Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Bayesian modelling of salinity profiles in the Volturno River mouth under depth dependence

Agostino Gnasso

agostino.gnasso@unina.it

University of Naples Federico II

Lucia Rita Pacifico

National Research Council

Raffaele Mattera

University of Campania "Luigi Vanvitelli"

Raffaele D'Adamo

National Research Council

Fabio Matano

National Research Council

Germana Scepi

University of Naples Federico II

Research Article

Keywords: Saltwater intrusion, Bayesian hierarchical modelling, errors-in-variables regression, Gaussian process, depth-resolved profiles, Volturno River, measurement error

Posted Date: June 9th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-9859471/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Version of Record: A version of this preprint was published at Stochastic Environmental Research and Risk Assessment on October 21st, 2026. See the published version at https://doi.org/10.1007/s00477-026-03390-0.

Bayesian modelling of salinity profiles in the Volturno River mouth under depth dependence

Agostino Gnasso[1*], Lucia Rita Pacifico[2], Raffaele Mattera[3], Raffaele D'Adamo[2], Fabio Matano[2], Germana Scepi[1]

[1*]Department of Economics and Statistics, University of Naples "Federico II", Naples, 80126, Italy.
[2]National Research Council, Institute of Marine Science (CNR-ISMAR), Naples, 80133, Italy.
[3]Department of Mathematics and Physics, University of Campania "Luigi Vanvitelli", Caserta, 81100, Italy.

*Corresponding author(s). E-mail(s): agostino.gnasso@unina.it;

Abstract

Saltwater intrusion causes a recurrent environmental hazard for Mediterranean coastal systems with severe economic risks, and understanding which physico-chemical variables are associated with salinity is essential for supporting monitoring strategies. Stimating these associations from depth-resolved field profiles requires accounting for two key issues: uncertainty in the salinity and covariate measurements after depth-level aggregation, and residual vertical dependence along the water column. We address these issues by proposing a Bayesian errors-in-variables model augmented with a Gaussian Process component for residual depth-dependent structure. The model is applied to two synoptic profiling campaigns in the Volturno River mouth (Italy), carried out under contrasting winter high-discharge and summer low-discharge conditions. Temperature and dissolved oxygen emerge as the most robust physico-chemical variables associated with salinity, while the residual depth structure is negligible in winter and substantial in summer. A simulation study calibrated on the empirical setting, including scenarios with and without residual depth dependence, shows that the model improves the calibration of parameter uncertainty and reduces misleading coefficient estimates relative to simpler specifications. Cross-month prediction further shows that the full model transfers across seasons better than the baseline.

Keywords: Saltwater intrusion, Bayesian hierarchical modelling, errors-in-variables regression, Gaussian process, depth-resolved profiles, Volturno River, measurement error

1 Introduction

Saltwater intrusion is the inland penetration of denser marine water beneath the lighter freshwater outflow in river estuaries and in coastal groundwater. In Mediterranean basins, where river discharge varies strongly between seasons, the extent of the salt wedge can shift by several kilometres over timescales of weeks (Savenije 2005). Identifying which physico-chemical variables co-vary with salinity is relevant both for process-based modelling and for the design of monitoring networks (Zhou et al. 2017).

The Volturno River originates in the Apennine mountains and flows for about 175 km before discharging into the Tyrrhenian Sea at Castel Volturno (41.02° N, 13.92° E) (see Figure 1). The lower reach is tidally influenced over the last 8 to 10 km, and the river mouth shows a stratified two-layer structure during low-discharge periods: denser marine water (salinity of about 37 PSU) intrudes along the bottom while fresh river water flows seaward at the surface. The depth of the halocline and the landward extent of the salt wedge vary seasonally in response to river discharge (Sorgente et al. 2020; Gaiolini et al. 2023). Saltwater intrusion in the Volturno coastal area has been studied through hydrodynamic modelling (Gaiolini et al. 2023), vulnerability indices (Busico et al. 2021) and geophysical methods (Tarallo et al. 2023), but a statistical characterisation of how vertical salinity variability is associated with the physico-chemical covariates measured by multi-parameter probes has not been attempted. We emphasise that the term *characterisation* is used here in its statistical-association sense and not in the causal-attribution sense familiar from climate-attribution studies: the analysis quantifies conditional associations under a flexible Bayesian model, not causal effects.

