A New Harmonic Regression Approach to Interpret and Predict Estuarine Salinity Variation

van Keulen, D.; Kranenburg, W. M.; Hoitink, A. J.F.

DOI
10.1029/2024JC022185

Publication date
2025

Document Version
Final published version

Published in
Journal of Geophysical Research: Oceans

Citation (APA)
van Keulen, D., Kranenburg, W. M., & Hoitink, A. J. F. (2025). A New Harmonic Regression Approach to Interpret and Predict Estuarine Salinity Variation. *Journal of Geophysical Research: Oceans*, *130*(5), Article e2024JC022185. https://doi.org/10.1029/2024JC022185

JGR Oceans

RESEARCH ARTICLE
10.1029/2024JC022185

A New Harmonic Regression Approach to Interpret and Predict Estuarine Salinity Variation

D. van Keulen[1,2], W. M. Kranenburg[2,3], and A. J. F. Hoitink[1]

[1]Department of Environmental Sciences, Hydrology and Quantitative Water Management Group, Wageningen University, Wageningen, The Netherlands, [2]Deltares, Delft, The Netherlands, [3]Department of Hydraulic Engineering, Section Environmental Fluid Mechanics, Delft University of Technology, Delft, The Netherlands

Key Points:

- In well-mixed systems, horizontal displacement of the longitudinal salinity profile is the dominant response to storm surges
- Effects of unsteadiness on salinity dynamics become increasingly important toward the downstream sections of alluvial, well-mixed estuaries
- The empirical Van der Burgh constant scales with the spatial scales of dispersion and advection, relative to changes in channel geometry

Supporting Information:

Supporting Information may be found in the online version of this article.

Correspondence to:

D. van Keulen,
Daan.vankeulen@wur.nl

Citation:

Keulen, D. V., Kranenburg, W. M., & Hoitink, A. J. F. (2025). A new harmonic regression approach to interpret and predict estuarine salinity variation. *Journal of Geophysical Research: Oceans*, *130*, e2024JC022185. https://doi.org/10.1029/2024JC022185

Received 26 NOV 2024
Accepted 29 MAR 2025

Author Contributions:

Conceptualization: D. van Keulen, W. M. Kranenburg, A. J. F. Hoitink
Formal analysis: D. van Keulen
Funding acquisition: A. J. F. Hoitink
Investigation: D. van Keulen
Methodology: D. van Keulen
Project administration: A. J. F. Hoitink
Supervision: W. M. Kranenburg, A. J. F. Hoitink
Visualization: D. van Keulen
Writing – original draft: D. van Keulen
Writing – review & editing: W. M. Kranenburg, A. J. F. Hoitink

Abstract In this paper, we introduce a physics-inspired harmonic regression model to capture the nonstationary salinity dynamics at monitoring stations in well-mixed estuarine systems. Building on existing hybrid harmonic regression approaches, which modify the classical harmonic analysis to cope with nonstationary signals to predict tidal water levels, our model captures tidal and subtidal salinity variations using a simplified analytical salt intrusion model. The harmonic regression model was tested in the well-mixed Ems and Scheldt estuaries using data sets spanning 2–4 years, explaining 87.4%–96.4% of the observed salinity variance at upstream stations. A key finding is that storm surge effects typically have longer wavelengths than the estuary's length scale, which justifies using a linear relation between vertical and horizontal excursions. In alluvial estuaries, where the system widens, unsteadiness of the river discharge shows to be increasingly important for more downstream stations. The model quantifies the characteristic response time of salinity to variation in discharge. Based on a critical evaluation of the model equations, we offer a physical interpretation of the optimized parameters. Specifically, we discuss the Van der Burgh constant, which is an empirical coefficient commonly used in salt intrusion models. Our findings reveal that the Van der Burgh coefficient scales with the spatial scales of dispersion and advection, relative to changes in channel geometry.

Plain Language Summary In this paper, we introduce a tool to analyze nonstationary salinity levels at monitoring stations in estuaries. The model requires minimal input and allows for quick analysis and prediction. It builds on traditional methods for predicting tidal water levels, extending them to capture salinity variations caused by storm surges and variations in river discharge. We demonstrate the model's accuracy using data from the Ems and Scheldt estuaries. One key insight is that the model described the dominant salinity response to storm surges as a linear relationship between surge water levels and resulting currents, which is used to demonstrate the importance of storm surges for salt intrusion. Application of the model and interpretation of the results reveal how the response of salinity levels to discharge variations increases for stations near the coast, where the estuary is wider. Finally, through a critical analysis of the model equations, we provide new insights into the physical meaning of a commonly used empirical parameter. Our findings show how this parameter depends on the estuary's geometry and hydrodynamic properties, offering insight into how salt intrusion behaves in different estuaries.

1. Introduction

Estuaries are dynamic water bodies forming the transition zone between rivers and the coastal ocean, characterized by a region with brackish water resulting from mixing of fresh water and saline seawater. The longitudinal salinity distribution and the landward salt intrusion length change over time, resulting from the variable forcing at the boundaries. In this paper, we propose a new model to analyze estuarine salinity variations at monitoring stations.

Estuarine salinity levels are the result of a continuous competition between flushing by freshwater river discharge and inward transport by dispersive mechanisms. The dispersive salt flux comprises density-driven and tidal contributions (Geyer & Signell, 1992; Hansen & Rattray, 1966; Jay & Smith, 1990; MacCready & Geyer, 2010), which intrinsically involve two- and three-dimensional processes. The density-driven contribution results from a baroclinic pressure gradient along the river to sea continuum, which drives a vertically sheared exchange flow (Chatwin, 1976; Hansen & Rattray, 1965; Pritchard, 1952). This circulation can be strengthened by tidal straining (Burchard et al., 2011; Jay & Musiak, 1994, 1996), which can enhance stratification during ebb and suppress it

during flood. This asymmetry in stratification leads to enhanced vertical mixing during flood and suppressed mixing during ebb, which induces momentum exchange enhancing the vertical circulation. Tidal dispersion encompasses multiple mechanisms including tidal trapping, horizontal residual circulation, and tidal pumping (Fischer et al., 1979). Although the underlying processes are different, each of these mechanisms is the result of vertical and/or lateral velocity and salinity variations. The tide-driven dispersive mechanisms strongly depend on the local geometry, flow, and density, making it difficult to parameterize the resulting dispersion process (Geyer & Signell, 1992). Which mechanism is dominant strongly depends on the setting of the estuary, and often changes along the estuary.

Although the estuarine salinity variations involve processes that are intrinsically three-dimensional and that depend on the local geometry, various analytical formulations have been derived to describe and predict the cross-sectional averaged steady-state salinity distribution (Brockway et al., 2006; Chatwin, 1976; Gay & O'Donnell, 2007; Ippen & Harleman, 1961; Kuijper & Van Rijn, 2011; MacCready, 1999; Prandle, 1981; Savenije, 1986; Xu et al., 2019). All these analytical expressions greatly simplify the estuarine dynamics by reducing the dispersive mechanisms involved, assuming a simplified channel geometry and using calibration parameters related to the dispersion coefficients. Arguably, the best substantiated models are derived for estuaries with a constant cross-sectional area where density-driven circulation is the dominant exchange mechanism. For converging systems, Prandle (1981) was among the first to derive a predictive model without a predictive expression for the bulk dispersion coefficient at the mouth. A full set of predictive equations for the salinity balance in converging estuaries was proposed by Savenije (1989), and applied for a variety of systems (Cai et al., 2015; Gisen & Savenije, 2015; Nguyen & Savenije, 2006; Savenije, 1986, 1989, 1993; E. Zhang et al., 2011). The model uses the empirical Van der Burgh relation (Van de Burgh, 1972) to obtain an expression for the longitudinal dispersion coefficients and resulting salinity. Although the model proposed by Savenije (1989) has been successfully applied for a variety of systems, its physical validity has limitations. Similar empirical expressions that avoid the use of the Van der Burgh formulations have been derived by Kuijper and Van Rijn (2011) and Z. Zhang and Savenije (2019), which resulted in nearly identical relationships. Alternatively, Gay and O'Donnell (2007) derived a theoretically well-substantiated expression for the salinity distribution for linearly tapered systems, adopting a constant dispersion coefficient in segments of the estuary. Their model provides insight into the along-channel salinity distribution based on the physical parameters.

On an intertidal timescale, the tidal averaged salinity distribution is altered by the tides. Ippen and Harleman (1961) showed that for stratified systems experiencing a weak tidal forcing, the longitudinal salinity distribution migrates between the two periods of slack water over a distance approximately equal to the tidal excursion length, without significantly changing its shape. For vertically well-mixed systems, Savenije (1989) used his steady-state model to derive an expression for the high- and low-water salinity distribution, adopting a Lagrangian approach to describe the salinity variations over the excursion length. For these well-mixed systems, a longitudinal salinity distribution function that does not change its shape between the slacks can be devised. However, for many systems, the longitudinal distribution may change over the tide, for example, due to tidal asymmetries like tidal straining (Simpson et al., 1990). Analytical models for the unsteady salt distribution have been derived by Ippen and Harleman (1961) for stratified prismatic channels and Xu et al. (2019) for well-mixed converging systems. These analytical models do not account for tidal asymmetries. However, they clearly indicate how the intertidal salinity variations depend on the boundary conditions, tidal excursion length, and longitudinal coordinate.

Next to the tidal motion, storm surges also influence salt intrusion lengths. In literature, the important influence of storm surges on the along-estuary salinity distribution has received little attention, especially compared to their associated wind effects (Chen & Sanford, 2009; Jongbloed et al., 2022; Scully et al., 2005). Gong et al. (2018) studied storm wave-current interaction and effects on salt intrusion in the stratified Modaomen estuary during typhoons. They show that advective transport resulting from a barotropic pressure gradient prevailed along the different inlets during the surge events. These results agree with W. Kranenburg et al. (2022), who showed that during surge events in the partially mixed Rhine-Meuse estuary the advective salt flux reversed. The wave dynamics of storm surge waves in estuaries was explored by Familkhalili et al. (2020), who found that the dynamics are similar to those of tidal waves. Hence, tidal dynamics may be used to describe the salt dynamics resulting from storm surge waves. Most of the aforementioned studies focused on stratified systems where destratification is short-lived and predominately attributed to wind effects (Cho et al., 2012; M. Li et al., 2006, 2007), with restratification occurring within days.

The main aim of this paper is to develop a harmonic regression model that captures nonstationary salinity level time series influenced by weekly or seasonal discharge variations, as well as aperiodic variations caused by storm surges at a monitoring station. This objective is achieved through the following steps: (a) We extend the concept that estuarine salinity levels depend on the intertidal dynamics as addressed by Xu et al. (2019). To achieve this, we reformulate the approach used by Matte et al. (2013, 2014) to adjust the classical harmonic analysis to allow for predictions of nonstationary water levels (referred to as hybrid harmonic models (Hoitink et al., 2017)), and apply the result to estuarine salinity records. This involves extending the analytical salt intrusion model by Savenije (1986), which describes nontidal salinity variations with a hybrid harmonic model that captures the intertidal variability. (b) We explore the predictability of salt intrusion resulting from storm surges and vertical excursions of the estuarine water mass in well-mixed systems, and show how these dynamics can be integrated into this framework. (c) We modify the harmonic analysis approach to handle a clipped harmonic signal (periodically truncated tidal record) and focus on optimal constituent selection. (d) The model is then extended to account for the unsteadiness term in the salt balance equation.

The resulting model is applicable to analyze time series of yearly salinity variations that are long enough to distinguish between the dominant tidal constituents (i.e., the time series should typically cover at least 1 year). For these periods, the model helps to unravel the contribution of various physical forcings to site-specific salinity variations, and provides insight into how salt intrusion behaves. The mechanistic salt transport processes are modeled using a parameterized approach, where model coefficients offer a synoptic understanding of the system's functioning. The salinity time series may include prolonged periods of low salt levels, representing river salinity, as well as occasional periods when the salt front reaches the station, although this limits the identification of tidal constituents. Finally, the model allows for short- to medium-term predictions (days–weeks in advance), which is the typical period for which predictions of discharge and water levels at the seaward boundary are available.

The structure of the remainder of this paper is as follows. Section 2 outlines the theoretical background of the harmonic regression model, including the steady-state salinity distribution as described by Savenije (1986), and discusses storm surge effects. Section 3 focuses on the harmonic regression model itself. In Section 3.1, the regression equations are derived, accounting for subtidal, intertidal, storm surge, and unsteadiness effects. Section 3.2 provides a summary of the regression and optimization procedures. In Section 4, a concise description of the Ems and Scheldt estuary systems is provided, along with a discussion of the observations used. Subsequently, the steady-state harmonic regression model is applied, followed by a sensitivity analysis of the results to the model parameters. Section 5 assesses the influence of the down-estuary boundary condition and unsteadiness effects, and tests the model formulation that accounts for unsteadiness. A critical evaluation of the adopted model is presented in Section 6, where the formulation of Savenije (1986) is compared with that of Gay and O'Donnell (2007) to discover the physical meaning of the optimized model powers. Section 7 discusses the broader implications of the model results, and conclusions drawn in Section 8.

2. Theoretical Background

2.1. The One-Dimensional Salt Balance

Temporal variations of estuarine salinity are generally the result of flushing by the river discharge, longitudinal exchange mechanisms, and dispersion of salt through tidal dynamics. In a one-dimensional perspective, temporal salinity variations are the result of advection of the salinity gradient and dispersion through a Fickian dispersion coefficient. The latter describes the down-gradient salt flux resulting from exchange flows and/or tidal dispersion. The one-dimensional cross-sectionally and tidally averaged salt balance reads as follows:

$$A_x\frac{\partial S_x}{\partial t}+\frac{\partial (QS_x)}{\partial x}-\frac{\partial}{\partial x}\left[A_xD_x\frac{\partial S_x}{\partial x}\right]=0, \tag{1}$$

with S_x being the cross-sectional averaged salinity, A_x the average cross-sectional area, x the longitudinal coordinate, Q the discharge, and D_x the longitudinal dispersion coefficient. If the subtidal salinity distribution is in a quasi-steady state, it may be described by the balance between advection and dispersion as follows:

$$Q_rS_x-A_xD_x\frac{\partial S_x}{\partial x}=0. \tag{2}$$

Equation 2 describes the subtidal salt balance best when the discharge has remained constant long enough for the salt front to adjust (Kuijper & Van Rijn, 2011; Monismith et al., 2002) and requires that there is a zero net salt flux trough the estuary.

2.2. Salt Intrusion Model

The harmonic regression model requires a shape function for the longitudinal salinity distribution, for which the empirical steady-state formulation of Savenije (1986) is adopted. To obtain an expression for the longitudinal profiles of salinity, Savenije (1986) employs an empirical relation for the longitudinal gradient of the dispersion coefficient (Van de Burgh, 1972), which reads as follows:

$$\frac{dD_x}{dx} = -K\frac{|Q_r|}{A_x}, \tag{3}$$

where K is a dimensionless coefficient. Substituting Equation 2 into Equation 3 and integrating over x yield a relationship between the salinity and dispersion coefficients, as in

$$\frac{S_x - S_r}{S_0 - S_r} = \left(\frac{D_x}{D_0}\right)^{\frac{1}{K}}, \tag{4}$$

where D_0 is the dispersion at the mouth, S_x is the local salinity, S_0 is the salinity at the downstream boundary, and S_r is the river salinity. An equation for the longitudinal dispersion coefficient is obtained by integrating Equation 3. Depending on whether the cross-sectional area is assumed to be constant or converging in the landward direction, an equation for the longitudinal dispersion coefficient is obtained for a prismatic or a converging channel. Substituting the relation for an exponentially decaying cross-section into Equation 4 yields

$$\frac{S_x - S_r}{S_0 - S_r} = \left(1 - \frac{KL_a|Q_r|}{A_0 D_0}\left(e^{\frac{x}{La}} - 1\right)\right)^{\frac{1}{K}}, \tag{5}$$

where A_0 is the cross-sectional area at the mouth and L_a is the convergence length of the cross-sectional area. In Equation 5, the parameters D_0 and K are used to tune the salinity curve and their influence is illustrated in Figure S1 in Supporting Information S1.

Focusing on the dispersion coefficient at the mouth, various predictors have been proposed (Gisen & Savenije, 2015; Kuijper & Van Rijn, 2011; Savenije, 1993). Gisen and Savenije (2015) derived a predictive equation from bulk estuarine parameters, based on reanalysis of the dimensionless parameters used by Savenije (1993). Using data from 30 different estuaries, the following expression for the nondimensional dispersion coefficient was found:

$$\frac{D_0}{v_0 E_0} = 0.40 N_{r,0}^{0.57}\left(\frac{g}{C^2}\right)^{0.21}, \tag{6}$$

with v_0 the tidal velocity amplitude, E_0 the tidal excursion length, $N_{r,0}$ the horizontal estuarine Richardson number, g the gravitational acceleration, and C the Chézy coefficient. The subscript 0 indicates that the parameters have been evaluated at the downstream boundary. The horizontal estuarine Richardson number, introduced by Fischer (1972), expresses the balance between the potential energy required for full mixing against buoyancy and the available kinetic energy from the tidal currents, to accomplish mixing, and reads as follows:

$$N_{r,0} = \pi\frac{\Delta\rho_0}{\rho}\frac{gh_0}{v_0^3}\frac{Q_r}{A_0}, \tag{7}$$

with h_0 the average depth, $\Delta\rho_0$ the horizontal density difference over the intrusion length, and ρ the density of fresh water. The estuarine Richardson number can be used to quantify the degree of stratification for an estuary (Fischer, 1976).

2.3. Storm Surges and Their Effects

Advective transport (i.e., excursion of the salinity curve) is assumed to be the dominant mechanism of storm surges for well-mixed systems. Tidal theory can be adopted to estimate the relation between surge water levels and the resulting excursion length, because storm surges have typical periods close to those of tides and both are long waves (Familkhalili et al., 2020).

