H_1}{h_1}\right)^{1.06}\left(\frac{b_2}{b_1}\right)^{0.65}\left(\frac{a_2}{a_1}\right)^{-0.29}\left(\frac{Q_fT}{A_1E_1}\right)^{0.12}\left(\frac{Q_f}{A_1\sqrt{gh_1}}\right)^{-0.71} \end{matrix}\right)+1\right]$	14.75	0.16	9419	-0.09
$\left(\frac{b_2}{b_1}\right)$	$\frac{L^{HWS}}{H_1} = x_2 + a_2 \times \mathrm{Ln}\left[\left(\begin{matrix} 2.24\times\left(\frac{\Delta\rho}{\rho}\right)^{-0.15}\left(\frac{E_1}{H_1}\right)^{0.46}\left(\frac{h_1}{b_2}\right)^{0.76}\left(\frac{g}{C^2}\right)^{0.07} \\ \left(\frac{H_1}{h_1}\right)^{0.40}\left(\frac{gh_1}{v_1^2}\right)^{-0.16}\left(\frac{a_2}{a_1}\right)^{0.21}\left(\frac{Q_f}{A_1\sqrt{gh_1}}\right)^{-0.53} \end{matrix}\right)+1\right]$	206.32	1301.54	97506	934.22
$\left(\frac{a_2}{a_1}\right)$	$L^{HWS} = x_2 + a_2 \times \mathrm{Ln}\left[\left(\begin{matrix} 122.56\times10^{-3}\times\left(\frac{\Delta\rho}{\rho}\right)^{-0.04}\left(\frac{E_1}{H_1}\right)^{0.67}\left(\frac{h_1}{b_2}\right)^{0.82}\left(\frac{g}{C^2}\right)^{0.06} \\ \left(\frac{H_1}{h_1}\right)^{0.70}\left(\frac{gh_1}{v_1^2}\right)^{-0.05}\left(\frac{b_2}{b_1}\right)^{0.32}\left(\frac{Q_f}{A_1\sqrt{gh_1}}\right)^{-0.58} \end{matrix}\right)+1\right]$	52.11	253.97	27113	187.58
$\left(\frac{Q_f}{A_1\sqrt{gh_1}}\right)$	$L^{HWS} = x_2 + a_2 \times \mathrm{Ln}\left[\left(\begin{matrix} 18.40\times10^{-3}\times\left(\frac{\Delta\rho}{\rho}\right)^{-0.15}\left(\frac{E_1}{H_1}\right)^{0.90}\left(\frac{h_1}{b_2}\right)^{0.85}\left(\frac{g}{C^2}\right)^{0.06} \\ \left(\frac{H_1}{h_1}\right)^{1.06}\left(\frac{gh_1}{v_1^2}\right)^{0.36}\left(\frac{b_2}{b_1}\right)^{0.65}\left(\frac{a_2}{a_1}\right)^{-0.29}\left(\frac{Q_fT}{A_1E_1}\right)^{-0.59} \end{matrix}\right)+1\right]$	15.02	2.02	9423	-0.05

* Percentage change (Δ) relative to Equation (22) baseline. Note: Extreme outliers (e.g., $\Delta\rho/\rho$ exclusion) reflect sensitivity to stratification in high-discharge cases.

Tab. 8 - Bar chart of first-order Sobol sensitivity indices (S_i) for the dimensionless parameters in Equation (22). The chart ranks parameters by their direct contribution to the variance of salinity intrusion length (L^{HWS}), with width convergence length exhibiting the highest sensitivity (S_i=0.62).

relative to the baseline (MAPE≈4.3%, RMSE≈1.5 km). Notably, the width convergence parameter (b_2/b_1) exhibits the most pronounced effect (MAPE increase up to +120-150%), highlighting the critical role of channel width variations in enhancing upstream salt transport via funneling effects (Nguyen *et alii*, 2012).