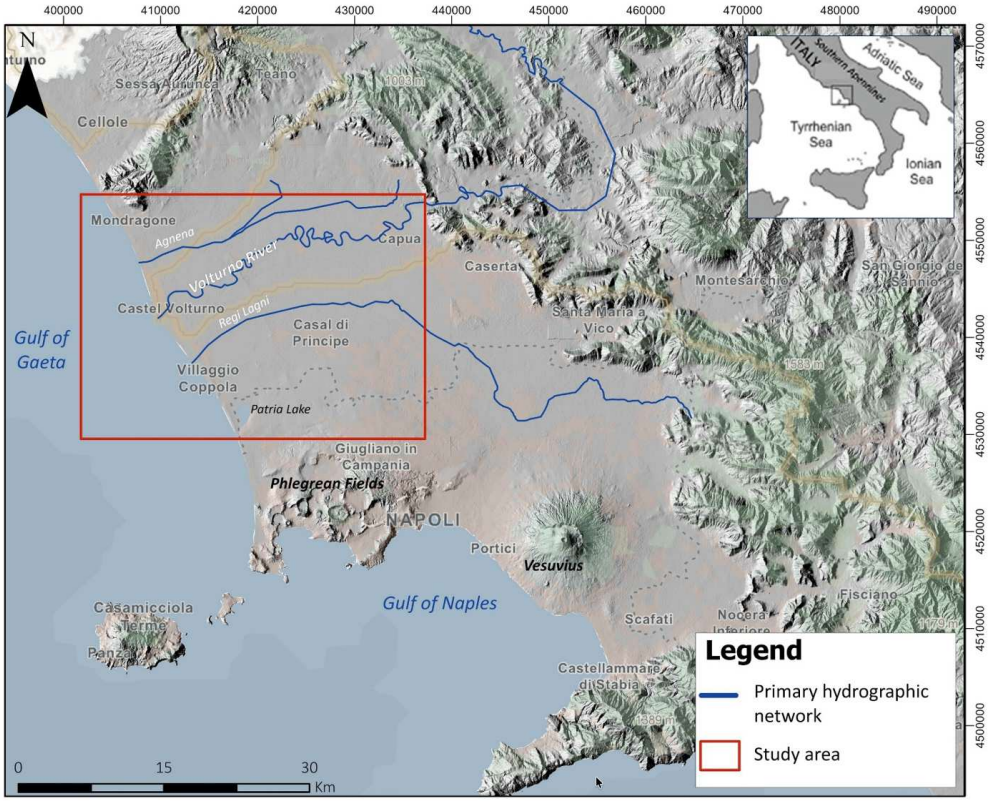


Fig. 1: Location of the Volturno River mouth study area in Campania, southern Italy.

From a geological and geomorphological perspective, the Volturno mouth belongs to the Volturno alluvial-coastal plain, one of the major low-lying coastal plains of the eastern Tyrrhenian Sea coast. This plain developed during the Quaternary under the combined influence of tectonic subsidence, sea-level fluctuations, fluvial sedimentation, coastal dynamics, and volcanic inputs from the Phlegraean Field volcanic district (Corrado et al. 2020; Buffardi et al. 2021). Its shallow subsurface is characterised by alluvial, coastal, lagoonal, and marsh deposits overlying older volcanic and volcaniclastic units, including the Campanian Ignimbrite. This stratigraphic and geomorphological setting favours the development of low-gradient river reaches and shallow coastal aquifers that are particularly sensitive to marine ingression, salinisation, and changes in freshwater discharge. In order to investigate the variability of the main chemical and physical

parameters, particularly salinity, that characterize the transition between freshwater river water and coastal seawater, a series of survey campaigns were conducted along the mouth of the Volturno River.