A relation between the water level variations, estuarine shape, and the resulting vertical excursion length for funnel-shaped and long estuaries is given by the geometry-tide relationship (Savenije, 2006; Toffolon et al., 2006):

$$E = H\left(\frac{r_s L_a}{h}\right)\left(\frac{\cos\varepsilon}{1-\delta_u L_a}\right), \tag{8}$$

with E the tidal excursion length, H the tidal range, L_a the convergence length of the cross-sectional area, δ_u the damping number of the horizontal tide, ε the phase lag between high water and high-water slack (describing the wave character), and r_s the dimensionless ratio of the storage width to the stream width. The wave has a standing character when ε equals 0, a mixed character when $0 < \varepsilon < \frac{\pi}{2}$, and is progressive when ε is $\frac{\pi}{2}$. Toffolon et al. (2006) derived explicit formulations for ε and δ_u for mixed and standing waves, expressed in terms of the estuary number $\gamma = \left(\frac{c_0}{\omega L_a}\right)$ and the friction number $X = \left(\frac{g}{C^2}\frac{c_0}{h\omega}\frac{\eta}{h}\right)$, where c_0 is the classical wave celerity, ω is the angular frequency ($\omega = \frac{2\pi}{T}$), η is the tidal amplitude, and T is the tidal period. Based on these equations, the wave character of storm surges and the resulting excursion length may be estimated when the amplitude and duration are known. Figure S2 in Supporting Information S1 visualizes the dimensionless phase lag and the damping number in the X-γ plane. As γ increases, the phase lag ε becomes independent of the estuary and the friction number. Similarly, δ becomes independent of the friction number, and the damping is solely determined by the shape of the estuary and the surge wavelength. When the typical wavelengths of storm surge events are larger than the length scale of the estuary, the relation between the vertical and horizontal excursion may be approximated by the classical relation $E \approx H\left(\frac{r_s L_a}{h}\right)$. Avoiding a dependence on ε and δ_u has the advantage that a (near) unique relation exists between storm surge water levels and the resulting horizontal excursion over a range of frequencies.

3. Regression Model for Salinity Time Series

This section proposes a harmonic regression model for estuarine salinity. First, Section 3.1 describes the modified version of Savenije (1986) and the extensions to capture tidal and surge variations. Second, Section 3.2 describes the implementation and optimization procedure.

3.1. The Regression Equations

3.1.1. Tidal Averaged Salinity Variations

For width-converging systems, the tide averaged (TA) salinity distribution can be described by Equation 5. To describe the salinity variations at a monitoring station, indicated by the subscript s, Equation 5 is simplified by lumping all time-independent variables in a single term as follows:

$$\left(\frac{S_s - S_r}{S_0 - S_r}\right)^K = \frac{D_s}{D_0} = 1 - a_0\frac{K|Q_r|}{D_0}. \tag{9}$$

The coefficient a_0 describes the degree in which the relative dispersion coefficient is reduced at x_s. Note that K is maintained in Equation 9 because it appears in the power of the RHS of Equation 9. The tide-averaged salinity variations at the location of the monitoring station are described by Q_r and D_0.

3.1.2. Intertidal Salinity Variations

Savenije (1989) approximated the TA and low-water slack (LWS) salinity distribution by assuming that the high-water slack distribution migrates over the horizontal excursion length (E). Following this reasoning, the salinity

variations observed over a tidal cycle, relative to the tide-averaged salinity, approximate the salinity observed at $x_t = x_s + \frac{E}{2}\cos(\omega t + \phi)$, where t is time, and ω and ϕ are the angular frequency and phase of the dominant tidal constituent, respectively. To obtain an expression for the intertidal salinity variations and parameters that can be determined through regression, x in Equation 5 is substituted by x_t. We simplify the equation to separate the tide averaged from the intratidal variations, which yields the following:

$$\left(\frac{S_s - S_r}{S_0 - S_r}\right)^K = \frac{D_s}{D_0} \approx 1 - K\frac{|Q_r|}{D_0}\frac{L_a}{A_0}\left(e^{\frac{x_s}{L_a}} - 1\right) + \left(\frac{1}{2}K\frac{|Q_r|}{D_0 A_0}e^{\frac{x_s}{L_a}}\right)E\cos(\omega t + \phi). \tag{10}$$

The third term on the RHS of Equation 10 is obtained by subtracting the TA expression from the modified version, including the tidal excursion. Subsequently, a first-order Taylor approximation for $E\left(e^{(\delta)} \approx 1 + \delta\right)$ around 0 is used in the natural exponent, to separate the tide averaged from the tidal variation. By lumping all time-independent variables in Equation 10, rewriting the harmonic part in terms of tidal constituents yields the following:

$$\left(\frac{S_s - S_r}{S_0 - S_r}\right)^K = 1 - a_0 K\frac{|Q_r|}{D_0} + \sum_{k=1}^{n}\left[a^c_{1,k}K\frac{|Q_r|}{D_0}\cos(\omega_k t) + a^s_{1,k}K\frac{|Q_r|}{D_0}\sin(\omega_k t)\right]. \tag{11}$$

In Equation 11, the amplitudes $a^c_{1,k}$ and $a^s_{1,k}$ determine the phase and amplitude of each tidal constituent k. Note that based on the Taylor approximation, an error is made, which is small when $\frac{E}{La} < 1$. Equation 11 is more accurate when the channel geometry is closer to being prismatic. For prismatic channels, the equation is exact.

3.1.3. Storm Surge Variations

Sections 2.3 discussed how storm surge waves, under the assumption that tidal theory applies to storm surges, depend on the dimensionless friction number and the estuary shape number. Because storm surge waves do not act on a predefined frequency and amplitude, some simplifications are required. Here, we assume that storm surges act as standing waves, which applies when the typical wavelengths are larger than the length scale of the estuary. For a standing wave, high and low water is reached everywhere at the same time (i.e., no phase lag and $\varepsilon = 0$). Damping becomes independent of friction, causing that damping coefficients vary little between different periods. Under these conditions, the surge-related excursion lengths may be approximated by $E = H\frac{L_a r_s}{h}$. The salinity variation introduced by storm surges, under the assumption of a linear relation between water levels and currents, are implemented similarly as the intertidal variations as follows:

$$\begin{aligned}\left(\frac{S_s - S_r}{S_0 - S_r}\right)^K &= 1 - \underbrace{a_0 K\frac{|Q_r|}{D_0}}_{\text{TA}} + \underbrace{a_1 K\frac{|Q_r|\delta_w}{D_0}}_{\text{SURGE}} \\ &+ \underbrace{\sum_{k=1}^{n}\left[\left(a^c_{2,k}K\frac{|Q_r|}{D_0} + a^c_{3,k}K\frac{|Q_r|\delta_w}{D_0}\right)\cos(\omega_k t) + \left(a^s_{2,k}K\frac{|Q_r|}{D_0} + a^s_{3,k}K\frac{|Q_r|\delta_w}{D_0}\right)\sin(\omega_k t)\right]}_{\text{TIDAL}}.\end{aligned} \tag{12}$$

In Equation 12, δ_w denotes the surge water level. In addition, the surge-related water level fluctuation amplifies or attenuates the amplitudes of the intertidal salinity variation. Hence, the surge-induced salinity variation also appears in the harmonic model.

3.1.4. Accounting for Unsteadiness

Equations 11 and 12 rely on that the system can be adequately described by a steady-state model. However, in many systems, variations in river discharge can occur over periods shorter than the time required for the salinity field to fully adjust. To account for this, empirical models for the salt intrusion length often incorporate these effects trough an autocorrelation process (Monismith et al., 2002; Reyes-Merlo et al., 2013). This approach is also adopted here. Specifically, Equations 11 and 12 can be extended with an autoregressive AR(1) term, which for Equation 12 yields

$$\left(\frac{S_s - S_r}{S_0 - S_r}\right)^K = 1 - \underbrace{a_0 K \frac{|Q_r|}{D_0}}_{\text{TA}} + \underbrace{a_1 K \frac{|Q_r|\delta_w}{D_0}}_{\text{SURGE}} + \underbrace{a_2 \tilde{D}_t}_{\text{LAG}} + \underbrace{\sum_{k=1}^{n}\left[\left(a^c_{3,k} K \frac{|Q_r|}{D_0} + a^s_{4,k} K \frac{|Q_r|\delta_w}{D_0} + a^c_{5,k}\tilde{D}_t\right)\cos(\omega_k t) + \left(a^s_{3,k} K \frac{|Q_r|}{D_0} + a^s_{4,k} K \frac{|Q_r|\delta_w}{D_0} + a^s_{5,k}\tilde{D}_t\right)\sin(\omega_k t)\right]}_{\text{TIDAL}}, \tag{13}$$

where $\tilde{D}_t = \left(\frac{S_s - S_r}{S_0 - S_r}\right)^K - 1\Big|_{t-\Delta t}$, and where Δt is set to 1 day. Furthermore, Equation 13 implicitly assumes a time step Δt, which is lumped into the regression parameters coefficients. An empirical response time, T_R, can be defined as $T_R = 1/(1 - a_1)$. Typically, this response time varies between high- and low-discharge conditions. However, in this study, a single-response time is used. Although Equation 13 allows for more accurate predictions when unsteadiness is significant, this approach does not handle truncated salinity signals effectively.

3.1.5. Modelling the Dispersion Coefficient at the Mouth

The dispersion coefficient at the downstream boundary D_0 is not yet specified in Equations 9–13. A modified version of Equation 6 is here proposed. First, based on Equation 6, it can be seen that D_0 is proportional to $N_{r,0}$, as in the power law $\frac{D_0}{v_0 E_0} = \alpha N^P_{r,0}$, where α is a constant. In the regression model we devise, the power P is allowed to deviate from the empirically derived value of 0.57 to optimize the model. Second, assessing $N_{r,0}$ requires velocity time series. Generally, only records of tidal elevation are available. Therefore, the horizontal tidal amplitude at the downstream boundary (v_0) is replaced by an appropriate velocity scale $v = \eta\omega\frac{L_a}{h}$ and the tidal excursion length is written as $E = \frac{vT}{\pi}$. For ω, the typical tidal frequency of the system is used. Substituting these relations into Equation 6 and rearranging the variables related to $N_{r,0}$ into a new dimensionless group yields the following:

$$D_0 \propto \left(\underbrace{\frac{h_0^3 \pi}{L_a^3 B_0}\left(\frac{g}{\omega^2}\right)}_{\text{constant}}\right)^P \left(\underbrace{\frac{\Delta\rho_0}{\rho}\frac{Q_r}{\eta_0^3\omega}}_{\text{varying}}\right)^P \underbrace{\left(\frac{L_a}{h_0}\right)^2}_{\text{constant}} \underbrace{\left(\eta_0^2\omega\right)}_{\text{varying}}. \tag{14}$$

In Equation 14, two constant and two variable terms are indicated. The constant terms in Equation 14 can be absorbed in the coefficients α_i of Equations 9–13, and are resolved based on regression.

3.1.6. Salinity at the Boundaries

The relative salinity in Equations 9–13 depends on the salinity input from the river, which generally varies with the discharge. These variations are smaller (i.e., within 0–1 psu), but are important to distinguish when the salinity is influenced by the oceanic boundary or is solely determined by the river (also see Section A1). In order to estimate the river salinity for the discharge, the following model is adopted:

$$S_r = \frac{C_b}{Q_r} + S_c, \tag{15}$$

where S_c is the salinity concentration of the discharge and C_b a background load that is diluted by the discharge. The values for S_c and C_b are determined by regression according to Equation 15, using observed river salinity.

Similarly, the salinity at the used down-estuary boundary S_0 can either be considered constant. Alternatively, measured values of S_0 can be used, or a parameterization which proceeds from a simplified version of Equation 13:

$$\left(\frac{S_0 - S_r}{S_{oc} - S_r}\right)^K = 1 - a_0 K \frac{|Q_r|}{D_0} + a_1 \tilde{D}_t, \tag{16}$$

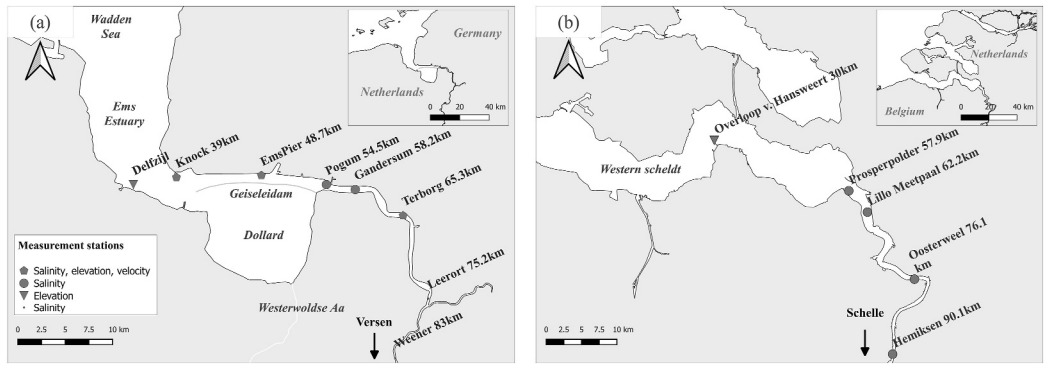

Figure 1. (a) Map of the Ems estuary and locations of the measurement stations. The river kilometers are referenced against the barrier island Borgum, located between the Wadden Sea and the North Sea. (b) Same as (a) but for the Scheldt estuary. The river kilometers are referenced against the city of Vlissingen.

where S_{oc} is the salinity at sea. The advantage of Equation 16 is that it allows to partially account for unsteadiness effect through the parameterized boundary condition such as used in Equations 9–12, which can deal with truncated salinity signals.

3.2. The Regression and Optimization Procedure

Estuarine salinity records often contain tidally periodic or prolonged periods of low salinity, where the salinity matches that of the river water (this introduces clipped harmonic records). To facilitate the optimization procedure, a priori assessment is conducted to isolate data windows where the salinity does not significantly exceed river salinity, excluding those windows from the optimization. The optimization procedure optimizes the regression coefficients and powers in Equations 9–12 through an iterative process that extends NS-TIDE (Matte et al., 2013, 2014). Furthermore, an adjusted version of the twofold strategy used by Matte et al. (2013) is applied for constituent selection and error estimates. However, for irregularly sampled data—introduced by excluding data points where salinity does not exceed river salinity—the degree to which tidal constituents can be properly resolved is influenced. Following Foreman et al. (2009), the correlation coefficients of the significance-tested constituents are monitored, discussed in Section 4.4.2. Details regarding constituent selection, optimization, and significance testing can be found in Appendix A.

4. Application to Two Estuaries

4.1. The Ems Estuary and Scheldt Estuary

The Ems estuary is located at the border between the Netherlands and Germany (see Figure 1a). It contains a large basin that consists of ∼80% tidal flats referred to as the Dollard. Exchange between the Dollard and the main channel of the Ems is limited by a permeable, longitudinal training dam (the Geise dam). Tides in the Ems estuary have a semidiurnal character with a tidal range of about 2.5 m near the entrance, but significantly amplify upstream of the Dollard (Winterwerp et al., 2017). The tidal wave upstream of the Dollard has a standing wave character. A storm surge barrier is located upstream of the city of Pogum and is occasionally closed. Fresh water input predominantly originates from the Ems River, which has an average discharge of ∼125 m^3/s. The Westerwoldse Aa constitutes another source of fresh water input, which drains ∼12.5 m^3/s into the Dollard basin.

The Scheldt estuary runs from Belgium to the south of the Netherlands (Figure 1b). The Dutch part of the Scheldt (the Western Scheldt) is wider and consists of multiple flood-ebb channel loops (Nguyen et al., 2008). In the Belgium part, the channels converge into a single channel that is kept at depth for the port of Antwerp. The mean tidal range is about 3.8 m near the entrance of the system and amplifies in the upstream direction. The wave propagation reflects that of an apparent standing wave (Savenije et al., 2008). The major contributors to the Scheldt are the Dender and the Rupel (Struyf et al., 2004); the average discharge of the Scheldt is estimated to be ∼85 m^3/s.

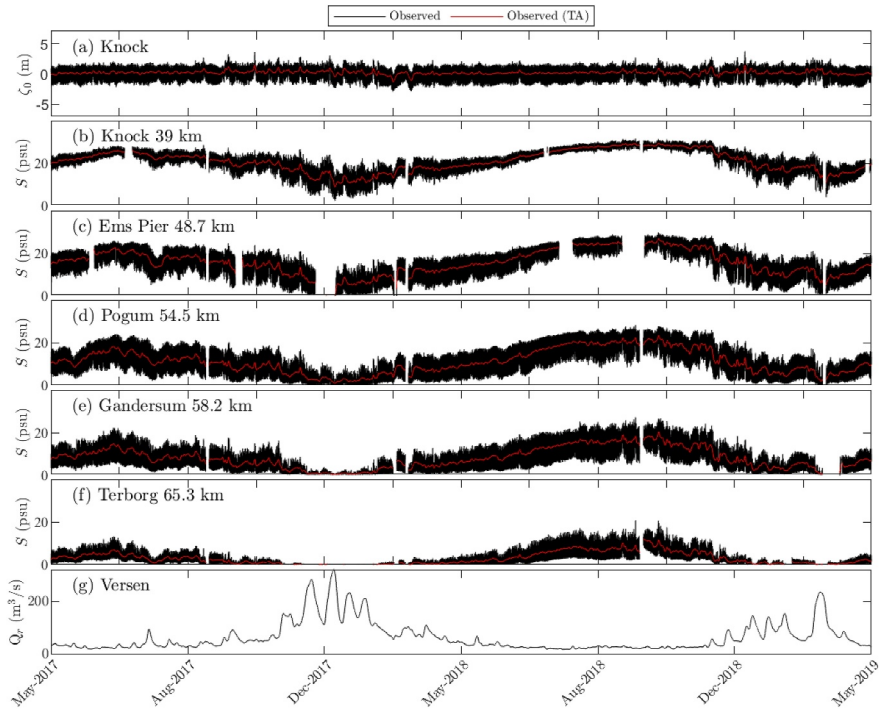

Figure 2. Overview of the observed salinity time series in the Ems estuary and the forcing at the up- and downstream boundaries. (a) Observed and Godin-filtered water levels at Knock. (b–f) Observed and tide-averaged salinity at the individual stations. Distance is relative to the barrier island of Borkum. (g) River discharge measured at Versen.