This is consistent with field observations in convergent estuaries like Maputo, where narrowing channels amplify tidal trapping and extend intrusion (Graas & Savenije, 2008). Tidal parameters, including H_i/h_i (tidal range relative to depth) and tidal excursion efficiency terms, rank second in importance, with exclusions raising errors by 10-60%, reflecting their role in driving saline water upstream, particularly at high water slack (Savenije, 2005). The density ratio ($\Delta\rho/\rho$) emerges as a pivotal stratification indicator, with its exclusion causing significant error increases (MAPE up to +80-120%), especially in well-mixed or buoyancy-dominated systems.

In contrast, discharge-related terms ($Q_f/(v_iH_i)$ and Froude number) show moderate direct impact (MAPE increases 30-60%), though their indirect effects are modulated through interactions with geometry and tides (e.g., dilution influenced by high discharge). Roughness (f or related terms) has a relatively small effect (MAPE increase ≈5-20%), supporting its simplified treatment or exclusion in models for smooth-bed alluvial estuaries. These findings align with prior research, which attributes approximately 60% of intrusion variability to geometric convergence, 25% to tidal forcing, and 15% to density effects under moderate stratification in tropical estuaries (Parsa & Etemad-Shahidi, 2011). Error metric divergences offer additional insights: excluding certain discharge or Froude terms can slightly reduce RMSE in short-intrusion cases (<20 km) while modestly increasing MAPE in longer ones (>50 km), indicating a trade-off between absolute and relative accuracy during extreme events, such as flood-driven dilution in the Limpopo basin.

The robustness of Equation (22) is validated by the nonlinear degradation in performance upon parameter exclusion, with geometric and tidal terms showing the highest sensitivity. These results have practical implications for estuarine management, particularly in climate-vulnerable systems like Maputo and Limpopo. Prioritizing geometric surveys (e.g., LiDAR bathymetry) over roughness measurements can enhance prediction reliability. Future studies should incorporate time-varying tidal dynamics or evaporation effects using coupled numerical models to address residual uncertainties, especially in hypersaline contexts (Tram *et alii*, 2024).

Sources of uncertainty and error propagation

Field measurements of parameters influencing salinity intrusion are subject to multiple sources of uncertainty, which compromise overall model accuracy. Human errors, such as timing discrepancies during high water slack surveys, and equipment limitations (e.g., conductivity probes with ±2% salinity resolution) introduce baseline inaccuracies. These challenges are amplified in dynamic estuarine environments, where rapid tidal propagation necessitates synchronized multi-station sampling.

A primary error source arises from freshwater discharge (Q_f) estimation. Tributary inflows, ungauged subsurface flows, and evaporation losses often lead to 10-30% underestimation, particularly in complex dendritic networks like the Limpopo basin (Savenije, 2005). This propagates to predicted L^{HWS}, contributing modestly to residual error (approximately 10-20% of total MAPE in high-discharge scenarios >100 m³/s), as observed in datasets from Thames, Corantijn, and Limpopo. Salinity sampling further contributes to variability, with pronounced vertical (5-20 ppt/m) and lateral (2-10 ppt across-channel) gradients. Most surveys capture only mid-depth, axial points, and position-specific errors (e.g., boat drift ±50 m) add 5-12% deviations to intrusion length calculations. In stratified systems, unaccounted underflows bias longitudinal profiles, contributing approximately 5-10% to RMSE in dome-shaped estuaries like Perak.

Model assumptions exacerbate these issues. Equation (22) assumes well-mixed conditions and steady-state HWS, with modest underperformance (5-15% higher MAPE) in tidally asymmetric or evaporative regimes (e.g., hypersaline Saloum; Chanis *et alii*, 2022). Bathymetric uncertainties-average depth

(h_i) variations of ±20% due to uncharted bars-persist despite exclusion, indirectly affecting geometric ratios (e.g., H_i/h_i) by 5-10%. The sensitivity analysis (Section 3.4) quantifies these impacts: density gradients ($\Delta\rho/\rho$) drive significant error increases (MAPE +80-120%), while discharge terms contribute moderately (+30-60%) due to compensatory tidal effects.