Multi-parameter probe surveys provide vertical salinity profiles together with temperature, dissolved oxygen, pH, and phytoplankton-related fluorescence at multiple stations along the longitudinal gradient of the river mouth. These observations allow the salt-wedge structure to be analysed not only through salinity, but also through associated physico-chemical variables that may act as tracers of water-mass contrasts. However, such variables should not be interpreted as direct causal controls on salinity; rather, they co-vary with salinity because they reflect the position of the sampled water parcel within the marine-to-freshwater transition layer. Beyond process understanding, the same physico-chemical information is the empirical basis for environmental risk assessment in river mouth coastal-aquifer settings. The vulnerability literature cited above relies, at its lowest level, on a statistical layer that connects sensor readings to salinity, and the credibility of any downstream risk product (vulnerability map, hazard alarm, monitoring-network design) is bounded by the calibration of that layer.

The objective of this paper is twofold. First, we identify which measured physico-chemical variables are statistically associated with salinity in the Volturno river mouth. Second, we assess whether and how the estimated relationships change between contrasting seasons, and we define season-specific models that can be transferred across seasons by cross-month prediction. Addressing both objectives requires a statistical model that accounts for two properties of the data. First, each bin-level observation is measured with uncertainty, obtained by aggregating several readings from multiple sensors. The associated standard error is non-negligible and must be included in the statistical model to avoid overconfident inference on the regression coefficients (Kim et al. 2016). Second, salinity profiles are also vertically structured, because the halocline between saline bottom water and fresher surface water induces smooth

dependence across neighbouring depth bins. Treating these bins as conditionally independent can therefore underestimate uncertainty and confound covariate effects with unexplained vertical structure.

To model salinity profiles accounting for these two properties, we propose a Bayesian model with error-in-variables. The Bayesian framework is useful in this setting because it provides full posterior uncertainty for the regression coefficients and hyperparameters, and because observational uncertainty and residual depth dependence can be embedded directly into the generative model (Carroll et al. 2006; Rasmussen and Williams 2006; Zheng and Han 2016). The two ingredients of the proposal both have an established literature. Bayesian treatment of measurement error in regression dates back to Richardson and Gilks (1993) and has since been extended to nonlinear and semiparametric settings (Berry et al. 2002; Carroll et al. 2006), with more recent contributions integrating it into general latent-Gaussian inference engines (Muff et al. 2015). Gaussian processes for environmental and oceanographic profile data have been used both for locally stationary interpolation of Argo float profiles (Kuusela and Stein 2018) and for scalable spatio-temporal modelling of large geophysical datasets (e.g. Banerjee et al. 2008; Colombo et al. 2025). Our contribution combines the two strands inside a single hierarchical specification tailored to depth-resolved estuarine profiles. An errors-in-variables layer propagates the bin-level uncertainty of both salinity and the covariates into the likelihood, rather than treating the depth-bin averages as error-free quantities. A Gaussian process with an exponential covariance kernel represents the residual depth dependence after conditioning on the physico-chemical covariates. The resulting model separates three components of variation: the covariate to salinity association of scientific interest, the uncertainty induced by bin-level observation and aggregation, and the smooth vertical structure not explained by the measured covariates.

We assess the usefulness of this proposal by comparing the full model with four restricted models, obtained by crossing the presence or absence of the errors-in-variables layer with the presence or absence of the Gaussian process depth-dependence layer. The comparison uses two criteria: within-campaign model adequacy, assessing how well each model describes the survey used for fitting, and cross-month predictive transferability, assessing whether the estimated covariate to salinity relationship remains informative under contrasting seasonal conditions. The field-data results show that the full specification provides a better in-sample representation of the salinity profiles and improved cross-month predictive performance relative to simpler models. A simulation study, calibrated on the empirical setting, supports this conclusion. Empirically, temperature and dissolved oxygen emerge as the most robust covariates associated with salinity, while residual depth dependence is substantial in the summer low-discharge campaign but negligible in the winter high-discharge campaign, consistently with a deeper and more vertically structured summer salt wedge.