4.2. Data and Observations

To evaluate the proposed regression model, time series data sets from the Ems estuary (2 years) and the Scheldt estuary (4 years) are used. These data sets include information on salinity, water levels, velocity, and discharge (locations are shown in Figure 1). Here, the data are used to estimate the importance of the unsteadiness term at different stations and to assess the extent to which a steady-state approach is justified. Furthermore, the along-channel variations in geometry, salinity, and of the dispersion coefficients are discussed, and salinity variations due to surges are examined.

Figures 2 and 3 show the measured salinity levels and forcing variables. Both estuaries exhibit strong seasonality, with the lowest salinity during winter. In the Ems estuary, downstream of Pogum, salinity is affected by partial closures of the storm surge barrier, such as the closure on 7–9 October, which elevated salinity for about 2–3 weeks (Figure 2). Figure 4 shows the geometry, the typical salinity distributions, and associated dispersion coefficients for different discharge conditions from the measurements in Figures 2 and 3. The geometry, in both systems can be described by an exponential decay (Figures 4a and 4b), with a break in the channel convergence observed in both systems (most pronounced in the Scheldt). Figures 4c and 4d shows the average salinity distributions for different discharge conditions, which are bin averages of spline-interpolated salinity values at monitoring stations. The longitudinal salinity distribution in the Ems is characterized by a hyperbolic shape, whereas that of the Scheldt estuary changes from concave to a convex shape with increasing discharge. Figures 4e and 4f shows the corresponding steady-state dispersion coefficients, estimated using Equation 2. The variations associated with these values are quite large and therefore are only used qualitatively. For the Ems estuary, the resulting dispersion coefficients show a (local) amplification with increasing discharge, which corresponds to estimates from Helder and Ruardij (1982) and De Swart et al. (1997). In the Scheldt estuary, the dispersion coefficient remains nearly constant in the downstream reach, but amplify slightly in the upstream region.

To assess the importance of the unsteadiness term $\left(\frac{\partial S}{\partial t}\right)$ relative to the advective term $\left(u\frac{\partial S}{\partial x}\right)$ in the salt balance at various monitoring sites, the ratio of the unsteadiness term to the advective term is evaluated (Figures 5a–5h). In both estuaries, unsteadiness is most pronounced at the down-estuary stations. Figures 5g and 5h evaluate the (potential) errors in dispersion coefficients when the unsteadiness term is neglected, highlighting persistent errors

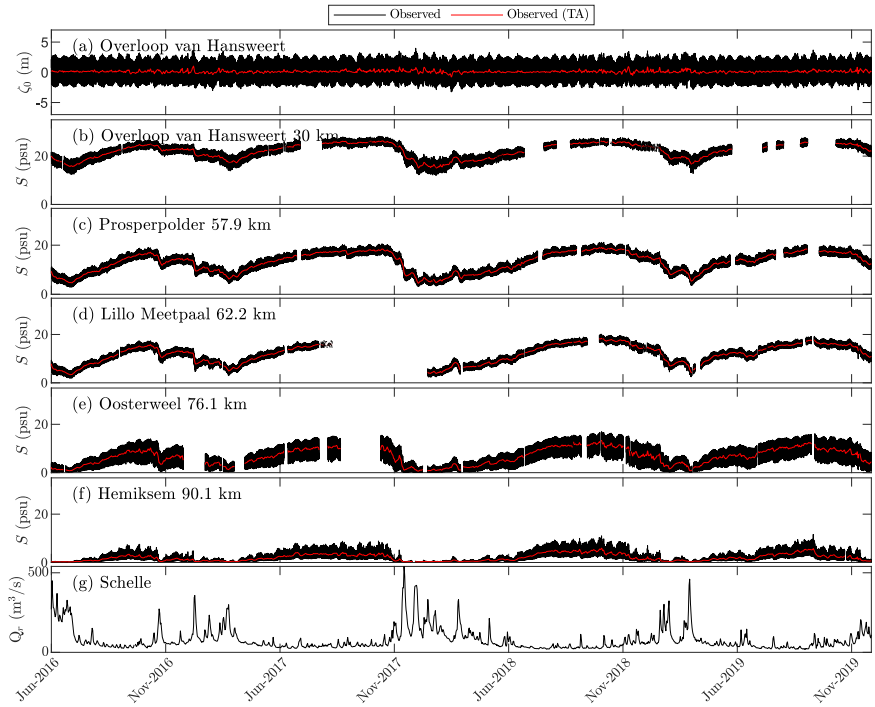

Figure 3. Overview of the observed salinity time series in the Scheldt estuary and the forcing at the up- and downstream boundaries. (a) Observed and Godin-filtered water levels at the Overloop van Hansweert station. (b–f) Observed and tide-averaged salinity at the individual stations. Distance is relative to the city of Vlissingen. (g) River discharge calculated at the Schelle station.

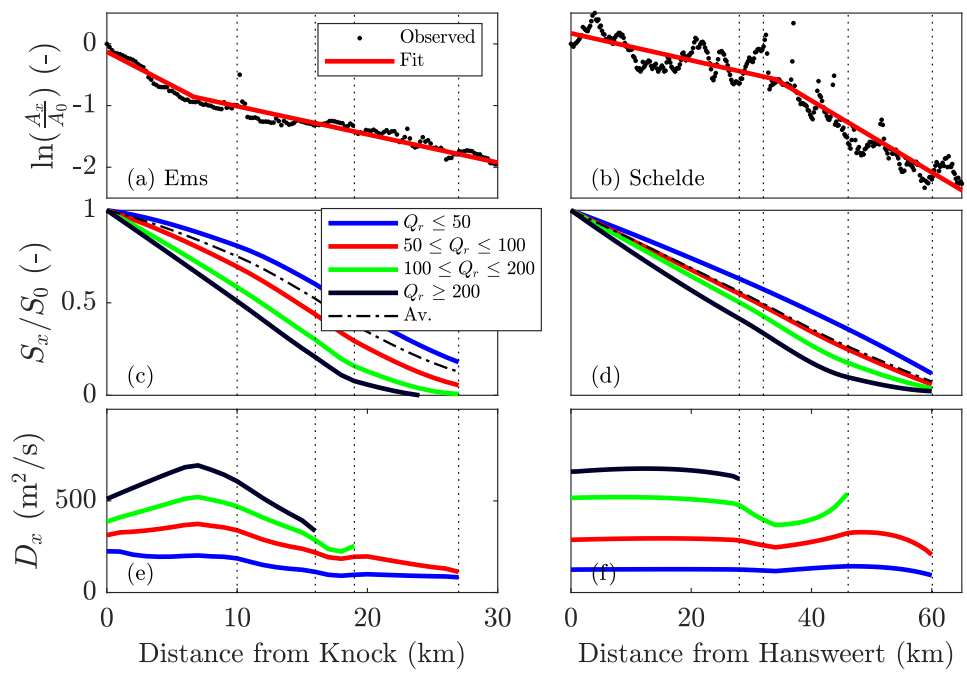

Figure 4. Overview of the geometry, typical along-channel salinity distributions, and estimated dispersion coefficients for the Ems and Scheldt estuaries. (a and b) Cross-sectional area and fitted exponential decay model $\left(A_x = A_0 e^{\frac{x}{L_a}}\right)$. (c and d) Typical along-channel salinity distributions under different discharge conditions. (e and f) Steady-state dispersion coefficients corresponding to the shown along-channel salinity distributions. Dashed vertical lines indicate the location of the salinity stations.

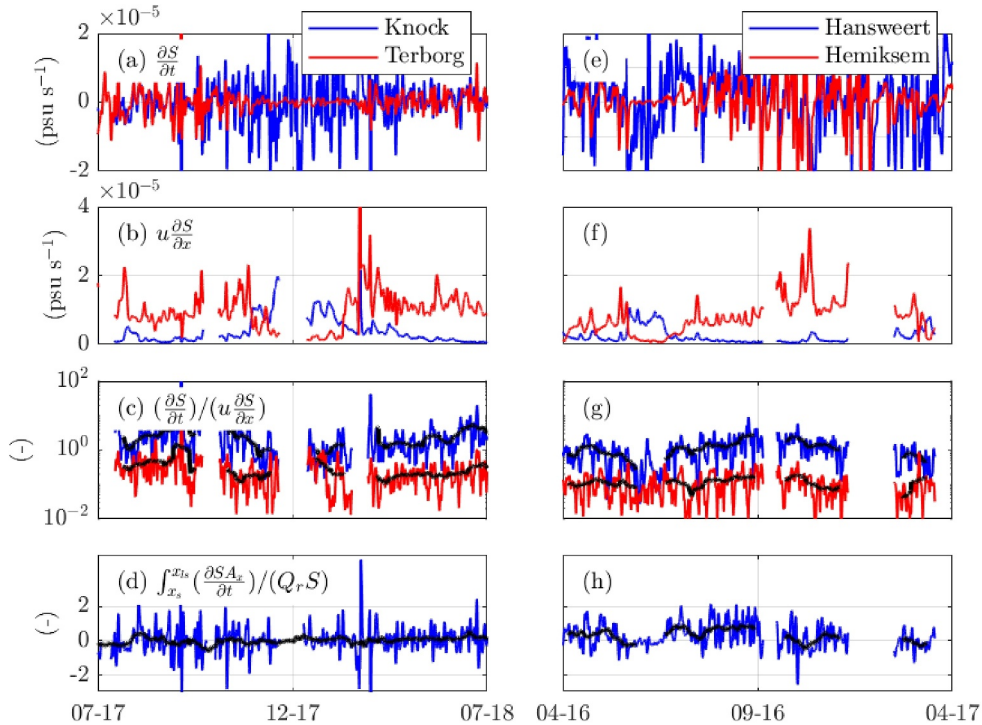

Figure 5. Evaluation of the importance of the (neglected) unsteadiness term. (a) Comparison of the unsteadiness term estimated from the tidally averaged salinity signal at the most down- and upstream stations in the Ems estuary. (b) Same estimation for the advection of salt. (c) Ratio of the unsteadiness term to the advection term, used to measure the importance of the unsteadiness term in the salt balance equation (Equation 1). The solid black lines indicate the monthly average. (d) Estimation of the error introduced by neglecting the unsteadiness term in the dispersion coefficient. Note: When using Equation 2, the dispersion coefficient is calculated as $D_x = \frac{Q}{A_x}\frac{S}{\partial S/\partial x}$. When using Equation 1, it is given by $D_x = \left(\int_{x_s}^{x_l}\frac{\partial S}{\partial t}A_x dx + Q_r S\right)\frac{1}{A_x \partial S/\partial x}$, where x_s is the location of the station and x_{ls} is the location where the salinity approaches zero; the most upstream station is used here. (e–h) Same as (a–d), but for the Scheldt estuary.

in the Scheldt due to the influence of unsteadiness. Thus, while the unsteadiness term is larger in the Ems, its sign is more persistent in the Scheldt, indicating a longer adjustment response time. Given these findings, a steady-state approach is likely more suitable for upstream stations and may introduce systematic errors for the down-estuary station. Furthermore, the admittance of the semidiurnal vertical tide (i.e., water level elevations), determined through a continuous wavelet transform, was tested for the Ems and Scheldt estuaries, but was found to show little temporal variations and is assumed stationary.

An analysis of water level, salinity, and discharge variations occurring in the 1.2–6-day range is performed to investigate salinity response during surge events. The lower limit of 1.2 days is used to distinguish the tidal signal from the surge variations. However, storm surges can act in the same frequency range as tides. The upper limit was set at 6 days, as choosing higher values caused that the filtered signal started to show variations that did not resembled the water level variations. The filtered salinity variations covary with the filtered water levels (Figures 6a and 6c). The water level variations are, however, not directly proportional to the observed salinity variations. This is because the resulting salinity variations depend both on the local, instantaneous salinity gradient and the salt water excursion resulting from storm surges. Dividing subtidal salinity variations by the local salinity gradient yields a length scale for the surge-induced excursion (Figures 6b and 6d). The relation between surge water levels and the surge-induced excursions shows a near-linear relation for both systems. This implies that the surge water level variations that occur over the 1.2–6-day range are directly proportional to the horizontal excursion of the estuarine water mass (Section 2.3).

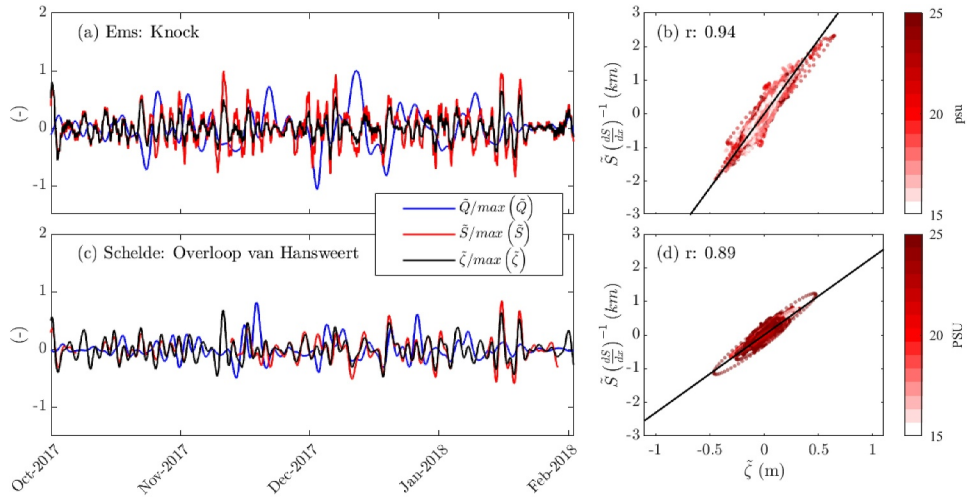

Figure 6. Relationship between storm surge water levels and horizontal excursion for the Ems and Scheldt estuaries. (a and c) Observed band-pass filtered variation of salinity $\tilde{S}$, water elevation $\tilde{\zeta}$, and discharge $\tilde{Q}_r$, normalized by the maximal observed value. The band-pass filter retains variation in between 1.2 and 6 days. (b and d) Linear relationship between water elevations $\tilde{\zeta}$ and estimated horizontal excursions. The latter is approximated by dividing $\tilde{S}$ by the local salinity gradient $\frac{dS}{dx}$.

4.3. Performance of the Regression Model

This section tests the predictive skill of the steady-state formulations (Equations 11 and 12) in predicting salinity variations at the stations where salinity often matches river levels and unsteadiness effects between the station and the downstream boundary are assumed minimal.

The model inputs consist of river discharge, river salinity, and tidal elevation. Furthermore, observed salinity at a down-estuary station, where unsteadiness effects become important, are provided for the downstream boundary condition. The data sets are divided into a calibration period and a prediction period. For the Ems, the first 15 months are used to optimize model coefficients, with the remaining 7 months testing predictive skill. Similarly, for the Scheldt, the last year of data is used for prediction testing. The low-pass filtered salinity signals at Knock and Prosperpolder provide the boundary conditions, S_0, for the Ems and Scheldt estuaries, respectively.

In the Ems estuary, the best results were achieved using Equation 12 (Figure 7). This model, fitted with 40–32 regression coefficients for the various stations (used tidal constituents are listed in Table S1 in Supporting Information S1), had an unexplained variance of 4.9%–11.1% and an RMSE of 1.2–1.8 psu during calibration (Table 1). The unexplained variance increased slightly to 4.8%–11.4% with an RMSE of 0.9–1.6 psu during the predictions, with the greatest decrease in predictive skill observed at the most upstream station. Figure 9 provides additional details for the Gandersum station. Figure 9a shows that the model effectively captures the harmonic variations in intertidal salinity, supported by the power spectra comparison (Figure 9d). It also accurately represents the period where the salinity signal was clipped (Figure 9c) and two significant surge events with sea level increases of 0.96 and 1.45 m (Figure 9b). When surge variations are not included (Equation 11), the unexplained variance increased by approximately 1%–2% during the prediction period (Figure S3 in Supporting Information S1).

Similar results were obtained for the Scheldt estuary (Figure 8 and Figure S4 in Supporting Information S1). Using 24–58 regression coefficients, depending on the station, the model using Equation 12 achieved an unexplained variance of 1.4%–8.8% and an RMSE of 0.5–0.9 psu during calibration (Table 2). Prediction variance increased to 3.6–12.0% with an RMSE of 0.6–1.1 psu. Compared to the Ems case, excluding storm surge effects had a smaller impact on the regression model's performance in the Scheldt estuary, likely due to its longer salt intrusion length and weaker sensitivity to storm surge water levels (see Figure 6).