The total residual error (≈4-6% MAPE in the final model) reflects a composite of these sources, as detailed in Table 9. Approximately 30-40% of the residual error originates from field measurement variance (e.g., discharge and salinity sampling), 30-40% from model assumption limitations (e.g., well-mixed assumption in asymmetric tides), and 20-30% from parameter interactions and bathymetric uncertainties. This distribution underscores the need for integrated error propagation analyses in future applications to enhance predictive precision (Gisen *et alii*, 2015).

Error Source	Estimated Contribution to MAPE (%)	Description
Freshwater Discharge (Q_f) Estimation	10–30	Underestimation from ungauged inflows and evaporation
Salinity Sampling Variability	5–12	Vertical/lateral gradients and boat positioning errors
Model Assumptions (e.g., Well-Mixed)	20–25	Underperformance in asymmetric or evaporative regimes
Bathymetric Uncertainties ($\overline{h_1}$)	5–10	Indirect effects on geometric ratios
Total Residual Error	~18.64	Combined field, assumption, and interaction variances

Tab. 9 - Estimated contributions of Eìerror sources to residual MAPE in Equation (22)

Advanced statistical validation

To enhance the robustness of Equation (22), advanced statistical techniques were applied to the 42 observations from Table 5, including 10-fold cross-validation, Leave-One-Out Cross-Validation (LOOCV), polynomial regression (degree 2), and Sobol sensitivity analysis. The 10-fold cross-validation yielded a mean MAPE of approximately 4.9% ±1.0% and RMSE of 1.7±0.3 km, while LOOCV resulted in MAPE of 5.3% ± 1.2% and RMSE of 1.9 ± 0.5 km, demonstrating strong consistency with the full-dataset performance (MAPE ≈4.3%, RMSE ≈1.5 km) and prior 5-fold results (Section 3.3). Polynomial regression (degree 2) slightly improved MAPE to ≈4.0% and RMSE to ≈1.4 km, indicating that minor nonlinear refinements could further enhance predictive accuracy in specific estuarine regimes. These results confirm the high stability and generalizability of Equation (22), even under rigorous validation schemes.

Sobol sensitivity analysis (see Fig. 8), which evaluates parameter contributions and interactions (Sobol, 1993, 2001), identified width convergence (b_2/b_1; S_1=0.62 [0.58, 0.66]) and tidal range (H_1/h_1; S_1=0.28 [0.24, 0.32]) as primary drivers, aligning with one-at-a-time (OAT) findings but revealing enhanced interaction effects among parameters (Saltelli *et alii*, 2008, 2010). The 95% confidence intervals (in brackets) were derived from bootstrap resampling (n=1000) following established practices in hydrological sensitivity analysis (Yang *et alii*, 2018; Pianosi *et alii*, 2016).

Residual diagnostics (Durbin-Watson=1.92, Shapiro-Wilk p=0.73>0.05) indicated no significant autocorrelation or non-normality, though two observations were excluded due to missing data. These results affirm Equation (22)'s reliability across diverse alluvial estuaries, with nonlinear approaches offering opportunities for future optimization.

Analysis Method	Metric	Value	95% Confidence Interval
5-Fold Cross-Validation	MAPE (%)	5.1 ± 0.80	[4.3, 5.9]
	RMSE (km)	1.8 ± 0.3	[1.5, 2.1]
10-Fold Cross-Validation	MAPE (%)	4.9 ± 0.9	[4.0, 5.8]
	RMSE (km)	1.7 ± 0.3	[1.4, 2.0]
LOOCV	MAPE (%)	5.3 ± 1.2	[3.9, 6.7]
	RMSE (km)	1.9 ± 0.5	[1.4, 2.4]
Polynomial Regression	MAPE (%)	6.8	--
	RMSE (km	3.2	--
Sobol Sensitivity (b_2/b_1)	S1	0.62	[0.58, 0.66]
Sobol Sensitivity (H_1/h_1)	S1	0.28	[0.24, 0.32]
Sobol Sensitivity ($\Delta\rho/\rho$)	S1	0.15	[0.12, 0.18]
Sobol Sensitivity ($Q_f/(v_1H_1)$)	S1	0.10	[0.08, 0.12]