The paper is organised as follows. Section 2 describes the field campaigns and the aggregation to depth bins. Section 3 presents the statistical model, the four restricted specifications and the cross-month prediction framework. Section 4 reports the simulation study. Section 5 reports the results on the field data and discusses the physical and methodological implications. Section 6 concludes with final remarks and a discussion on the current work limitations and possible future research directions.

2 Field campaigns and salinity data processing

Two synoptic monitoring campaigns were conducted within the framework of the SAL.WE. (SAline WEdge intrusion) project to characterise the vertical and longitudinal structure of saltwater intrusion along the terminal reach of the Volturno River under contrasting seasonal conditions. The winter survey (Jan-26) was carried out on

16 January 2026 under high-discharge conditions. The summer survey (Jul-25) was carried out in July 2025 under low-discharge conditions.

In Jan-26, 13 stations (FV-C01, FV-C0, FV-Foce Volturno, FV-01 through FV-10) were deployed along the river axis from the river mouth to 8.0 km inland (Table 1). The deepest station reached 10.0 m. The survey yielded 488 individual probe readings across 13 stations. In Jul-25, 20 stations were deployed over a substantially longer reach, from the river mouth to 44 km inland (FV-Capua). The additional upstream stations (FV-CeA, FV-Grazzanise, FV-Capua) served as freshwater reference points and lie well beyond the limit of marine influence. The survey yielded 443 readings.

At each station the probe recorded depth, temperature (°C), conductivity, salinity (PSU, Practical Salinity Unit), dissolved oxygen (percentage and mg kg^{-1}), pH, chlorophyll-a fluorescence (Trx-ChlA), phycocyanin, phycoerythrin and total dissolved solids. Distance from the most seaward station was computed as distance measured along the river course from FV-C01 (see Table 1). The spatial distribution of the

Table 1: Station coordinates and distances from the most seaward station (FV-C01). Jan-26 stations are those from FV-C01 to FV-10. Stations FV-11 through FV-Capua appear only in Jul-25.

Station	Longitude	Latitude	Distance (km)	Campaign
FV-C01	13.9085	41.0216	0.0	Both
FV-C0	13.9184	41.0242	2.7	Both
FV-Foce Volturno	13.9200	41.0200	3.0	Both
FV-01 to FV-10	see profile axis		3.5 to 8.0	Both
FV-11 to FV-16	see profile axis		5.5 to 10.0	Jul-25 only
FV-CeA	14.0287	41.0747	13.0	Jul-25 only
FV-Grazzanise	14.1055	41.0927	27.0	Jul-25 only
FV-Capua	14.2091	41.1094	44.0	Jul-25 only

sampling stations is shown in Figure 2.

2.1 Seasonal salinity contrast

The two campaigns show markedly different salinity regimes (see Table 2). In Jan-26,