Figures 7 and 8 show that the low-pass filtered variations closely align with observed averages, indicating that the mean signal is well captured and suggesting minimal unsteadiness effects upstream of the imposed boundary. However, residuals still exhibit autocorrelation. To assess the importance of the neglected unsteadiness term, it

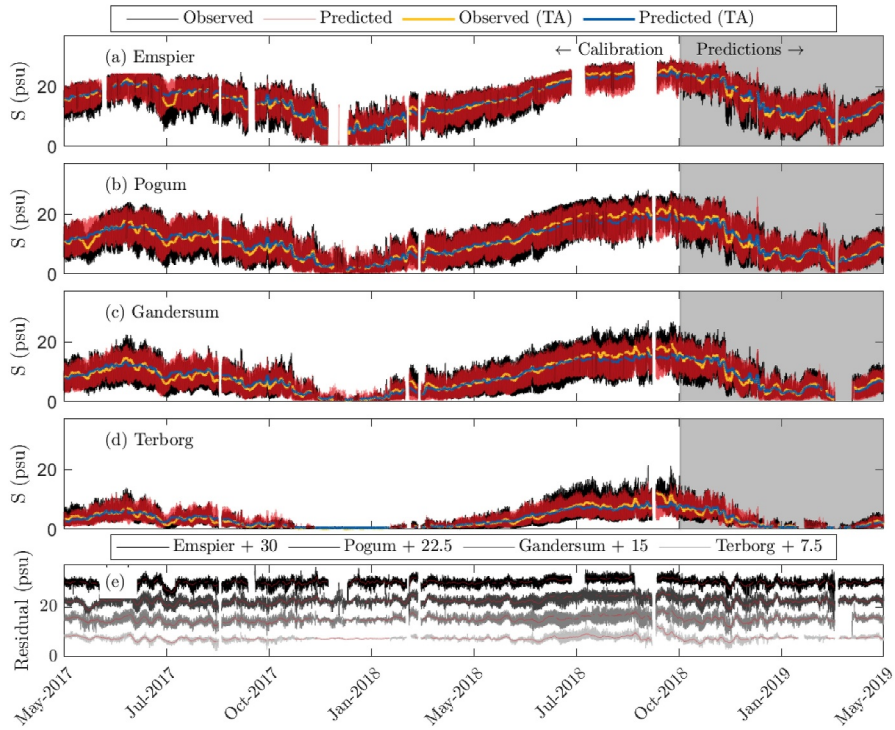

Figure 7. Data-model comparison for salinity in the Ems estuary, with salinity predicted using Equation 12 and the fortnightly averaged salinity at Knock as seaward boundary condition. (a–d) Observed and predicted salinity variations and tide averaged (determined by a moving average, 1-day window). The period for which the salinity is predicted is highlighted by a gray background. (e) Residuals (predictions minus observations) of panel (a–d). The residuals include an offset for clarity of the figure.

was assumed that the low-pass filtered residuals reflect the mismatch due to unsteadiness. This accounted for 3.2%–3.8% of the variance in the Ems, with little variation between stations, while in the Scheldt, this percentage increased upstream, ranging from 0.4% to 4.8%.

4.4. Model and Parameter Sensitivity

4.4.1. Optimized Powers and Fitted Coefficients

Section 4.3 demonstrated that accurate predictions can be achieved using Equations 11 and 12. In this section, we assess the robustness and consistency of the model, focusing on the optimized power coefficients and fitted regression parameters.

First, three experiments were conducted for both systems using Equation 12: one where both power coefficients were optimized and two where either P or K was held constant. These experiments aim to determine whether optimizing both powers is necessary. Second, for the model with both power coefficients optimized, four additional experiments were performed to evaluate the consistency of the best-fit values of the power and regression coefficients when using different periods for model optimization and different discharge conditions. The best-fit model coefficients are shown in Figures 10 and 11 for the Ems estuary and the Scheldt estuary, respectively.

In the first experiment, both P and K were optimized (Figures 10a–10c and 11a–11c). For the Ems estuary, the best-fit values for P increased from 0.56 ± 0.01 to 0.67 ± 0.02. The obtained values for $1/K$ varied significantly between stations, increasing from 1.26 ± 0.07 to 4.30 ± 0.06 at the upper station. In the Scheldt estuary, P decreased from 0.76 ± 0.01 to 0.62 ± 0.01, while $1/K$ increased upstream from 1.20 ± 0.04 to 5.00 ± 0.04 at the most upstream station. In the second experiment, P was optimized while $1/K$ was set to 2, consistent with the values found by Kuijper and Van Rijn (2011) and Z. Zhang and Savenije (2019). With $1/K$ fixed, the model performance was slightly adjusted, as along-channel variations in $1/K$ were compensated by changes in the best-

Table 1
Overview of the Model Experiments Performed for the Ems Estuary and Results Discussed in Sections 4.3

Regression model	Boundary	Calibration	Prediction	Station (–)	Missing (%)	Pars[d] (–)	P/K (–)	Cal. rmse (psu)	Cal. Un. Var. (%)	Pred. rmse (psu)	Pred. Un. var. (%)
Tides (Equation 11)	Obs[a]	17 months	7 months	Emspier	13	31	0.57/0.77	1.8	6.8	1.8	6.6
				Pogum	4	29	0.58/0.50	1.9	7.2	1.9	7.9
				Gandersum	8	31	0.61/0.46	1.8	8.4	1.9	10.6
				Terborg	37	25	0.66/0.29	1.3	12.2	1.0	12.6
Tides + surges (Equation 12)	Obs[a]	17 months	7 months	Emspier	13	40	0.60/0.82	1.7	6.0	1.6	4.9
				Pogum	4	30	0.62/0.53	1.8	6.5	1.6	5.9
				Gandersum	8	34	0.64/0.48	1.7	7.6	1.6	8.0
				Terborg	37	32	0.67/0.44	1.2	9.8	0.9	11.1
Tides + surges (Equation 12)	Con[b]	17 months	7 months	Emspier	13	22	0.57/0.71	2.9	18.8	3.3	23.6
				Pogum	4	24	0.50/0.48	2.9	17.4	2.8	17.7
				Gandersum	8	20	0.51/0.47	2.6	17.5	3.6	19.3
				Terborg	37	18	0.59/0.45	1.7	23.0	2.4	14.7
Tides + surges (Equation 12)	Pred[c]	17 months	7 months	Emspier	13	38	0.57/0.77	1.6	5.6	1.5	4.3
				Pogum	4	44	0.58/0.50	1.8	6.0	1.5	5.3
				Gandersum	8	36	0.61/0.46	1.7	7.0	1.6	7.2
				Terborg	37	28	0.66/0.30	1.2	10.2	0.8	8.4
Tides + surges (Equation 13)	Con[b]	17 months	7 months	Knock	10	15	0.29/0.83	1.1	4.5	1.9	8.9

Note. The missing percentage refers to the total percentage of missing and rejected data points, while Un. Var. refers to the unexplained variance in the model results. [a]Fortnightly averaged salinity at the Knock station for the seaward boundary condition. [b]Constant value of 35 psu for the seaward boundary condition. [c]Parameterized seaward boundary condition. [d]Total of the fitted proportionality coefficients.

fit value of P. Similarly, in the third experiment, K was optimized while adopting the power found by Gisen and Savenije (2015) for P. These experiments highlight the equifinality in the model, where alternative sets of K and P can yield similar outcomes. Fixing K influences the optimized value of P, and vice versa. However, optimizing both powers yields the most robust parameter set, and consistent regression coefficients.

In the second set of experiments (where both P and K were optimized), the model was optimized using either the entire data set or a portion of it (as done in Section 4.3). Additionally, experiments were conducted where only data points with below-average discharge were used for optimization (Figures 10d–10i and 11d–11i). These experiments show that the selected period for model optimization has little influence on the best-fit values of P. However, the best-fit values of K showed more variability between periods, particularly at the most upstream stations. Notably, when only below-average discharge conditions were used, the values of $1/K$ displayed a clear offset in both estuaries, but maintained a similar along-channel trend. These results suggest that K is spatially variable and depends on discharge conditions. The interpretation of K is further explored in Section 6.

Figures 10g, 10h, 11g, and 11h compare the best-fit regression coefficients a_0 and a_1. The values of these coefficients varied between experiments and were dependent on the values of P and K. Although K differed significantly, the longitudinal variations in a_0 primarily reflect the reduction in relative dispersion as a function of longitudinal coordinate and channel convergence, following a similar trend across all experiments. A *t test* (p-value <0.05) confirmed that the values of a_0 and a_1 were significantly different from zero for all experiments. Figures 10i and 11i show the constant phase α_k of the M_2 constituent (see Equation A9), which increases upstream and shows little variation between experiments.

4.4.2. Constituent Selection and the Influence of Missing Data Points

In the regression model, the constituents are selected using the adjusted Rayleigh criterion based on the length of record and postfit significance testing. In addition, for each tidal constituent, it is tested if the contribution of the

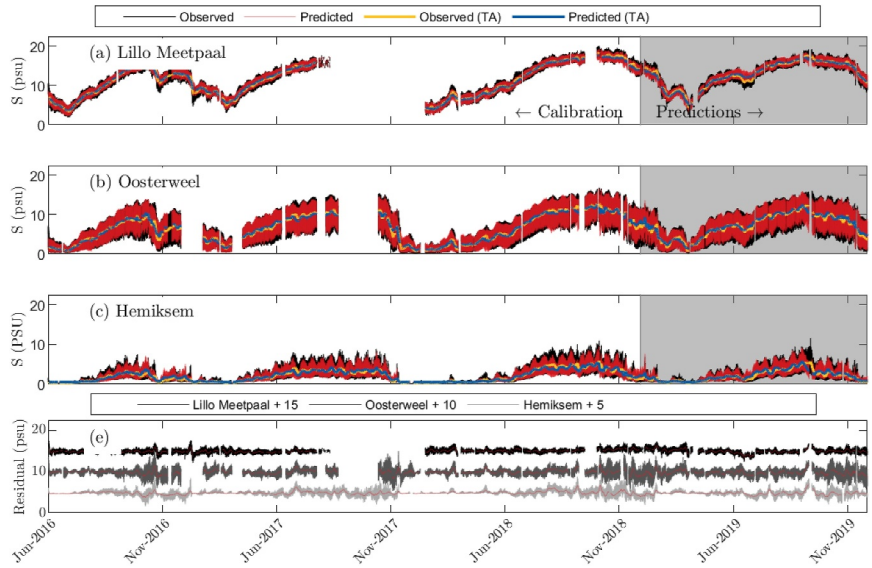

Figure 8. Data-model comparison for salinity in the Scheldt estuary, with salinity predicted using Equation 12 and the fortnightly averaged salinity at the Overloop van Hansweert station as seaward boundary condition. (a–d) Observed and predicted salinity variations and tide averaged (determined by a moving average, 1-day window). The period for which the salinity is predicted is highlighted by a gray background. (e) Residuals (predictions minus observations) of panels (a–d). The residuals include an offset for clarity of the figure.

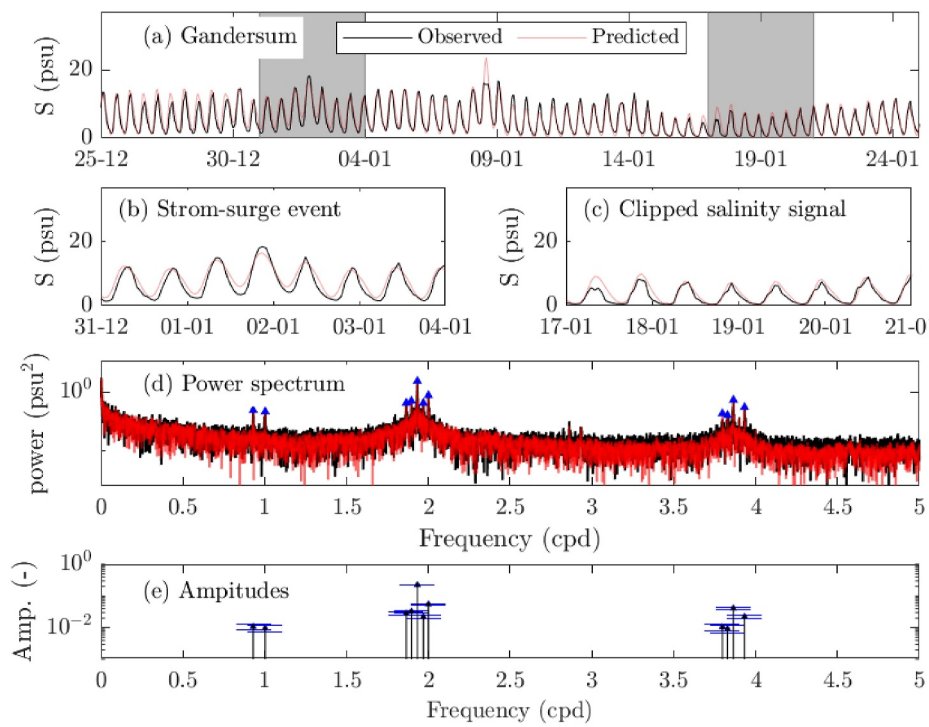

Figure 9. Data-model comparison for salinity at Gandersum station (data same as in Figure 7). (a–c) Zoom of the observed and predicted salinity. The gray boxes in panel (a) indicate the periods shown in panels (b and c). (d) Power spectrum of the predicted observed salinity and predicted salinity signal. The blue triangles indicate the included tidal coefficients. (e) Amplitude of the tidal constituents included (derived using Equation A7) and the blue error bars indicate the associated confidence bounds.

Table 2
Overview of the Model Experiments Performed for the Scheldt Estuary and Results Discussed in Section 4.3

Regression model	Boundary	Calibration	Prediction	Station (–)	Missing (%)	Pars[d] (–)	P/K (–)	Cal. rmse (psu)	Cal. Un. var. (%)	Pred. rmse (psu)	Pred. Un. var. (%)
Tides (Equation 11)	Obs[a]	2.5 years	1 year	Lillo Meetpaal	13	31	0.72/0.97	0.5	1.5	0.6	3.8
				Oosterweel	25	39	0.60/0.29	1.0	5.2	1.1	7.7
				Hemiksem	19	25	0.60/0.20	0.6	9.7	0.6	12.8
Tides + surges (Equation 12)	Obs[a]	2.5 years	1 year	Lillo Meetpaal	32	34	0.71/0.94	0.5	1.4	0.6	3.6
				Oosterweel	26	58	0.60/0.30	0.9	4.7	1.1	6.9
				Hemiksem	27	24	0.60/0.20	0.5	8.8	0.6	12.0
Tides + surges (Equation 12)	Con[b]	2.5 years	1 year	Lillo Meetpaal	26	8	0.65/0.29	2.8	43.0	3.2	69.2
				Oosterweel	26	28	0.66/0.45	2.2	26.3	2.2	27.9
				Hemiksem	27	18	0.63/0.30	1.0	27.6	1.0	34.4
Tides + surges (Equation 12)	Pred[c]	2.5 years	1 year	Lillo Meetpaal	26	28	0.73/0.94	0.5	1.4	0.6	3.6
				Oosterweel	26	48	0.60/0.29	0.9	4.6	1.0	6.9
				Hemiksem	27	30	0.60/0.14	0.5	8.9	0.6	12.2
Tides + surges (Equation 13)	Con[b]	2.5 years	1 year	Hansweert	19	31	0.46/0.98	0.5	2.3	0.9	11.6
				Prosperpolder	7	29	0.47/0.47	0.5	1.1	1.4	8.9
				Lilo Meetpaal	13	17	0.47/0.48	0.4	1.1	1.1	10.5

Note. The missing percentage refers to the total percentage of missing and rejected data points, while Un. Var. refers to the unexplained variance in the model results. [a]Fortnightly averaged salinity at Prosperpolder for the seaward boundary condition. [b]Constant value of 35 psu for the seaward boundary condition. [c]Parameterized seaward boundary condition. [d]Total of the fitted proportionality coefficients.

amplitudes related to the storm surge variations are significant (also see A5 in Appendix A). Furthermore, the selection procedure summarized above does not account for the possible influence of clipping on constituent selection. These two aspects are further addressed in this section.

Significant surge events occur only several times in a year. Therefore, the statistical uncertainty is higher for the coefficients associated to storm surge terms in Equation 12. In the performed experiments, the total number of included tidal constituents roughly varied between 10 and 18. However, only for a few constituents, the amplitude of storm surge terms were found to contribute significantly (Tables S1 and S2 in Supporting Information S1). By testing the individual amplitudes, instead of the overall resulting amplitudes, overfitting is prevented and the total number of model coefficients is reduced.

To illustrate the influence of missing data, the correlation matrix of the model coefficients of the constant tidal variations are visualized for the most down- and upstream stations of the Ems estuary in Figure 12. A similar figure is found for the Scheldt in Figure S5 in Supporting Information S1. For both stations shown in Figure 12, a 518-day period was available to optimize the regression model for the Ems, and based on the adjusted Rayleigh criterion, a frequency separation of 1.1 10^{-2} d^{-1} is required. Missing and rejected data points significantly differed between the two stations. At Emspier, in total, 12% data points were missing due to gaps in the data set, which increased by 1% after removing the data points where salinity is insignificantly different from the river salinity. At Terborg, 4% of the data points were missing, and an additional 33% of the data are rejected from the optimization of the harmonic model. This allows for the use and optimization of the harmonic model, but it influences the ability to distinguish between several tidal constituents. Constituents can become poorly identifiable because the low stages of the tidal cycles are less frequently sampled (i.e., these data points are rejected from the analysis).