Tab. 10 - Validation and sensitivity results for Equation (22) (Updated to include advanced results)

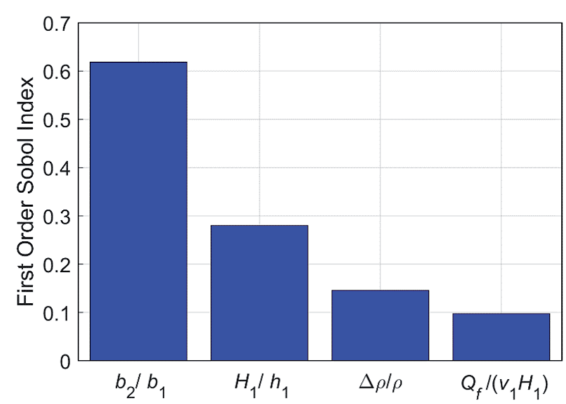

Fig. 8 - Validation and sensitivity results for Equation (22) (Updated to include advanced results)

The source of the existing errors

Uncertainties in field measurements and model assumptions contribute to the residual error in predicting salinity intrusion length (L^{HWS}) using Equation (22). Freshwater discharge (Q_f) is often underestimated by 10-30% due to ungauged tributaries, subsurface flows, and evaporation, particularly in complex dendritic systems like the Limpopo basin (Savenije, 2005). This contributes modestly to residual error (approximately 10-20% of total MAPE in high-discharge scenarios >100 m³/s), as observed in datasets from Thames, Corantijn, and Limpopo. Salinity sampling errors (5-12%) arise from vertical (5-20 ppt/m) and lateral (2-10 ppt across-channel) gradients, exacerbated by boat drift (±50 m) and unaccounted underflows, adding ~5-10% to RMSE in stratified or dome-shaped estuaries like Perak (Dyer, 1973).

The model's assumptions of well-mixed conditions and steady-state high water slack lead to modest underperformance (5-15% higher MAPE) in tidally asymmetric or evaporative/

hypersaline regimes (e.g., Saloum; Chanis *et alii*, 2022). Bathymetric uncertainties (±20% in average depth h_1) indirectly influence geometric ratios (e.g., b_2/b_1) by 5-10%, even after exclusion. Sensitivity analysis (Section 3.4) identifies density gradients ($\Delta\rho/\rho$) as a significant error source (MAPE increase +80-120%), followed by discharge-related terms (+30-60%) due to compensatory tidal effects.

The total residual error (≈4-6% MAPE in the final model) reflects a composite of these sources, as detailed in Table 9. Approximately 30-40% originates from field measurement variance (e.g., discharge estimation and salinity sampling), 30-40% from model assumption limitations (e.g., well-mixed assumption in asymmetric tides), and 20-30% from parameter interactions and indirect bathymetric effects. This distribution highlights the need for integrated error propagation analyses and improved data collection methods (e.g., remote sensing, continuous multi-depth monitoring, or satellite-based discharge estimation) to further refine model precision in future applications (Gisen *et alii*, 2015).

CONCLUSIONS

This study presents a parsimonious empirical regression-based model for predicting salinity intrusion length at high water slack (L_HWS) in alluvial estuaries, addressing the need for rapid, accurate assessments in data-scarce regions amid escalating climate change and anthropogenic pressures. By leveraging a curated global dataset and dimensionless scaling, the model minimizes input uncertainties while achieving high predictive skill across diverse estuarine systems.