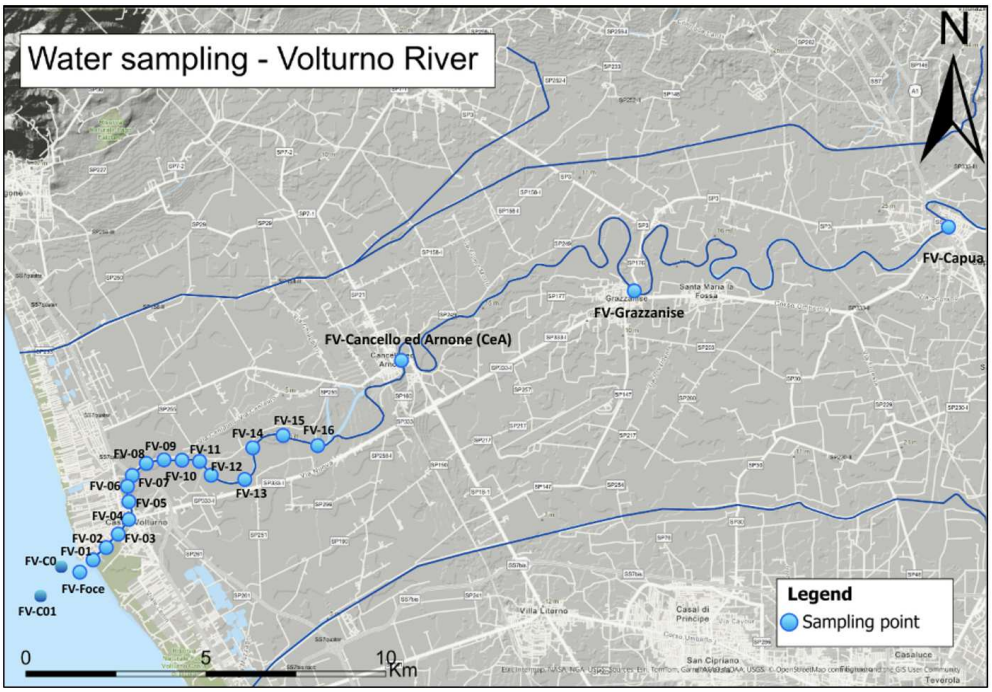


Fig. 2: Spatial distribution of CTD sampling stations along the terminal reach of the Volturno River, from the coastal sector and river mouth to the inner fluvial sector. Station coordinates, distances from FV-C01, and campaign occurrence are reported in Table 1.

Table 2: Descriptive statistics for the two campaigns. Sal: salinity (PSU). Temp: temperature (°C). O_2: dissolved oxygen (mg kg^{-1}).

Campaign	n	Sal mean	Sal sd	Sal median	Temp mean	O_2 mean
Jan-26	488	22.4	17.1	35.4	12.7	10.1
Jul-25	443	20.0	16.9	25.4	26.7	7.9

the salt wedge does not extend beyond 6.5 km from the mouth: stations FV-07 through FV-10 record a mean salinity of about 0.36 PSU, essentially fresh water. The transition zone (stations FV-01 to FV-06, 3.5 to 6.0 km) spans a full range from 2 to 37 PSU over less than 3 km, which indicates a sharp halocline. In Jul-25, salinity of 18.6 PSU is still recorded at 5 km (FV-10) and values above 1 PSU persist up to 9 km (FV-15), consistent with the deeper summer intrusion driven by reduced freshwater discharge.

The far-upstream stations (FV-CeA at 13 km, FV-Grazzanise at 27 km, FV-Capua at 44 km) confirm that marine influence does not reach beyond 10 km in summer, with all three stations recording salinities below 0.42 PSU. Temperature reflects the seasonal forcing: mean probe temperature in Jul-25 is 26.7°C against 12.7°C in Jan-26, with dissolved oxygen lower in summer (7.9 mg kg^{-1}) than in winter (10.1 mg kg^{-1}).

Figure 3 shows the raw vertical probe profiles of salinity, temperature and dissolved oxygen for a representative subset of stations in each campaign. The sharp January halocline and the deeper, more gradual July transition are visible at the level of the individual probe readings, before any binning or modelling. The shift in temperature regime between the two campaigns (winter cluster around $12 - 15$°C, summer cluster around $25 - 27$°C) is the dominant feature of the temperature panels and confirms that the two surveys sample distinct hydrological regimes.