For example, at Emspier, the correlation coefficients are found to be low (Figure 12a). As expected, the highest correlation coefficients occur between neighboring tidal constituents, peaking at 0.1 for O_1 and K_1. At Terborg, the correlation between neighboring tidal constituents increased, amounting to 0.18 for O_1 and K_1. However, a more distinct correlation exists between astronomical and shallow water constituents that differ in frequency by a

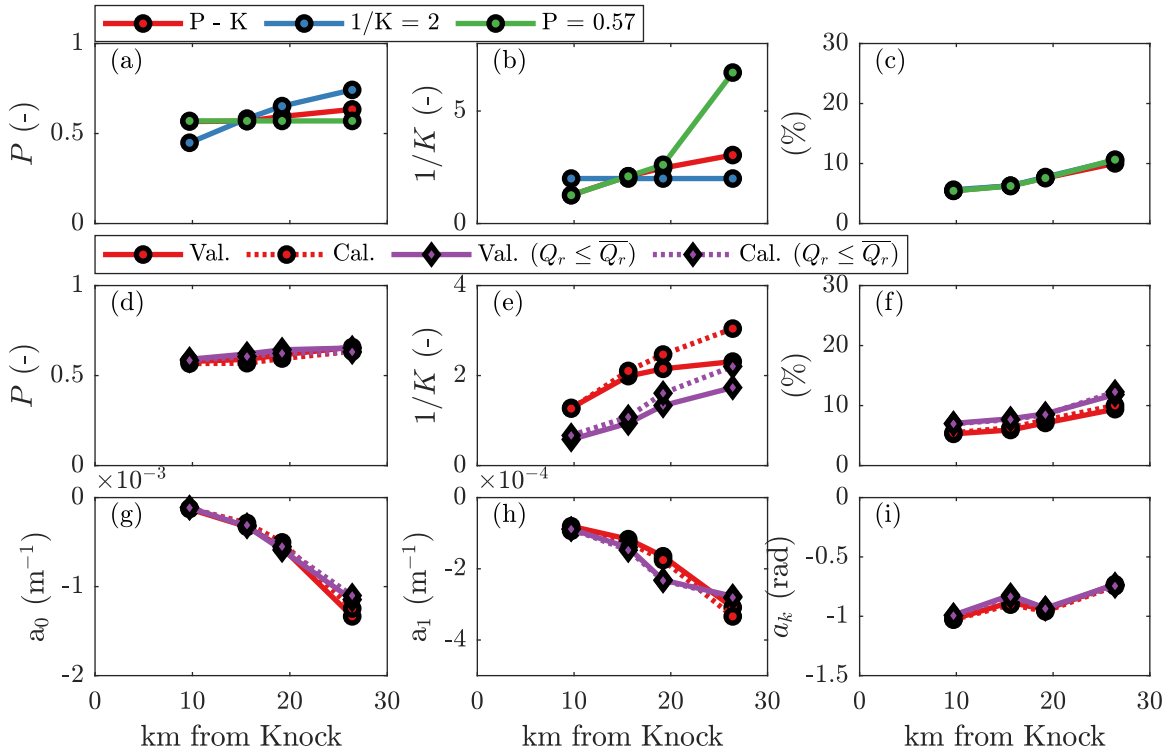

Figure 10. Overview of model coefficients for the Ems estuary and the resulting unexplained variance, using alternative settings to optimize the power coefficients. Results are derived from Equation 12 longitudinal profiles of the power coefficients P and K under different optimization conditions: when both powers are optimized (P - K), when $1/K = 2$, or when $P = 0.57$. (c) Resulting unexplained variance. (d–f) Same as panels (a–c), showing the difference in optimized powers P and K when part of the data set is used for optimization (Val.), compared to when the entire data set is used (Cal.). Additionally, the influence of discharge on the obtained power is analyzed by using only below-average discharge conditions (Val. $(Q_r \leq \overline{Q_r})$ and Cal. $(Q_r \leq \overline{Q_r})$), emphasizing the dependence of K on the discharge conditions. (g and h) Longitudinal development of the fitted model coefficients of the river and surge terms. (i) Amplitudes coefficients for the M_2 constituent.

factor of two. For example, at Emspier, the correlation between M_2 and M_4 is only 5.7 10^{-3}, but increases for the Terborg station to 0.22. Similar results are found for example, for S_2 and MS_4, and for N_2 and MN_4 (Figure 12b).

The correlation between these constituents is not caused by an insufficiently long data set; it only requires about 2 days to distinguish M_2 from M_4. The correlation arises from the fact that the salinity signal is less frequently sampled near LWS, which hampers the isolation of shallow water tides from astronomical tides. In our analysis, the correlations as obtained are weaker. However, additional synthetic experiments indicate that serious errors in the amplitudes and phases occur when ~50% of the data signal is clipped, which agrees with the results of Evans and Pugh (1982). When the signal contains substantial clipping, results can be improved by rejecting the shallow water constituents from the analysis.

5. Influence of the Downstream Boundary and Unsteadiness

Section 4.3 has demonstrated the ability of steady-state regression models to effectively predict truncated salinity signals with an appropriate boundary condition. In the evaluation so far, the model's predictive performance depends on the assumption of a steady state, and on salinity measurements from a down-estuary station a boundary condition. This section evaluates the impact of unsteadiness and of the down-estuary boundary condition. We explore the parameterizations that eliminate the need for measured salinity at the seaward boundary.

The influence of unsteadiness in the salt balance increases for stations located closer to the coast (see Section 4.2). In Section 4.3, good results were obtained by using the low-passed filtered measured salinity signal from a down-estuary station as a boundary condition. This approach captures the salinity variations caused by unsteadiness in the lower estuary, allowing for the use of a steady-state formulation. Relying on a constant salinity value ($S_0 = 35$ psu) instead of a time-varying boundary (significantly) reduces the model's predictive skill. In the Ems,

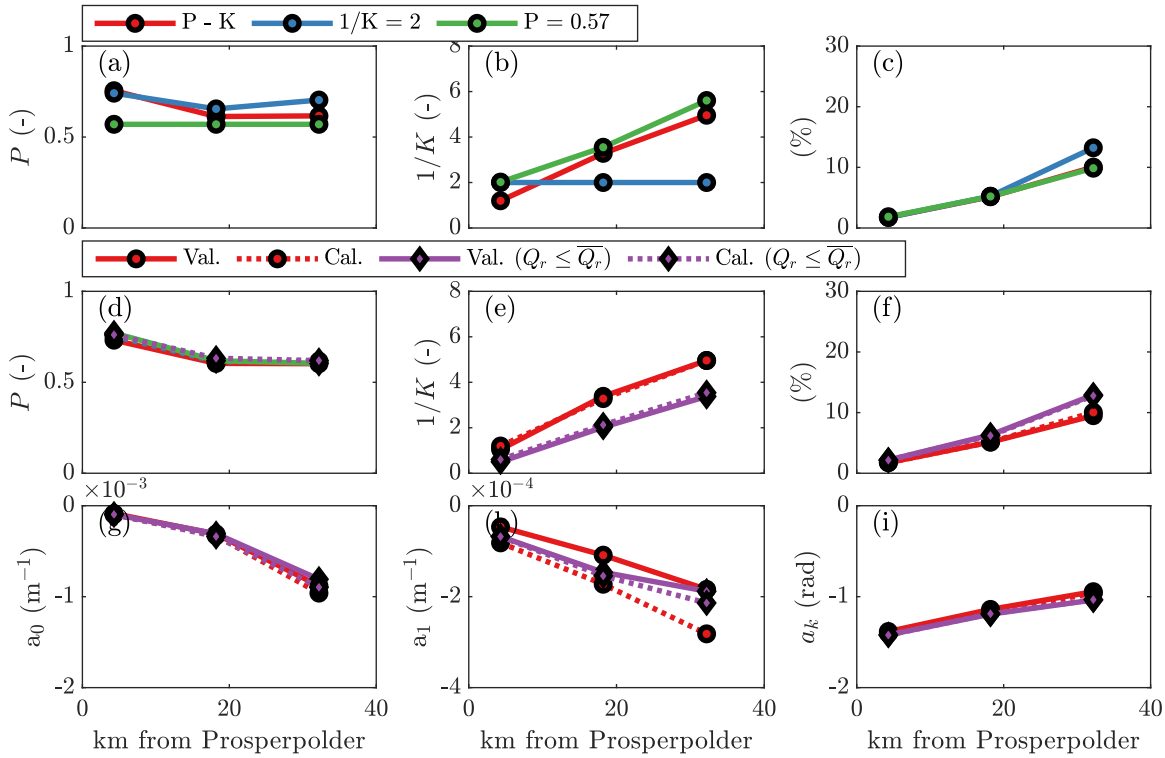

Figure 11. Same as Figure 10, but for the Scheldt estuary.

this leads to a reduction of about 15% in the explained variance. For the Scheldt, this reduction reached up to 50% at downstream stations and 15%–25% at upstream stations when using Equation 12 (see Tables 1 and 2).

Equation 16 parameterizes the salinity at the downstream station, where unsteadiness becomes significant. This approach enables accurate predictions of salinity at the upstream stations (with truncated signals) without the need to know the salinity values at the downstream boundary during the prediction phase. This approach led to a less

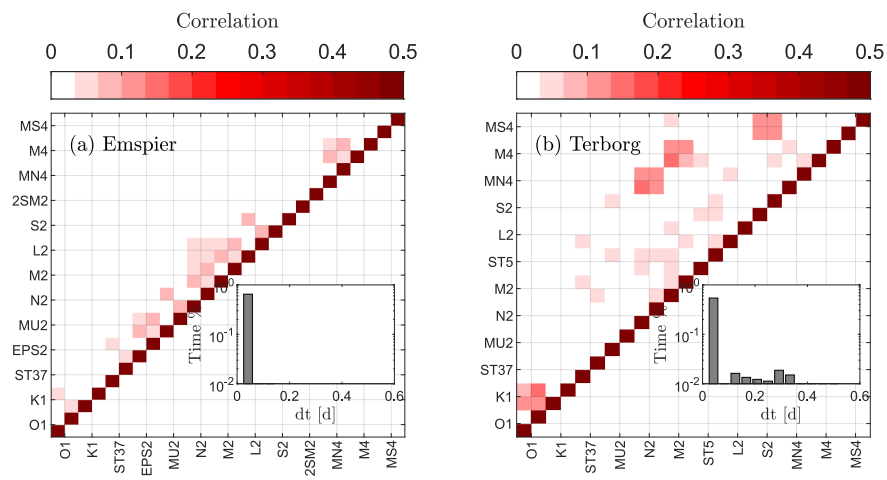

Figure 12. Correlation matrix of the model coefficients fitted for the harmonic model (storm surge terms are not included) at two stations in the Ems estuary, subpanels indicate the time step distribution after the data point selection. (a) Correlation matrix for the downstream station Emspier. Model coefficients very weakly correlate for neighboring tidal constituents. (b) Correlation matrix for the most upstream station Terborg. Weak–modest correlation is observed between neighboring constituents K_1/O_1 and between the astronomical and shallow water constituents with (approximately) twice the frequency of the astronomical constituent.

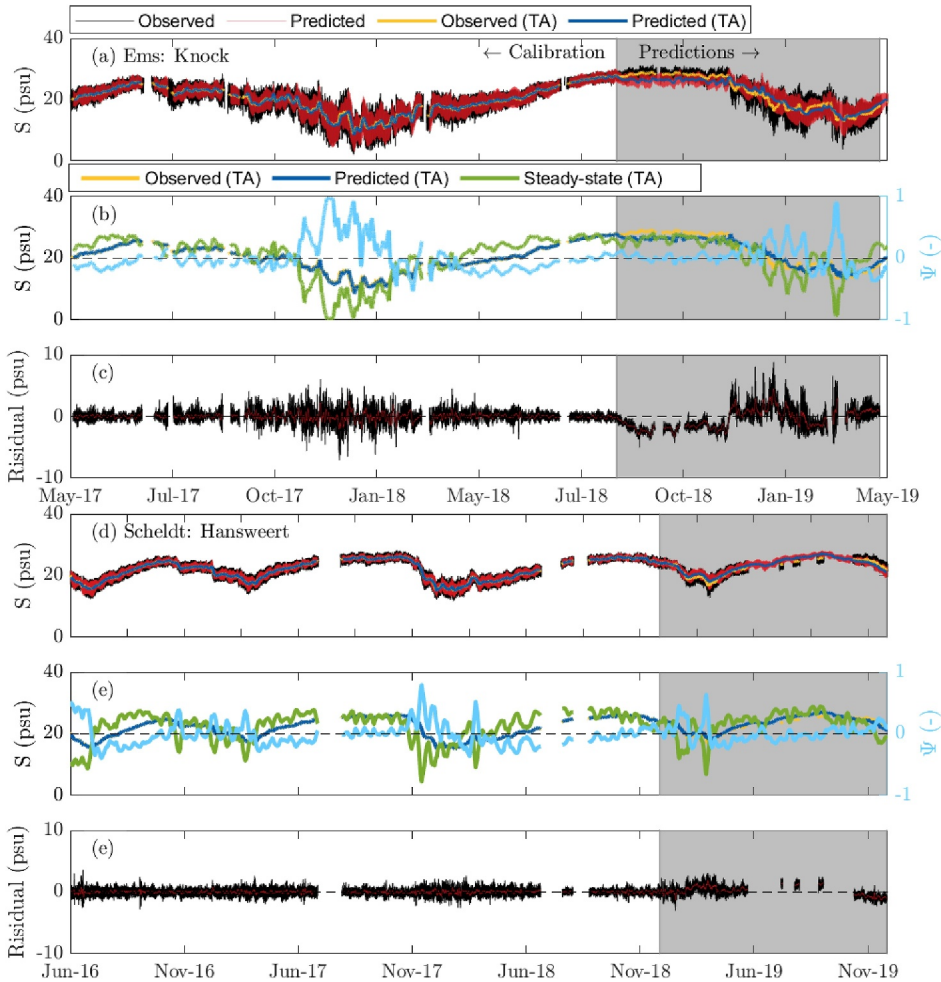

Figure 13. Data-model comparison using Equation 13 and a constant salinity (35 psu) for the seaward boundary condition. (a) Observed and predicted salinity variations and tide averaged (TA) (determined by a moving average, 1 day window) for the station Knock in the Ems estuary. The period for which the salinity is predicted is highlighted by a gray background. (b) Comparison between TA observed, TA predicted salinity variations, its equivalent steady-state salinity, and the TA predicted salinity when a steady-state approach is adopted (Equation 12). (c) Residuals. (d and e) Same but for the station Overloop v. Hansweert in the Scheldt estuary.

than 5% reduction in explained variance for the Ems estuary. In the Scheldt estuary, it improved the models accuracy especially for the downstream stations, reducing the explained variance by 2%–5% compared to the results in Section 4.3.

For nontruncated signals where unsteadiness effects are important, Equation 13, which directly incorporates the AR(1) lag into the model formulation, can be applied. This model is demonstrated for the two most downstream stations—Knock and Hansweert (see Figure 13)—where unsteadiness is most pronounced. For these stations, the unexplained variance was 8.9% and 11.5%, with RMSE values of 1.9 and 0.9 psu, respectively. In contrast, the steady-state approach explained only 34% and 77% of the variance, respectively. The empirical response times (T_r) for adjustment, derived from the model, were 24 and 35 days, which is notably longer than the typical period over which discharge variations occur. While Equation 13 better captures the salinity variations at these stations and significantly reduces the autocorrelation in the residuals, it still does not fully capture the dynamics, but presents an improvement over the steady-state approach.

In the steady-state approach, salt concentrations directly reflect the current forcing conditions. In contrast, the unsteady approach depends on both the previous state and the changes introduced by variations in the forcing conditions. The model embedding Equation 13 can be used to express the salinity that would be reached under

unchanged forcing conditions (i.e., steady-state), and infers whether the system is gaining or losing salt as follows:

$$\tilde{S}_s = \left(1 - a_0 K \frac{|Q_r|}{D_0} + a_1 \tilde{D}t\right)^{1/K}, \tag{17a}$$

$$\tilde{S}_{st} = \left(1 - \frac{a_0}{1 - a_1} K \frac{|Q_r|}{D_0}\right)^{1/K}, \tag{17b}$$

$$\Psi = \frac{\tilde{S}_s - \tilde{S}_{st}}{\tilde{S}}, \tag{17c}$$

where $\tilde{S}_s$ represents the predicted scaled subtidal salinity, $\tilde{S}_{st}$ is the equivalent steady-state salinity, and Ψ is the scaled difference. Specifically, $\Psi < 0$ indicates gaining salt, $\Psi > 0$ indicates a loss of salt, and $\Psi = 0$ indicates a steady state. Figures 13b and 13e illustrates the resulting values for Ψ, showing a clear seasonal cycle in both systems. Typically, during spring and summer, the estuary gains salt, while in autumn and winter, salt loss dominates. Similarly, the amplitudes of the harmonic component can be calculated, showing that the amplitudes are larger during periods of increased discharge ($\Psi > 0$), leading to stronger along-estuary gradients than periods of decreased discharge ($\Psi < 0$).

6. Critical Evaluation of Adopted Model Formulation

The previous sections show that the proposed model, which embeds the Savenije (1986) in formulation, is well capable of prediction of the salinity variations. This formulation is empirical rather than based on first principles. Therefore, in this section, we compare the Savenije (1986) model with the more theoretically substantiated Gay and O'Donnell (2007) formulation, and use the Gay and O'Donnell (2007) formulation to evaluate the obtained best-fit powers and interpret the meaning of the Van der Burgh constant.

6.1. Comparison With Gay and O'Donnell Model

Savenije (1986) assumes that the Van der Burgh coefficient (K) is constant. This assumption conflicts with our best-fit values of K, which exhibit significant site-to-site variations, specifically in the Scheldt estuary. Also, the physical interpretation of K provided by Savenije (2015) does not align with the general knowledge of both systems, nor is it reflected in the along-channel variations in the best-fit values.