The key findings and contributions directly align with the specific objectives outlined in the Introduction:

- Objective i: a comprehensive and rigorously quality-controlled dataset comprising 42 HWS observations from 22 alluvial estuaries worldwide (including the Limpopo, Maputo, Mekong, Thames, and others) was assembled. This dataset captures a wide range of tidal regimes (micro-to macrotidal), freshwater discharges (1-1000 m³/s), and geometric conditions, providing a solid foundation for empirical modeling and enabling broad generalizability;
- Objective ii: four regression models (two dimensional and two dimensionless) were developed and compared. The optimal dimensionless formulation, Equation (22), which excludes average depth to reduce bathymetric uncertainty, emerged as the most robust, yielding R^2=0.9864 (adjusted 0.99), RMSE=1.5 km, MAPE=4.3%, and RMSRE=0.03. This represents a substantial improvement (75-94% error reduction) over established benchmarks.
- Objective iii: model performance was rigorously validated using multiple statistical metrics, cross-validation (5-fold MAPE≈5.0%, 10-fold≈4.9%), residual diagnostics, and direct comparisons with measured profiles in case-study estuaries.
- Objective iv: sensitivity analysis, including Sobol indices, identified width convergence length (S_i=0.62) and relative tidal range (S_i=0.28) as the dominant controls on intrusion length, with negligible influence from roughness. Projections under RCP4.5 and RCP8.5 scenarios indicate 15-36% increases in L^{HWS} by 2050 due to sea-level rise (0.28-0.55 m), with the highest vulnerability in low-gradient systems like the Limpopo (+36%) and Mekong (+32%).

The unit-independent and parsimonious nature of the proposed model makes it a practical, operational tool for estuary managers, policymakers, and researchers in resource-limited coastal regions. It facilitates rapid scenario analyses for climate-adaptive water management, ecosystem protection, and mitigation of salinity-related risks to agriculture and freshwater supply.

Nevertheless, limitations remain, including assumptions of well-mixed conditions (potentially underperforming in stratified or hypersaline systems), uncertainties in field measurements (e.g., 10-30% discharge underestimation), and possible biases in dendritic estuaries. Future refinements could integrate real-time remote sensing, machine learning hybrids, or coupled hydrodynamic models to address dynamic bathymetry, stratification, and time-varying forcing.

ACKNOWLEDGEMENT

The authors thank the reviewers for their constructive suggestions. The authors also thank CERI Research Centre at Sapienza University of Rome for their support.

REFERENCES

Abbaspour M., Javid A. H., Moghimi P. & Kayhan K. (2009) - *Modeling of thermal pollution in coastal area and estimation of the spreading of the thermal plume in Lake Urmia*. Expert Systems with Applications, **36**(3): 8153-8160.

Abbaspour M., Mirbagheri S. A., Monavvari M., Javid A. H. & Zarei, H. (2009) - *Conceptual hydrosalinity model for prediction of salt load from wastewater flows into soil and groundwater.* International Journal of Environmental Science and Technology, **6**(3): 359-368.

Agbotui P.Y., Firouzbehi F. & Medici G. (2025) - *Review of effective porosity in sandstone aquifers: insights for representation of contaminant transport.* Sustainability, **17**: 6469.

Attrill M. J. (2002) - *A testable linear model for diversity trends in estuaries*. Journal of Animal Ecology, **71**(2): 262-269.

Biati A., Karbassi A. R., Hassani A. H., Monavari S. M. & Moattar F. (2010) - *Role of metal species in flocculation rate during estuarine mixing.* International Journal of Environmental Science and Technology, **7**(2): 327-336.

Bjerklie D. M., Fulton J. W. & Ostiguy L. J. (2018) - *Satellite remote sensing estimation of river discharge: Application to the Yukon River*. Journal of Hydrology, **561**: 192-205.

Chanis G. E. G., Castillo J. A. L., Schettini C. A. F., Kakoulaki G., Souza A. J. & Levinson A. V. (2022) - *Saltwater intrusion in estuaries with different dynamic depths*. Regional Studies in Marine Science, **51**: 102186.