Alternative models that avoid the use of Equation 3 have been proposed by Kuijper and Van Rijn (2011) and Z. Zhang and Savenije (2019), resulting in similar model structures, where K is set to 0.5. Alternatively, in the Gay and O'Donnell (2007) formulation, who derived a descriptive equation for along-channel salinity variations directly from Equation 1, the power in their formulation is described by physical properties. Assuming a linearly tapered estuary and a constant along-channel dispersion coefficient, at equilibrium, their model reduces to

$$\frac{S_x - S_r}{S_0 - S_r} = \left(1 - \frac{a}{A_0} x\right)^{\frac{Q_r}{aD_x}}, \tag{18}$$

where a represents the estuary tapering factor, and the dispersion coefficient is constant in this model ($D_0 = D_x$). A key difference with Equation 5 is that the exponent $\frac{Q_r}{aD_x}$ consists of physically interpretable variables. Here, Equation 18 is used to obtain analogous equations to Equations 11 and 12 for the harmonic regression model, resulting in Equations B1 and B2 in Appendix B. With Equation 18 as a base function, two power coefficients are optimized: again the power P for the dispersion coefficient, and γ, which determines the magnitude of the power $\gamma \frac{Q_r}{D_0}$. Extending this model with an AR(1) component is less straightforward due to the time-varying component appearing in the power. For this reason, we used the Savenije (1986) formulation in our approach, rather than the Gay and O'Donnell (2007) formulation.

The predictive skills of Equations B1 and B2 are comparable to the results discussed in Section 4.3, although the explained variance is approximately 2% lower for both estuaries (Tables B1 and B2). With both models resulting

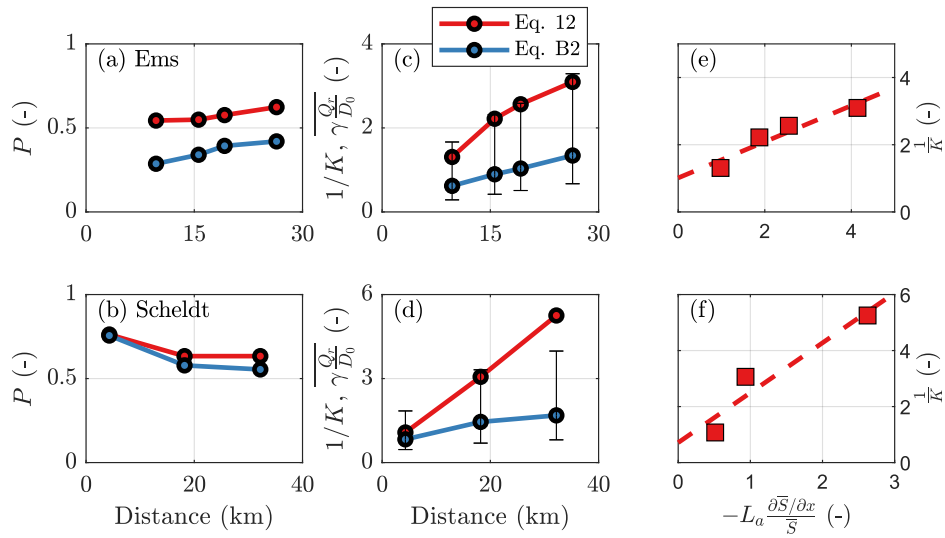

Figure 14. (a) Comparison of the power P obtained using Equation 12 and Equation B2 in the Ems estuary. (b) Same as (a), but for the Scheldt estuary. (c) Comparison of the power $1/K$ from Equation 12 with the time series-averaged power $\overline{\gamma \frac{Q_r}{D_0}}$ obtained with B2, along with its variability, in the Ems estuary. (d) Same as (c), but for the Scheldt estuary. (e) Relationship between $-L_a \frac{\partial \overline{S}/\partial x}{\overline{S}}$ and the best-fit values of K, expressed as $\frac{1}{K}$. The dashed lines represent the fitted regression line. (f) Same as (e), but for the Scheldt estuary.

in similar predictive skill, it is interesting to compare the best-fit value of the optimized powers (Figures 14a–14d). Figures 14a and 14b compare the best-fit values for P. The values show similar along-channel variations; however, the value obtained using Equations B1 and B2 are substantially lower in the Ems, whereas more comparable values are obtained for the Scheldt. Figures 14c and 14d compare the best-fit values of $1/K$ with the time series-averaged power $\left(\overline{\gamma \frac{Q_r}{D_0}}\right)$ and its variability, where the overbar denotes the long-term average. Again, both powers show a similar trend in the up-estuary direction, but their magnitude differs. This suggests that the locally optimized Van der Burgh constant may have a similar physical meaning as the power in the model by Gay and O'Donnell (2007). This is further explored in the next section.

6.2. Interpretation of the Van der Burgh Constant

In the model of Gay and O'Donnell (2007), the power governs the time-varying spatial scale of dispersion and advection relative to the imposed scale over which channel geometry changes. To evaluate whether the locally optimized Van der Burg constant could have a similar meaning as the power in the model of Gay and O'Donnell (2007), the values of $1/K$ are plotted against $-L_a \frac{\partial \overline{S}/\partial x}{\overline{S}}$. The latter can be expressed using Equation 2 as $-\left(\overline{Q_r} L_a\right)\left(\overline{D_x} A_x\right)^{-1}$, and the same group of parameters is recognized as in Equation 5, which determines the relative dispersion coefficient. Recognizing that $-L_a = \frac{A_x}{\partial A_x/\partial x}$ allows to write $1/K \propto \left(\frac{\overline{Q_r}}{\overline{D_x} \partial A_x/\partial x}\right)$ or $1/K \propto \left(\frac{\overline{u_r}}{\overline{D_x}} \frac{A_x}{\partial A_x/\partial x}\right)$, which results in the same parameter group as the power term in Equation 18. We calculate the long-term average salinity and its along-channel derivative from the average salinity distribution shown in Figure 4. For the cross-sectional convergence lengths (L_a), values of 2.2×10^3 m for the Ems and 1.8×10^3 m for the Scheldt estuaries were used.

Based on the above, Figures 14e and 14f show the ratio $-L_a \frac{\partial \overline{S}/\partial x}{\overline{S}}$ obtained from the data against the best-fit values for $1/K$ for both estuaries, which results in near-linear relation for the values within one system. The relationships differ between the Ems and the Scheldt estuaries, which is confirmed by a statistically significant difference in the slopes of the fitted regression lines. Whether this difference in slope is due to inherent physical properties of the estuaries or represents an artifact of the model, remains unclear. However, if the relationship seen in Figures 14e and 14f holds within one system, it suggests that the power $1/K \propto \left(\frac{\overline{u_r}}{\overline{D_x}} \frac{A_x}{\partial A_x/\partial x}\right)$ and the locally optimized Van der

Burgh constant indeed has a similar physical meaning as the power in the model of Gay and O'Donnell (2007). The inverse of our locally fitted Van der Burgh constant may then be interpreted as an empirical coefficient that scales with the average spatial scale of the dispersion and advection magnitude, relative to imposed scale of channel geometry change. Finally, this also explains the substantially lower values for $1/K$ when only using low discharge condition to optimize the model (see Section 4.3).

6.3. Sensitivity to the River Discharge

In Section 6.1, we observe that the best-fit values for the power P, derived from using either Equation 5 or Equation 18 as the base function, show comparable trends along the channel. However, in the Ems, values obtained with Equation 18 are noticeably lower, while in the Scheldt, both equations yield similar values. Here, we further assess the observed difference in the best-fit values of P, which governs the sensitivity to the river discharge.

To find out if the power P differs between the Ems and Scheldt estuaries, the relationship between Q_r and D_0 was examined by calculating the dispersion coefficient according to Equation 2, using salinity time series and the observed gradients between different stations. The obtained dispersion coefficients are subsequently used to obtain an estimate of P in $Q_r^P \propto D_x$, where the power is estimated by means of a regression (see Figures S6 and S7 in Supporting Information S1). For the Ems estuary, lower values of P (0.4–0.55) are obtained from the data compared to the Scheldt estuary (0.66–0.79), which is in agreement with a substantial difference in the value of P for the two estuaries that can be expected based on Equation 18. A likely reason for the fact that this effect is not fully captured by Equation 5 lies in the model structure. The sensitivity to discharge variations can be accommodated by modifying K in Equation 5. Hence, differences in the actual values of P are compensated for by an altered value of K. This model characteristic may partially account for the variations observed between the two systems in the previous section. However, in both systems, estimates of P derived from the data show an increase in the up-estuary direction. This aligns with the along-channel trend of best-fit values for P in the Ems. In the Scheldt, however, the best-fit values indicate a decrease. The elevated value at the first station in the Scheldt is likely due to its close proximity to the downstream boundary station, which limits the sensitivity to discharge fluctuations.

7. Discussion

7.1. Possibilities and Limitations of the Model

The proposed harmonic regression model to predict estuarine salinity variations at monitoring sites yields favorable results for two applications. To demonstrate in which type of systems the model works best, the Ems and Scheldt estuaries can be projected onto the classification diagram by Geyer and MacCready (2014). This diagram maps the estuarine parameter space based on a mixing parameter M and the freshwater Froude number Fr_f. M quantifies the effectiveness of tidal mixing for stratified estuaries and Fr_f is a measure for stratification driven by gravitational circulation, which scales as $Fr_f^{\frac{2}{5}}$ (MacCready, 1999). To classify the Ems and Scheldt estuaries in the $M - Fr_f$ plane, we used the observed forcings at the boundary and geometry estimates from Gisen and Savenije (2015), shown in Figure 15a. Both estuaries fall in the lower right quadrant of the diagram. The Scheldt estuary is classified as a well-mixed system, bordering the weakly stratified strain-induced periodic stratification (SIPS) regime. The Ems estuary falls in the SIPS regime and approximates the partially mixed regime for higher discharges. The fact that the Ems estuary is located in the SIPS regime may explain why the portion of unexplained variance was higher than the Scheldt.

The proposed regression model is based on a steady-state salinity distribution, best applicable when the adjustment timescales of the salinity front are shorter than the timescales at which the discharge varies (C. Kranenburg, 1986; Monismith et al., 2002). An expression for the adjustment time is $T_A = l_s \frac{1}{2} \frac{A}{Q_r}$ for systems dominated by tidal stirring and $T_A = l_s \frac{1}{6} \frac{A}{Q_r}$ for exchange-dominated systems (MacCready & Geyer, 2010). These expressions indicate that the adjustment time increases with the length of the system and is inversely related to the strength of the residual current. This suggests that the proposed steady-state model is best applicable to systems with shorter salt intrusion lengths and stronger residual currents, which allow for quicker adjustment to variations in discharge. For the systems evaluated in this study, spectral analysis of the discharge time series reveals that the

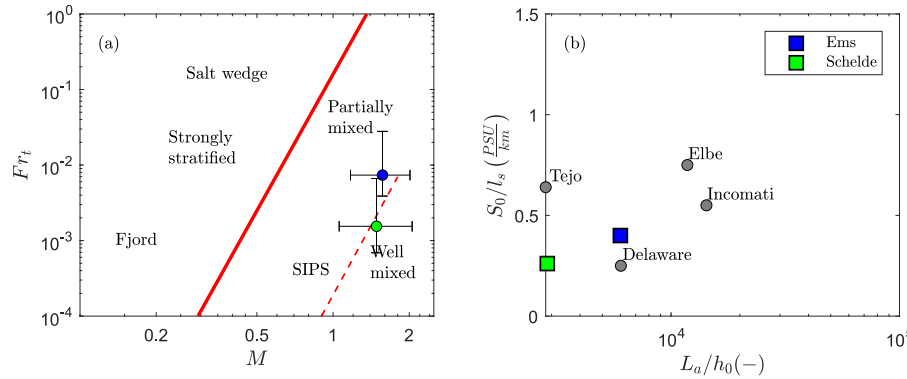

Figure 15. (a) Classification of the Ems and Scheldt estuaries in the Geyer and MacCready (2014) diagram, including variability around the average location in the diagram that results from spring-neap tide and discharge variations. (b) Matrix to assess the sensitivity of advective displacements by storm surges, based the dimensionless shape factor $\frac{L_a}{h_0}$ for converging estuaries and a typical salinity gradient $\frac{S_0}{l_s}$. Location of different estuaries are based on literature (Gisen & Savenije, 2015).

majority of discharge variations (over 75%) occurs at periods below 14 days, for both systems. The adjustment times calculated for exchange-dominated systems, based on average intrusion lengths (98 km for the Ems and 66 km for the Scheldt), discharges (79 m^3/s for the Ems, 66 m^3/s for the Scheldt), and cross-sectional areas ($1.4 \cdot 10^4$ m^2 at Knock, $1.1 \cdot 10^4$ m^2 at Hansweert), are 22 days for the Ems and 33 days for the Scheldt. These values agree reasonable well with the empirical derived response times. In both cases, the adjustment time exceeds the typical time scale of discharge variations.

The effects of unsteadiness were found to increase in the down-estuary direction, while remaining smaller at the more up-estuary station (Figure 5). Hence, although in both systems the adjustment timescale exceeds the typical timescale of discharge variations, a steady-state model can still be applied successfully in the more up-estuary sections. This is possible by using a time-varying boundary condition at a down-estuary station, where unsteadiness becomes more significant (see Section 4.3), or by employing a parameterized boundary condition (see Section 5). For stations where unsteadiness is significant and the salinity signal is not truncated, the formulation with the AR(1) extension can be effectively applied. Ideally, this approach would also be useful for stations with truncated signals. However, for these cases, we were unable to obtain reliable estimates of the AR(1) coefficient. Additionally, a constant empirical response time is currently used, but theoretically, an adjustment time that varies with discharge would be more appropriate.

The importance of storm surges varied between the two systems. Including the effect of storm surges improved the explained variance by 1%–2% and 0.2%–0.5%, respectively for the Ems and Scheldt estuaries (see Section 4.3). The latter suggests that the Scheldt estuary is less sensitive to storm surges than the Ems estuary. This can be explained by differences in the geometry and the intrusion lengths. A shorter intrusion length results in larger longitudinal salinity gradients, that is, a higher sensitivity to salinity variations caused by advective displacements. Similarly, the influence of tides and storm surges on local salinity often increases for higher discharges (Figure 16). The shape of the estuary largely determines the relation between the vertical and horizontal excursions (Equation 8), and scales with the channel convergence to the depth ratio $\left(\frac{L_a}{h}\right)$. This ratio is higher for the Ems than the Scheldt estuary, and hence, a larger vertical excursion length is found for the Ems estuary (Figures 6b and 6d). Figure 15b shows a simple matrix, based on the intrusion length and estuary shape, to demonstrate the sensitivity of a (converging) estuary to storm surges. Estuaries where the salt intrusion length is closer to the tidal excursion are likely more sensitive to storm surges.

The horizontal excursion resulting from storm surges is remarkably well captured by a simple linear relation. The appropriateness of this relation depends on the wave character of the occurring storm surge. When a storm surge wave has a standing wave character, the resulting horizontal tidal excursion varies near-linearly with the surge water levels (Figure 6). For the Ems and Scheldt estuaries, adding the extrahorizontal excursion resulting from storm surge water levels allowed us to reconstruct the observed salinity variations. However, it is likely that in other systems, such a simple relation cannot describe the dynamics. For example, C. Li et al. (2010), Gong et al. (2018), and W. Kranenburg et al. (2022) found that in partially and stratified systems, storm surge events can

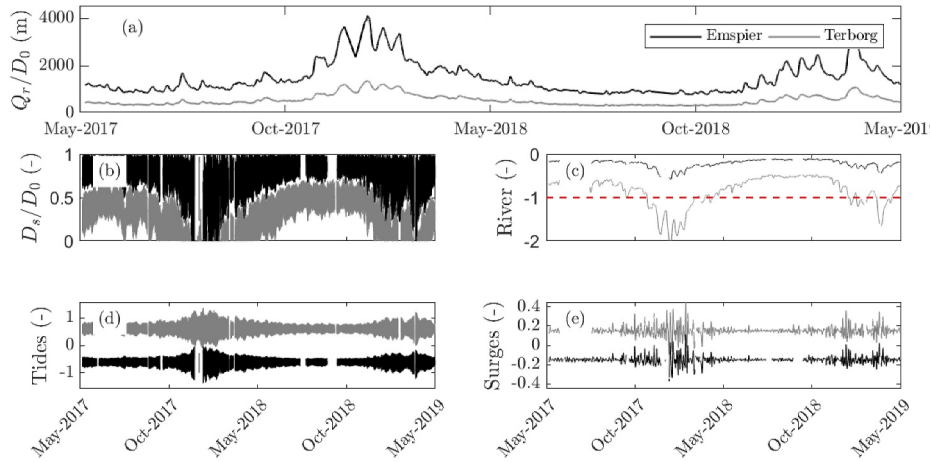

Figure 16. Components of the harmonic regression model, based on Equation 12, illustrated for the stations Emspier and Terborg. (a) Temporal variation in $\frac{Q_r}{D_0}$, which scales each term in the regression model. (b) Full signal (i.e., relative dispersion) of panels (c–e). (c) Contribution of the tide-averaged term, resulting from a balance between dispersion and flushing. The red dashed line indicates the threshold value at which the signal is truncated. (d and e) Contribution of the tidal and storm surge terms, resulting from a horizontal excursion of the salinity distribution. For clarity of the figure, an offset of ±0.6 in panel (d) and ±0.15 in panel (e) is used.

significantly increase mixing and reduce stratification, arresting the estuarine circulation. In our regression model, surge water levels increase the tidal range and mixing through Equation 14. Whether or not the regression model is capable of capturing storm surge dynamics in more stratified systems requires further testing. In long estuaries, where the typical wavelength of a surge wave does not exceed the length scale of the estuary, storm surge waves may have a progressive character. These systems may require a frequency specific approach to capture the storm surge variations.

7.2. Interpretation on Causes of Salinity Variations

The regression model facilitates the examination of various factors influencing along-channel salt distribution and variability in salt concentration at individual sites. As discussed in Section 6, the along-channel variations in the best-fit values of K seem to reflect changes in dispersive and advective magnitudes relative to the imposed scale of channel geometry change $\left(1/K \propto \frac{\overline{u}_r}{D_x} \frac{A_x}{\partial A_x/\partial x}\right)$. Further research is needed to confirm this relationship across different systems and explore the sensitivity to the value of P. The harmonic regression model can elucidate the causes of salinity variations at specific sites. For example, predictions for the Ems estuary (see Section 4.3) are divided into contributions from tide-averaged oscillations, tidal dynamics, and surge-induced variations, as shown for the most down- and upstream stations in Figure 16. For these stations, approximately 75% of the variance is attributed to tidal dynamics, whereas for Pogum and Gandersum, tide-averaged oscillations contribute around 35%–40% of the variance. The overall contribution of storm surge variations is limited to about 2%. In the steady-state models, each term is scaled by $\frac{Q_r}{D_0} \propto Q^{P-1}$, which predominantly reflects the discharge variations. Hence, each term amplifies with increasing discharge. With the model to include unsteadiness effects (trough autoregressive term) as proposed, the interpolation of the terms changes, but allows for a similar decomposition of the salinity signal.

7.3. Comparison With Other Predictive Methods

The goal of this paper is to develop a tool capable of providing accurate short- to medium-term predictions, while also providing insights into the contributing mechanisms. An alternative approach for obtaining short to medium-term predictions is the use of artificial neural networks (ANNs). Several studies showed that ANNs are well capable of providing accurate estimates of estuarine salinity (Rath et al., 2017; F. Zhou et al., 2020; X. Zhou et al., 2017). A well-trained ANN may outperform the regression model proposed herein in terms of predictive skill, especially when the salinity signal contains nontidal variations and the unsteadiness of the estuary is high.

Compared to an ANN, the advantage of the proposed harmonic regression methodology is the transparency in the model structure, offering insight into the contribution of various physical mechanisms. This is in contrast to an ANN approach, in which the relation between the input and the explained variables is unclear. Our approach may better predict events of salinity intrusion, in which conditions are more extreme than what is captured by the training data.

8. Conclusion

A harmonic regression model has been developed and evaluated to analyze and predict salinity variation at monitoring stations. The model is based on the steady-state salt intrusion model of Savenije (1986) and is extended to capture variations caused by tides and storm surges. Tidal variations are represented using harmonic functions, while storm surge-induced changes are modeled through a linear relationship, which is theoretically justified when the typical wavelengths of storm surges have longer wavelengths than the estuary's length scale. An autoregressive term is introduced to capture unsteadiness effects. To accurately reproduce local salinity, the model optimizes two power coefficients and proportionality constants. Salinity predictions are truncated to account for periods when the intrusion length occasionally reaches the monitoring station, though this truncation affects the identifiability of higher harmonics. The model's predictive performance was demonstrated at seven stations in the upper sections of the Ems and Scheldt estuaries, explaining between 87.4% and 96.4% of the observed variance. The dominant salinity response to storm surge variations was well captured, with horizontal displacement of the longitudinal salinity profile being the primary response. This reveals the sensitivity of salinity in well-mixed systems to storm surges, based on a simple matrix relating geometry and intrusion length. Close to the coast, where the estuary has widened, effects of river discharge unsteadiness increases. For stations with untruncated signals, the extended autoregressive model enabled accurate predictions at these sites, and assessed the lagged response. The best-fit power coefficients were critically evaluated, providing insights into the physical interpretation of the Van der Burgh constant, showing that it scales with the spatial scales of dispersion and advection, relative to changes in channel geometry. The proposed harmonic regression model is expected to accurately predict estuarine salinity in well-mixed systems. For those systems, the model can act as an analytical tool to interpret and predict nonstationary salinity signals at monitoring stations, requiring minimal input.

Appendix A: Details Regarding the Solution Procedure

A1. Data Preprocessing

The harmonic regression models are optimized using data points that exceed a minimum salinity threshold, $S_{\min}$. Salinity sensors often exhibit deviations in their minimal detection range and accuracy, which may make the salinity values from an upstream station unsuitable for $S_{\min}$. Instead, $S_{\min}$ can be based on the 95% confidence upper bounds derived from the regression model for river salinity (Equation 15). Alternatively, a constant threshold value may be used, although this could lead to the unnecessary rejection of data points from the analysis.

The use of Equation 12 requires storm surge-induced water level variations (δ_w). In de model, δ_w is inferred from provided water levels record at the downstream boundary. The astronomical water level variations are estimated using T-TIDE (Pawlowicz et al., 2002); the nonexplained variations (residuals) are assumed to be atmospherically forced, and subsequently band-pass filtered to isolate the range of oscillations at periods between 1.2 and 14 days. Surge variations at timescales smaller than 1.2 days and exceeding 14 days cannot properly be distinguished from the tidal motion, including fortnightly variations. For the experiment discussed in Section 4, the upper limit is set to 6 days. Furthermore, salinity variation are translated to density series based on the UNESCO (1981) formulation, which is required for Equation 14.

A2. The Regression Matrix

For the optimization to obtain the best-fit regression coefficients and model powers, Equations 9–12 are rewritten in matrix notation as follows:

$$\mathbf{S} = \mathbf{D} \cdot \mathbf{x}, \tag{A1}$$

where **S** is a vector of the resulting relative dispersion over time minus one (i.e., the LHS minus the first term on the RHS):

$$\mathbf{S} = \left(\frac{S_{s,i} - S_{r,i}}{S_{0,i} - S_{r,i}}\right)^K - 1, \tag{A2}$$

with the index i ($i = 1{:}\, m$) for the moment in time, used in the regression analysis (see A1 in Appendix A). The design matrix **D** for Equation 12 reads as follows:

$$\mathbf{D} = K\frac{Q_r}{\left(\frac{\Delta\rho}{\rho}\frac{Q_r}{\eta^3\omega}\right)^P(\eta^2\omega)}\Big[1 \quad \delta_{w,t_i} \quad \cos(\omega_1 t_i) \quad \ldots \quad \cos(\omega_n t_i) \quad \sin(\omega_1 t_i) \quad \ldots \quad \sin(\omega_k t_i) \quad \delta_{w,t_i}\cos(\omega_1 t_i) \quad \ldots \quad \delta_{w,t_i}\cos(\omega_n t_i) \quad \delta_{w,t_i}\sin(\omega_n t_i) \quad \ldots \quad \delta_{w,t_i}\sin(\omega_n t_i)\Big], \tag{A3}$$

with the index k ($k = 1{:}\, n$) for a tidal constituent added in the model. The unknown coefficients **X** are given by

$$\mathbf{X} = \begin{bmatrix} a_0 & a_1 & a^c_{2,1} & \ldots & a^c_{2,n} & a^s_{2,1} & \ldots & a^s_{2,n} & a^c_{3,1} & \ldots & a^c_{3,n} & a^s_{3,1} & \ldots & a^s_{3,n} \end{bmatrix}, \tag{A4}$$

which contains the proportionality coefficients of the subtidal (a_0) and storm surge variations (a_1), and the amplitudes for the intertidal variations ($a^c_{2,n}, a^s_{2,n}, a^c_{3,n}, a^s_{3,n}$).

A3. The Optimization Procedure

The proportionality coefficients **X** and the powers P and K of the model are iteratively optimized. First, an initial estimate for the proportionality coefficients is determined by minimizing the merit function through an ordinary least squares regression (OLSR) as follows:

$$\mathbf{R}^2 = |\mathbf{D}\cdot\mathbf{x} - \mathbf{S}|^2, \tag{A5}$$

with **R** being the residuals. Hereafter, the best-fit values for the powers are found using MATLAB's *lsqcurvefit* function, which minimizes the least squared error between the predicted and the observed salinities. For the OLSR, only data points where the observed salinity exceeds $S_{\min}$ are used. For the predicted salinity, all valid data points are considered. However, the predicted relative dispersion (RHS Equations 9–12) may fall outside the range yielding physical solutions. To prevent this, a truncation is applied as follows:

$$\frac{D_{s,i}}{D_{0,i}} = \begin{cases} \dfrac{D_{s,i}}{D_{0,i}}, & \text{if } \dfrac{D_{s,i}}{D_{0,i}} \geq 0 \\ 0, & \text{otherwise,} \end{cases} \tag{A6}$$

which avoids below-zero values. For each new estimate for P and K, a new estimate for the proportionality coefficients is made through Equation A5, until a (local) minimum is found in the sum of squares. For the initial estimate, $P = 0.57$ (from Gisen and Savenije (2015)) is used and varies between 0 and 1 in the optimization, while K has a default for $K = 0.5$ and is optimized within $0.05 < K < 2$. The included tidal constituents are determined based on the adjusted Rayleigh criterion and depend on the power P (see A5 in Appendix A).

A4. Amplitudes and Phases of the Tidal Constituents

The coefficients $a^c_{2,n}, a^s_{2,n}, a^c_{3,n} a^s_{3,n}$ in Equation A4 are associated to the intertidal dynamics. The coefficients related to the intertidal variations can be expressed in terms of (time-varying) amplitudes and phases for each individual tidal constituent added in the model (Matte et al., 2013). For this purpose, the amplitudes (A_k and B_k) and corresponding phases (α_k and β_k) are defined as follows:

$$A_k = \sqrt{a_{2,k^c}^2 + a_{2,k^s}^2}, \tag{A7}$$

$$B_k = \delta_{w,T}\sqrt{a_{3,k^c}^2 + a_{3,k^s}^2}, \tag{A8}$$

$$\alpha_k = \arctan\frac{a_{2,k^s}}{a_{2,k^c}}, \tag{A9}$$

$$\beta_k = \arctan\frac{a_{3,k^s}}{a_{3,k}^c} \tag{A10}$$

where the constant amplitude A_k represents the contribution to the constant intertidal variations and B_k is the time-varying amplitude resulting from by storm surges introduced intertidal variations. The constant phases are given by α_k and β_k, respectively. As shown by Matte et al. (2013), for each tidal constituent, the resulting time-varying amplitude and phases can be derived from A4–A7, and they are given by

$$|D_k(t)| = |d_k(t)| + |d_{-k}(t)| \tag{A11}$$

$$\psi(t) = \arctan\frac{\mathrm{Im}(d_{-k})}{\mathrm{Re}(d_{-k})} \tag{A12}$$

with d_{-k} and its complex conjugate d_{-k}^* obtained from

$$d_{-k} = d_{-k}^* = \frac{1}{2}\left(A_k\cos(\alpha_k) + B_k\cos(\beta_k)\right) + i\frac{1}{2}\left(A_k\sin(\alpha_k) + B_k\sin(\beta_k)\right). \tag{A13}$$

The resulting amplitudes and phases in Equation 11 are equivalent to the amplitudes A_k and phases α_k.

A5. Constituent Selection and Error Estimates

An adjusted version of the twofold strategy (Matte et al., 2013) is used for the constituent selection and error estimates. The importance of the constituents is obtained from an initial harmonic analysis performed on salinity time series at the target station. The lower bound frequency that can be resolved under the assumption of a regularly sampled signal is set by the Rayleigh criterion, which states two frequencies can be distinguished when $\frac{|\omega_1 - \omega_2|}{LOR} < R$, with R set to one and LOR is the length of record. In addition, the degree in which constituents can properly be distinguished is also affected by nonstationary influences (Matte et al., 2013). In the case of salinity, time series broadening of spectral lines can either be introduced by nonstationarity of the salinity distribution itself or indirectly via a nonstationary tidal signal. For the frequency selection, only the influence of non-stationarity in the salinity distribution is considered (additional to the traditional Rayleigh criterion). The minimal allowable frequency separation is evaluated using an adjusted Rayleigh criterion (Matte et al., 2013), for which the normalized spectrum of Q^{1-P} is evaluated. If the time series contain gaps, the two frequencies cannot properly be identified and R should be adjusted accordingly. There is, however, no accepted practice to adjust the Rayleigh criterion for irregularly sampled data (Codiga, 2011). For irregularly sampled data, the degree in which tidal constituents can properly be resolved is monitored through the correlation between the model coefficients, following Foreman et al. (2009).

The significance of added constituents and error estimates of the amplitudes or phases of the harmonic model are based on the uncorrelated noise model of NS-TIDE (Matte et al., 2013). The uncorrelated noise is used to calculate a distribution realization for each model coefficient. Subsequently, for each replicate, the resulting amplitude and phase are calculated based on Equations A11 and A12, and used to compute the time-varying confidence bounds for each tidal constituent. In NS-TIDE, the significance of each constituent is estimated using the squared ratio of the resulting amplitude and amplitude error. The drawback of this approach is that only the contribution of the overall resulting amplitude is evaluated, while the contribution of the amplitudes related to individual mechanisms, based on Equations A7 and A8, may be insignificant. Therefore, first, the significance of each tidal constituent based on the overall resulting amplitude is tested, adopting a signal-to-noise ratio (SNR)

bound of 4. Hereafter, it is tested whether the amplitudes related to the individual mechanisms are actually significant, adopting the same SNR bound. If not, the terms related to the individual mechanisms are removed from the model structure for that specific tidal constituent.

Appendix B: Alternative Base Function for the Harmonic Regression Model

In this paper, we have chosen to use the analytical model for steady-state salinity distribution derived by Savenije (1986). However, an analogous set of equations, Equations 9–12, can also be derived using the more theoretically substantiated model proposed by Gay and O'Donnell (2007). Equation 18 can be extended to account for tidal and storm surge variations in a similar manner as described in Section 3.1. The harmonic regression including tidal variations is expressed as

$$\left(\frac{S_s - S_r}{S_0 - S_r}\right)^{\left(\gamma \frac{D_0}{Q_r}\right)} = 1 - c_0 + \sum_{k=1}^{n} [c^c_{2,k} \cos(\omega_k t) + c^s_{2,k} \sin(\omega_k t)], \tag{B1}$$

and the extension with storm surge variations is given by

$$\left(\frac{S_s - S_r}{S_0 - S_r}\right)^{\left(\gamma \frac{D_0}{Q_r}\right)} = 1 - \underbrace{c_0}_{\text{TA}} + \underbrace{c_1 \delta_w}_{\text{SURGE}} + \underbrace{\sum_{k=1}^{n} [(c^c_{2,k} + c^c_{3,k}\delta_w)\cos(\omega_k t) + (c^s_{2,k} + a^s_{3,k}\delta_w)\sin(\omega_k t)]}_{\text{TIDAL}}. \tag{B2}$$

Equations B1 and B2 include two power coefficients that need to be optimized: P associated with D_0 and γ, which scales the power $\left(\gamma \frac{D_0}{Q_r}\right)$. To obtain the best-fit values for these parameters, the same optimization procedure

Table B1
Overview of the Model Experiments Performed for the Ems Estuary and Results Discussed in Sections 4.3

Regression model	Boundary	Calibration	Prediction	Station (–)	Missing (%)	Pars[b] (–)	P/γ (–)	Calibration rmse (psu)	Calibration Un. var. (%)	Predictions rmse (psu)	Predictions Un. var. (%)
Tides (Equation B1)	Obs[a]	17 months	7 months	Emspier	13	31	0.37/ $8.7{\cdot}10^{-5}$	1.9	7.8	2.2	8.9
				Pogum	4	31	0.35/ $7.4{\cdot}10^{-5}$	1.9	7.2	2.0	7.6
				Gandersum	8	31	0.39/ $1.2{\cdot}10^{-4}$	1.7	7.7	2.1	9.2
				Terborg	37	31	0.44/ $2.1{\cdot}10^{-4}$	1.2	11.7	1.6	16.3
Tides + surges (Equation B2)	Obs[a]	17 months	7 months	Emspier	13	44	0.39/ $9.4{\cdot}10^{-5}$	1.8	7.2	2.1	7.6
				Pogum	4	48	0.36/ $8.5{\cdot}10^{-5}$	1.8	6.6	1.9	6.3
				Gandersum	8	48	0.41/ $1.4{\cdot}10^{-4}$	1.6	7.1	1.9	7.5
				Terborg	37	48	0.42/ $2.1{\cdot}10^{-4}$	1.1	11.4	1.5	14.8

Note. The missing percentage refers to the total percentage of missing and rejected data points. [a]Fortnightly averaged salinity at the Knock station for the seaward boundary condition. [b]Total of the fitted proportionality coefficients.

Table B2
Overview of the Model Experiments Performed for the Scheldt Estuary and Results Discussed in Section 4.3

Regression model	Boundary	Calibration	Prediction	Station (–)	Missing (%)	Pars[b] (–)	P/γ (–)	Calibration rmse (psu)	Calibration Un. var. (%)	Predictions rmse (psu)	Predictions Un. var. (%)
Tides + surges (Equation B1)	Obs[a]	2.5 years	1 year	Lillo Meetpaal	13	27	0.75/ $7.5{\cdot}10^{-4}$	0.5	1.5	0.64	3.9
				Oosterweel	25	29	0.57/ $4.0{\cdot}10^{-4}$	1.2	7.5	1.3	10.5
				Hemiksem	19	25	0.55/ $4.2{\cdot}10^{-4}$	0.6	11.8	0.56	13.0
Tides + surges (Equation B2)	Obs[a]	2.5 years	1 year	Lillo Meetpaal	13	36	0.75/ $7.2{\cdot}10^{-4}$	1.3	10.8	1.4	3.8
				Oosterweel	25	28	0.57/ $4.2{\cdot}10^{-4}$	1.4	11.1	1.6	9.8
				Hemiksem	19	34	0.56/ $4.3{\cdot}10^{-4}$	0.6	11.0	0.54	12.0

Note. The missing percentage refers to the total percentage of missing and rejected data points. [a]Fortnightly averaged salinity at Prosperpolder for the seaward boundary condition. [b]Total of the fitted proportionality coefficients.

described in Appendix A is applied. The results for the experiments performed in Section 4.3 using Equations B1 and B2 are tabulated in Tables B1 and B2 and provide comparable results.