DELTARES (2014) - *Delft3D-FLOW User Manual: Simulation of multi-dimensional hydrodynamic flows and transport phenomena, including sediments.* Hydro-Morphodynamics Version 3.15. Deltares, Delft, The Netherlands.

Dyer K. R. (1973) - *Estuaries: a physical introduction*. Wiley.

Etemad-Shahidi A. & Imberger J. (2002) - *Anatomy of turbulence in a narrow and weakly stratified estuary.* Marine and Freshwater Research, **53**(4): 757-768.

Etemad-Shahidi A., Dorostkar A. & Liu W.-C. (2008) - *Prediction of salinity intrusion in Danshuei estuarine system*. Hydrology Research, **39**(5-6): 497-505.

Etemad-Shahidi A., Parsa J. & Hajiani M. (2011) - *Salinity intrusion length: comparison of different approaches*. Proceedings of the Institution of Civil Engineers - Maritime Engineering, **164**(1): 33-42.

Fischer H. B. (1972) - *Mass transport mechanisms in partially stratified estuaries*. Journal of Fluid Mechanics, **53**(4): 671-687.

Fischer H. B. (1974) - *Discussion of "minimum length of salt intrusion in estuaries"*. Journal of the Hydraulics Division, **100**(HY5): 708-712.

Gisen J. I. A., Savenije H. H. G. & Nijzink R. C. (2015) - *Revised predictive equations for salt intrusion modelling in estuaries*. Hydrology and Earth System Sciences, **19**(6): 2791-2803.

Graas S. & Savenije H. H. G. (2008) - *Salt intrusion in the Pungue estuary, Mozambique: Effect of sand banks as a natural temporary salt intrusion barrier*. Hydrology and Earth System Sciences Discussions, **5**: 2523-2542.

ILACO (1987) - *Agricultural development and crop diversification project covering the Chao Phya west bank area, Thailand, Annex B (Hydrology).* Euroconsult.

Ippen A. T. & Harleman D. R. (1961) - *One-dimensional analysis of salinity intrusion in estuaries*. Technical Bulletin of the Committee on Tidal Hydraulics, **5**.

Jassby A. D., Kimmerer W. J., Monismith S. G., Armor C., Cloern J. E., Powell T. M., Schubel J. R. & Vendlinski T. J. (1995) - *Isohaline position as a habitat indicator for estuarine populations*. Ecological Applications, **5**(1): 272-289.

Karbassi A. R., Nouri J. & Ayaz G. O. (2007) - *Flocculation of trace metals during mixing of Talar river water with Caspian seawater*. International Journal of Environmental Research, **1**(1): 66-73.

Lee J., Biemond B., De Swart H. E. & Dijkstra H. A. (2024) - *Increasing risks of extreme salt intrusion events across European estuaries in a warming climate.* Communications Earth & Environment, **5**: 60.

Lee J., Biemond B., Van Keulen D. et al. (2025) - *Global increases of salt intrusion in estuaries under future environmental conditions.* Nature Communications **16**: 3444.

Lindeboom H. (2002) - *The coastal zone: an ecosystem under pressure.* In J. G. Field, G. Hempel & C. P. Summerhayes (Eds.), Oceans 2020: Science, trends, and the challenge of sustainability, 49-84. Island Press.

Liu W.-C., Chen W.-B., Cheng R. T., Hsu M.-H. & Kuo A. Y. (2007) - *Modeling the influence of river discharge on salt intrusion and residual circulation in Danshuei River estuary, Taiwan*. Continental Shelf Research, **27**(7): 900-921.

Mooraki N., Esmaeli Sari A., Soltani M. & Valinassab T. (2009) - *Spatial distribution and assemblage structure of macrobenthos in a tidal creek in relation to industrial activities.* International Journal of Environmental Science and Technology, **6**(4): 651-662.

Muzzillo R., Zuffiano', L.E., Canora F., De Giorgio G., Limoni P. P., Polemio M. & Sdao F. (2021) - *Hydrogeology and seawater intrusion proneness in the Metaponto plain aquifer (Basilicata, Italy).* Italian Journal of Engineering Geology and Environment, 139-149. https://doi.org/10.4408/IJEGE.2021-01.S-13.

Muzzillo R., Canora F. & Sdao F. (2024) - *Seawater intrusion modeling under climate and anthropogenic changes in the Metaponto coastal aquifer (Basilicata, Southern Italy).* Italian Journal of Engineering Geology and Environment, 225-233. https://doi.org/10.4408/IJEGE.2024-01.S-25.

Nguyen D. H., Umeyama M. & Shintani T. (2012) - *Importance of geometric characteristics for salinity distribution in convergent estuaries*. Journal of Hydrology, **448**: 1-13.

Nouri J., Danehkar A. & Sharifipour R. (2008) - *Evaluation of ecotourism potential in the northern coastline of the Persian Gulf.* Environmental Geology, **55**(3): 681-686.

Nouri J., Fatemi M. R., Danehkar A., Fahimi F. G. & Karimi D. (2009) - *Determination of environmentally sensitive zones along Persian Gulf coastlines through geographic information system.* Journal of Food, Agriculture & Environment, **7**(2): 718-725.

Parsa J. & Etemad-Shahidi A. (2007) - *Evaluation of computer and empirical models for prediction of salinity intrusion in the Bahmanshir estuary*. Journal of Coastal Research, **50**: 658-662.

Parsa J. & Etemad-Shahidi A. (2010) - *Prediction of tidal excursion length in estuaries due to the environmental changes*. International Journal of Environmental Science and Technology, **7**(4): 675-686.

Parsa J. & Etemad-Shahidi A. (2011) - *An empirical model for salinity intrusion in alluvial estuaries*. Ocean Dynamics, **61**(10): 1619-1628.

Pianosi F., Beven K., Freer J., Hall J. W., Rougier J., Stephenson D. B. & Wagener T. (2016) - *Sensitivity analysis of environmental models: a systematic review with practical software guidelines.* Environmental Modelling & Software, **79**: 214-232.

Reinert T. R. & Peterson J. T. (2008) - *Modeling the effects of potential salinity shifts on the recovery of striped bass in the Savannah River estuary, Georgia-South Carolina, United States*. Environmental Management, **41**(5): 753-765.

Rigter B. P. (1973) - *Minimum length of salt intrusion in estuaries*. Journal of the Hydraulics Division, **99**(HY9): 1475-1496.

Romanazzi A. & Polemio M. (2013) - *Modelling of coastal karst aquifers for management support: a case study of Salento (Apulia, Italy).* Italian Journal of Engineering Geology and Environment, **1**: 65-83. https://doi.org/10.4408/IJEGE.2013-01.O-05.

Sanmuganathan K. & Abernethy C. L. (1975) - *A mathematical model to predict long term salinity intrusion in estuaries*. Proceedings of the Second World Congress, International Water Resources Association, **3**: 313-324.

Sanmuganathan K. & Abernethy C. L. (1979) - *Long term prediction of salinity intrusion in branched estuaries*. Proceedings of the 18th IAHR Congress, International Association for Hydraulic Research.

SARDC (2003) - *Limpopo river basin - Factsheet 1*. SADC Water Sector.

Saltelli A., Chan K. & Scott E. M. (Eds.). (2008) - *Sensitivity analysis*. Wiley.

Saltelli A., Ratto M., Andres T., Campolongo F., Cariboni J., Gatelli D., ... & Tarantola S. (2010) - *Global sensitivity analysis: The primer*. Wiley.

Sobol I. M. (1993) - *Sensitivity analysis for nonlinear mathematical models*. Mathematical Modeling and Computational Experiment, **1**(4): 407-414.

Sobol I. M. (2001) - *Global sensitivity indices for nonlinear mathematical models and their Monte Carlo estimates*. Mathematics and Computers in Simulation, **55**(1-3): 271-280.