Data Availability Statement

The salinity records used in this study for the Ems estuary were provided by Niedersächsischer Landesbetrieb für Wasserwirtschaft, Küsten- und Naturschutz (NLWKN). Data from the Dutch part of the Ems and Scheldt estuaries are publicly available at https://waterinfo.rws.nl, while salinity records for the Belgian part of the Scheldt can be downloaded from https://waterinfo.vlaanderen.be/Meetreeksen. In the data analysis, T-TIDE is used (Pawlowicz et al., 2002), a MATLAB toolbox for harmonic analysis, which can be downloaded from https://www-old.eoas.ubc.ca/~rich/. The data, code for the regression model, model setup for the experiments, and code to generate the figures are available from https://doi.org/10.5281/zenodo.14905063 (Van Keulen, 2025).

Acknowledgments
This project is part of the SALTISolutions Perspective Programme (2022/TTW/01344701P18-32 project4) and is funded by Dutch Research Council NWO, Rijkswaterstaat, Deltares, Port of Rotterdam, Hydrologic, Hoogheemraadschap van Schieland De Krimpenerwaard, Hoogheemraadshap Rijnland, and Royal BAM. The authors wish to acknowledge two reviewers for providing suggestion to improve the initial manuscript, Niedersächsischer Landesbetrieb für Wasserwirtschaft, Küsten- und Naturschutz (NLWKN) for providing the used measurements and V. Meulenberg for providing suggestions and feedback on the manuscript.

References

Brockway, R., Bowers, D., Hoguane, A., Dove, V., & Vassele, V. (2006). A note on salt intrusion in funnel-shaped estuaries: Application to the Incomati estuary, Mozambique. *Estuarine, Coastal and Shelf Science*, *66*(1–2), 1–5. https://doi.org/10.1016/j.ecss.2005.07.014

Burchard, H., Hetland, R. D., Schulz, E., & Schuttelaars, H. M. (2011). Drivers of residual estuarine circulation in tidally energetic estuaries: Straight and irrotational channels with parabolic cross section. *Journal of Physical Oceanography*, *41*(3), 548–570. https://doi.org/10.1175/2010jpo4453.1

Cai, H., Savenije, H. H., Zuo, S., Jiang, C., & Chua, V. P. (2015). A predictive model for salt intrusion in estuaries applied to the Yangtze estuary. *Journal of Hydrology*, *529*, 1336–1349. https://doi.org/10.1016/j.jhydrol.2015.08.050

Chatwin, P. (1976). Some remarks on the maintenance of the salinity distribution in estuaries. *Estuarine and Coastal Marine Science*, *4*(5), 555–566. https://doi.org/10.1016/0302-3524(76)90030-x

Chen, S.-N., & Sanford, L. P. (2009). Axial wind effects on stratification and longitudinal salt transport in an idealized, partially mixed estuary. *Journal of Physical Oceanography*, *39*(8), 1905–1920. https://doi.org/10.1175/2009jpo4016.1

Cho, K.-H., Wang, H. V., Shen, J., Valle-Levinson, A., & Teng, Y.-C. (2012). A modeling study on the response of Chesapeake Bay to hurricane events of Floyd and Isabel. *Ocean Modelling*, *49*, 22–46. https://doi.org/10.1016/j.ocemod.2012.02.005

Codiga, D. L. (2011). Unified tidal analysis and prediction using the UTide Matlab functions.

De Swart, H., De Jonge, V., & Vosbeek, M. (1997). Application of the tidal random walk model to calculate water dispersion coefficients in the ems estuary. *Estuarine, Coastal and Shelf Science*, *45*(1), 123–133. https://doi.org/10.1006/ecss.1996.0171

Evans, J., & Pugh, D. (1982). Analysing clipped sea-level records for harmonic tidal constituents. *International Hydrographic Review*. Retrieved from https://journals.lib.unb.ca/index.php/ihr/article/view/23551

Familkhalili, R., Talke, S. A., & Jay, D. A. (2020). Tide-storm surge interactions in highly altered estuaries: How channel deepening increases surge vulnerability. *Journal of Geophysical Research: Oceans*, *125*(4), e2019JC015286. https://doi.org/10.1029/2019jc015286

Fischer, H. B. (1972). Mass transport mechanisms in partially stratified estuaries. *Journal of Fluid Mechanics*, *53*(4), 671–687. https://doi.org/10.1017/s0022112072000412

Fischer, H. B. (1976). Mixing and dispersion in estuaries. *Annual Review of Fluid Mechanics*, *8*(1), 107–133. https://doi.org/10.1146/annurev.fl.08.010176.000543

Fischer, H. B., List, J. E., Koh, C. R., Imberger, J., & Brooks, N. H. (1979). *Mixing in inland and coastal waters*. Academic Press.

Foreman, M. G., Cherniawsky, J. Y., & Ballantyne, V. (2009). Versatile harmonic tidal analysis: Improvements and applications. *Journal of Atmospheric and Oceanic Technology*, *26*(4), 806–817. https://doi.org/10.1175/2008jtecho615.1

Gay, P. S., & O'Donnell, J. (2007). A simple advection-dispersion model for the salt distribution in linearly tapered estuaries. *Journal of Geophysical Research*, *112*(C7), C07021. https://doi.org/10.1029/2006jc003840

Geyer, W. R., & MacCready, P. (2014). The estuarine circulation. *Annual Review of Fluid Mechanics*, *46*(1), 175–197. https://doi.org/10.1146/annurev-fluid-010313-141302

Geyer, W. R., & Signell, R. P. (1992). A reassessment of the role of tidal dispersion in estuaries and bays. *Estuaries*, *15*(2), 97–108. https://doi.org/10.2307/1352684

Gisen, J. I. A., & Savenije, H. H. (2015). Estimating bankfull discharge and depth in ungauged estuaries. *Water Resources Research*, *51*(4), 2298–2316. https://doi.org/10.1002/2014wr016227

Gong, W., Chen, Y., Zhang, H., & Chen, Z. (2018). Effects of wave–current interaction on salt intrusion during a typhoon event in a highly stratified estuary. *Estuaries and Coasts*, *41*(7), 1904–1923. https://doi.org/10.1007/s12237-018-0393-8

Hansen, D. V., & Rattray, M., Jr. (1965). Gravitational circulation in straits and estuaries. *Journal of Marine Research*, *79*(2), 104–122. https://doi.org/10.1357/002224021834614399

Hansen, D. V., & Rattray, M., Jr. (1966). New dimensions in estuary classification 1. *Limnology and Oceanography*, *11*(3), 319–326. https://doi.org/10.4319/lo.1966.11.3.0319

Helder, W., & Ruardij, P. (1982). A one-dimensional mixing and flushing model of the ems-Dollard estuary: Calculation of time scales at different river discharges. *Netherlands Journal of Sea Research*, *15*(3–4), 293–312. https://doi.org/10.1016/0077-7579(82)90060-6

Hoitink, A., Wang, Z., Vermeulen, B., Huismans, Y., & Kästner, K. (2017). Tidal controls on river delta morphology. *Nature Geoscience*, *10*(9), 637–645. https://doi.org/10.1038/ngeo3000

Ippen, A. T., & Harleman, D. R. (1961). *One-dimensional analysis of salinity intrusion in estuaries. (Tech. Rep.)*. Army Engineer Waterways Experiment Station Vicksburg Ms.

Jay, D. A., & Musiak, J. D. (1994). Particle trapping in estuarine tidal flows. *Journal of Geophysical Research*, *99*(C10), 20445–20461. https://doi.org/10.1029/94jc00971

Jay, D. A., & Musiak, J. D. (1996). Internal tidal asymmetry in channel flows: Origins and consequences. *Coastal and Estuarine Studies*, 211–249. https://doi.org/10.1029/ce050p0211

Jay, D. A., & Smith, J. D. (1990). Circulation, density distribution and neap-spring transitions in the Columbia River estuary. *Progress in Oceanography*, *25*(1–4), 81–112. https://doi.org/10.1016/0079-6611(90)90004-l

Jongbloed, H., Schuttelaars, H. M., Dijkstra, Y. M., Donkers, P. B., & Hoitink, A. J. (2022). Influence of wind on subtidal salt intrusion and stratification in well-mixed and partially stratified estuaries. *Journal of Physical Oceanography*, *52*(12), 3139–3158. https://doi.org/10.1175/jpo-d-21-0291.1

Kranenburg, C. (1986). A time scale for long-term salt intrusion in well-mixed estuaries. *Journal of Physical Oceanography*, *16*(7), 1329–1331. https://doi.org/10.1175/1520-0485(1986)016<1329:atsflt>2.0.co;2

Kranenburg, W., Van der Kaaij, T., Tiessen, M., Friocourt, Y., & Blaas, M. (2022). Salt intrusion in the Rhine Meuse delta: Estuarine circulation, tidal dispersion or surge effect. In *Proceedings of the 39th IAHR world congress June* (pp. 5601–5608). Granada.

Kuijper, K., & Van Rijn, L. C. (2011). Analytical and numerical analysis of tides and salinities in estuaries; part ii: Salinity distributions in prismatic and convergent tidal channels. *Ocean Dynamics*, *61*(11), 1743–1765. https://doi.org/10.1007/s10236-011-0454-z

Li, C., Weeks, E., & Blanchard, B. W. (2010). Storm surge induced flux through multiple tidal passes of lake Pontchartrain estuary during hurricanes Gustav and Ike. *Estuarine, Coastal and Shelf Science*, *87*(4), 517–525. https://doi.org/10.1016/j.ecss.2010.02.003

Li, M., Zhong, L., Boicourt, W. C., Zhang, S., & Zhang, D.-L. (2006). Hurricane-induced storm surges, currents and destratification in a semi-enclosed bay. *Geophysical Research Letters*, *33*(2), L02604. https://doi.org/10.1029/2005gl024992

Li, M., Zhong, L., Boicourt, W. C., Zhang, S., & Zhang, D.-L. (2007). Hurricane-induced destratification and restratification in a partially-mixed estuary. *Journal of Marine Research*, *65*(2), 169–192. https://doi.org/10.1357/002224007780882550

MacCready, P. (1999). Estuarine adjustment to changes in river flow and tidal mixing. *Journal of Physical Oceanography*, *29*(4), 708–726. https://doi.org/10.1175/1520-0485(1999)029<0708:eatcir>2.0.co;2

MacCready, P., & Geyer, W. R. (2010). Advances in estuarine physics. *Annual Review of Marine Science*, *2*(1), 35–58. https://doi.org/10.1146/annurev-marine-120308-081015

Matte, P., Jay, D. A., & Zaron, E. D. (2013). Adaptation of classical tidal harmonic analysis to nonstationary tides, with application to river tides. *Journal of Atmospheric and Oceanic Technology*, *30*(3), 569–589. https://doi.org/10.1175/jtech-d-12-00016.1

Matte, P., Secretan, Y., & Morin, J. (2014). Temporal and spatial variability of tidal-fluvial dynamics in the St. Lawrence fluvial estuary: An application of nonstationary tidal harmonic analysis. *Journal of Geophysical Research: Oceans*, *119*(9), 5724–5744. https://doi.org/10.1002/2014jc009791

Monismith, S. G., Kimmerer, W., Burau, J. R., & Stacey, M. T. (2002). Structure and flow-induced variability of the subtidal salinity field in northern San Francisco bay. *Journal of Physical Oceanography*, *32*(11), 3003–3019. https://doi.org/10.1175/1520-0485(2002)032<3003:safivo>2.0.co;2

Nguyen, A. D., & Savenije, H. H. (2006). Salt intrusion in multi-channel estuaries: A case study in the Mekong delta, Vietnam. *Hydrology and Earth System Sciences*, *10*(5), 743–754. https://doi.org/10.5194/hess-10-743-2006

Nguyen, A. D., Savenije, H. H., van der Wegen, M., & Roelvink, D. (2008). New analytical equation for dispersion in estuaries with a distinct ebb-flood channel system. *Estuarine, Coastal and Shelf Science*, *79*(1), 7–16. https://doi.org/10.1016/j.ecss.2008.03.002

Pawlowicz, R., Beardsley, B., & Lentz, S. (2002). Classical tidal harmonic analysis including error estimates in Matlab using t_tide. *Computers & Geosciences*, *28*(8), 929–937. https://doi.org/10.1016/s0098-3004(02)00013-4

Prandle, D. (1981). Salinity intrusion in estuaries. *Journal of Physical Oceanography*, *11*(10), 1311–1324. https://doi.org/10.1175/1520-0485(1981)011<1311:siie>2.0.co;2

Pritchard, D. W. (1952). Estuarine hydrography. In *Advances in geophysics* (Vol. 1, pp. 243–280). Elsevier. https://doi.org/10.1016/s0065-2687(08)60208-3

Rath, J. S., Hutton, P. H., Chen, L., & Roy, S. B. (2017). A hybrid empirical-Bayesian artificial neural network model of salinity in the San Francisco bay-delta estuary. *Environmental Modelling & Software*, *93*, 193–208. https://doi.org/10.1016/j.envsoft.2017.03.022

Reyes-Merlo, M. Á., Díez-Minguito, M., Ortega-Sánchez, M., Baquerizo, A., & Losada, M. Á. (2013). On the relative influence of climate forcing agents on the saline intrusion in a well-mixed estuary: Medium-term Monte Carlo predictions. *Journal of Coastal Research*, *65*, 1200–1205. https://doi.org/10.2112/SI65-203.1/192534

Savenije, H. H. (1986). A one-dimensional model for salinity intrusion in alluvial estuaries. *Journal of Hydrology*, *85*(1–2), 87–109. https://doi.org/10.1016/0022-1694(86)90078-8

Savenije, H. H. (1989). Salt intrusion model for high-water slack, low-water slack, and mean tide on spread sheet. *Journal of Hydrology*, *107*(1–4), 9–18. https://doi.org/10.1016/0022-1694(89)90046-2

Savenije, H. H. (1993). Predictive model for salt intrusion in estuaries. *Journal of Hydrology*, *148*(1–4), 203–218. https://doi.org/10.1016/0022-1694(93)90260-g

Savenije, H. H. (2006). *Salinity and tides in alluvial estuaries*. Elsevier.

Savenije, H. H. (2015). Prediction in ungauged estuaries: An integrated theory. *Water Resources Research*, *51*(4), 2464–2476. https://doi.org/10.1002/2015wr016936

Savenije, H. H., Toffolon, M., Haas, J., & Veling, E. J. (2008). Analytical description of tidal dynamics in convergent estuaries. *Journal of Geophysical Research*, *113*(C10), C10025. https://doi.org/10.1029/2007jc004408

Scully, M. E., Friedrichs, C., & Brubaker, J. (2005). Control of estuarine stratification and mixing by wind-induced straining of the estuarine density field. *Estuaries*, *28*(3), 321–326. https://doi.org/10.1007/bf02693915

Simpson, J. H., Brown, J., Matthews, J., & Allen, G. (1990). Tidal straining, density currents, and stirring in the control of estuarine stratification. *Estuaries*, *13*(2), 125–132. https://doi.org/10.2307/1351581

Struyf, E., Van Damme, S., & Meire, P. (2004). Possible effects of climate change on estuarine nutrient fluxes: A case study in the highly nutrified Schelde estuary (Belgium, The Netherlands). *Estuarine, Coastal and Shelf Science*, *60*(4), 649–661. https://doi.org/10.1016/j.ecss.2004.03.004

Toffolon, M., Vignoli, G., & Tubino, M. (2006). Relevant parameters and finite amplitude effects in estuarine hydrodynamics. *Journal of Geophysical Research*, *111*(C10), C10014. https://doi.org/10.1029/2005jc003104

UNESCO. (1981). Tenth report of the joint panel on oceanographic tables and standards-Sidney, BC, Canada, 1-5 September 1980, sponsored by UNESCO, ICES, SCOR, IAPSO.

Van de Burgh, P. (1972). Ontwikkeling van een methode voor het voorspellen van zoutverdelingen in estuaria, kanalen en zeeen.

Van Keulen, D. (2025). Harmonic regression model for estuarine salinity at monitoring stations (0.1) [Dataset and Software]. *Zenodo*. https://doi.org/10.5281/zenodo.14905063

Winterwerp, J. C., Vroom, J., Wang, Z.-B., Krebs, M., Hendriks, E. C., van Maren, D. S., et al. (2017). Spm response to tide and river flow in the hyper-turbid ems river. *Ocean Dynamics*, *67*(5), 559–583. https://doi.org/10.1007/s10236-017-1043-6

Xu, Y., Hoitink, A. J., Zheng, J., Kästner, K., & Zhang, W. (2019). Analytical model captures intratidal variation in salinity in a convergent, well-mixed estuary. *Hydrology and Earth System Sciences*, *23*(10), 4309–4322. https://doi.org/10.5194/hess-23-4309-2019

Zhang, E., Savenije, H. H., Wu, H., Kong, Y., & Zhu, J. (2011). Analytical solution for salt intrusion in the Yangtze estuary, China. *Estuarine, Coastal and Shelf Science*, *91*(4), 492–501. https://doi.org/10.1016/j.ecss.2010.11.008

Zhang, Z., & Savenije, H. (2019). Maximum power of saline and fresh water mixing in estuaries. *Earth System Dynamics*, *10*(4), 667–684. https://doi.org/10.5194/esd-10-667-2019

Zhou, F., Liu, B., & Duan, K. (2020). Coupling wavelet transform and artificial neural network for forecasting estuarine salinity. *Journal of Hydrology*, *588*, 125127. https://doi.org/10.1016/j.jhydrol.2020.125127

Zhou, X., Yang, T., Shi, P., Yu, Z., Wang, X., & Li, Z. (2017). Prospective scenarios of the saltwater intrusion in an estuary under climate change context using Bayesian neural networks. *Stochastic Environmental Research and Risk Assessment*, *31*(4), 981–991. https://doi.org/10.1007/s00477-017-1399-7