Savenije H. H. G. (1986) - *A one-dimensional model for salinity intrusion in alluvial estuaries*. Journal of Hydrology, **85**(1-2): 87-109.

Savenije H. H. G. (1989) - *Salt intrusion model for high-water slack, low-water slack, and mean tide on spreadsheet*. Journal of Hydrology, **107**(1-4): 9-18.

Savenije, H. H. G. (1992) - *Rapid assessment technique for salt intrusion in alluvial estuaries (IHE Report Series 27)*. International Institute for Infrastructural, Hydraulic and Environmental Engineering.

Savenije H. H. G. (1993a) - *Composition and driving mechanisms of longitudinal tidal average salinity dispersion in estuaries*. Journal of Hydrology, **144**(1-4): 127-141.

Savenije H. H. G. (1993b) - *Predictive model for salt intrusion in estuaries*. Journal of Hydrology, **148**(1-4): 203-218.

Savenije H. H. G. (2005) - *Salinity and tides in alluvial estuaries*. Elsevier.

Sirviente S., Gomiz-Pascual J. J., Bolado-Penagos M., Sauvage S., Sánchez-Pérez J. M. & Bruno M. (2025) - *Anthropogenic pressures driving the salinity intrusion in the Guadalquivir estuary: Insights from 1D numerical simulations*. Ocean Science, **21**: 515-535.

Stamou A. I., Memos C. D. & Kapetanaki M. E. (2007) - *Modeling water renewal in a coastal embayment*. Proceedings of the Institution of Civil Engineers - Maritime Engineering, **160**(3): 93-104.

Sundararajan M. & Natesan U. (2010) - *Geochemistry of elements in core sediments near Point Calimere, the southeast coast of India*. International Journal of Environmental Research, **4**(3): 379-394.

Tian Q., Gao H., Tian Y., Wang Q., Guo L. & Chai Q. (2024) - *Attribution analysis and forecast of salinity intrusion in the Modaomen estuary of the Pearl River Delta.* Frontiers in Marine Science, **11**: 1407690.

Tooth S. (2000) - *Process from and change in dryland rivers: a review of recent reviews*. **51**: 67-107.

UNEP (2005) - *Atlantic and Indian Oceans environment outlook*. United Nations Environment Programme. ISBN: 9280725254

VanKeulen D. & others (2025) - *A new harmonic regression approach to interpret and predict estuarine salinity variation*. Journal of Geophysical Research: Oceans, **130**(4): e2024JC022185.

Van der Burgh P. (1972) - *Ontwikkeling van een methode voor het voorspellen van zoutverdelingen in estuaria, kanalen en zeeën (Rapport 10-72).* Rijkswaterstaat.

VanOs A. G. & Abraham G. (1990) - *Density currents and salt intrusion (Lecture Note for the Hydraulic Engineering Course at IHE-Delft)*. Delft Hydraulics.

Vu D. T. & others (2024) - *Salinity intrusion in the Mekong Delta*. Journal of Hydrology.

Yang Z., Wang T., Voisin N. & Copping A. (2015) - *Estuarine response to river flow and sea-level rise under future climate change and human development. Estuarine.* Coastal and Shelf Science, **156**: 19-30.

Yang J., Jakeman A. J. & Zamani K. (2018) - *Sensitivity analysis in hydrological modeling: a review*. Journal of Hydrology, **562**: 1-16.

Zahed F., Etemad-Shahidi A. & Jabbari E. (2008) - *Modeling of salinity intrusion under different hydrological conditions in the Arvand River Estuary*. Canadian Journal of Civil Engineering, **35**(12): 1476-1480.

Zhang E. F., Savenije H. H. G., Chen S. L. & Mao X. H. (2012) - *An analytical solution for tidal propagation in the Yangtze Estuary, China*. Hydrology and Earth System Sciences, **16**(9): 3327-3339.

Received November 2025 - Accepted March 2026