Salinity is reported on the practical salinity scale commonly used for CTD-derived seawater properties (Fofonoff and Millard 1983). For interpretative purposes, salinity values were grouped into three broad classes: freshwater conditions (< 2 PSU), brackish/transitional waters (2 to 30 PSU), and saltwater conditions (> 30 PSU). The term transitional waters is used consistently with the Water Framework Directive, which defines them as surface waters near river mouths that are partly saline due to the proximity of coastal waters but substantially influenced by freshwater flows (European Parliament and Council 2000).

2.2 Longitudinal salinity sections

Longitudinal salinity sections for the summer and winter campaigns were reconstructed from the CTD profiles using Ocean Data View (ODV; Schlitzer 2022), allowing salinity to be interpolated and visualised along the depth–distance transect. Longitudinal salinity sections reconstructed from CTD profiles are shown in Figure 4. In the summer section, shown in Fig. 4a, saline and brackish waters extend farther

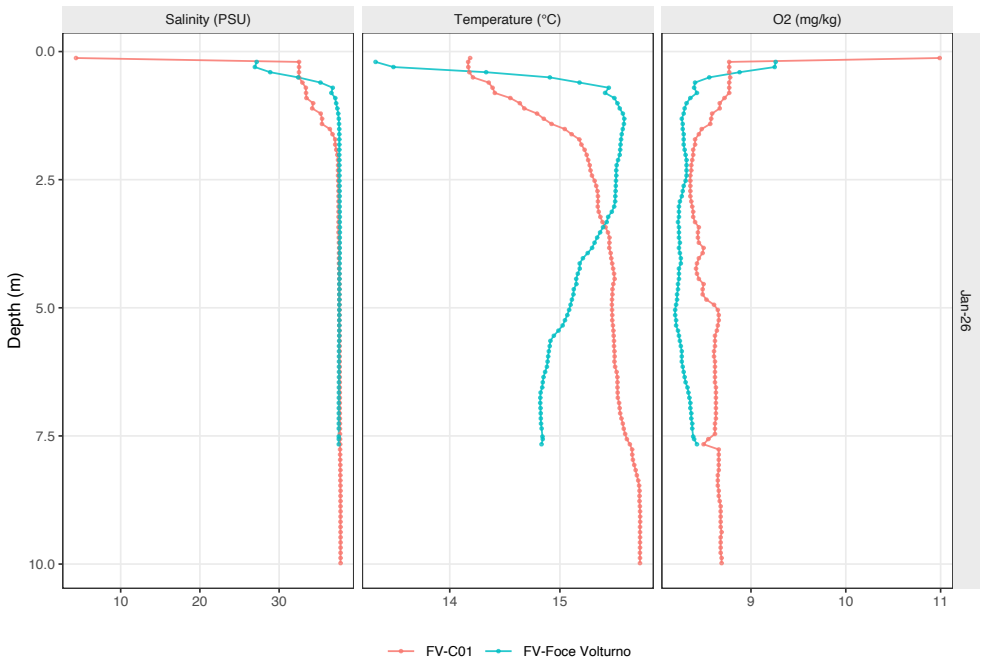


Fig. 3: Raw vertical probe profiles for a representative subset of stations in each campaign. Rows: Jan-26 (top) and Jul-25 (bottom). Columns: salinity, temperature and dissolved oxygen. Each line corresponds to one station; the y axis is depth (m, downward).

inland, indicating a more developed salt wedge during low-flow conditions. Saltwater-dominated conditions occur near the river mouth and in the adjacent coastal sector, while brackish waters progressively attenuate upstream before reaching freshwater conditions in the inner fluvial sector. This pattern highlights the landward propagation of marine influence during the summer campaign and the gradual transition from marine-dominated to freshwater conditions along the river axis. In the winter section, shown in Fig. 4b, the saline intrusion is more spatially restricted. The transition from saltwater or brackish conditions to freshwater occurs closer to the river mouth, consistent with the higher freshwater discharge during the January campaign. Compared with the summer section, the winter profile shows a shorter inland extent of brackish waters and a more limited vertical development of the saline wedge along the investigated

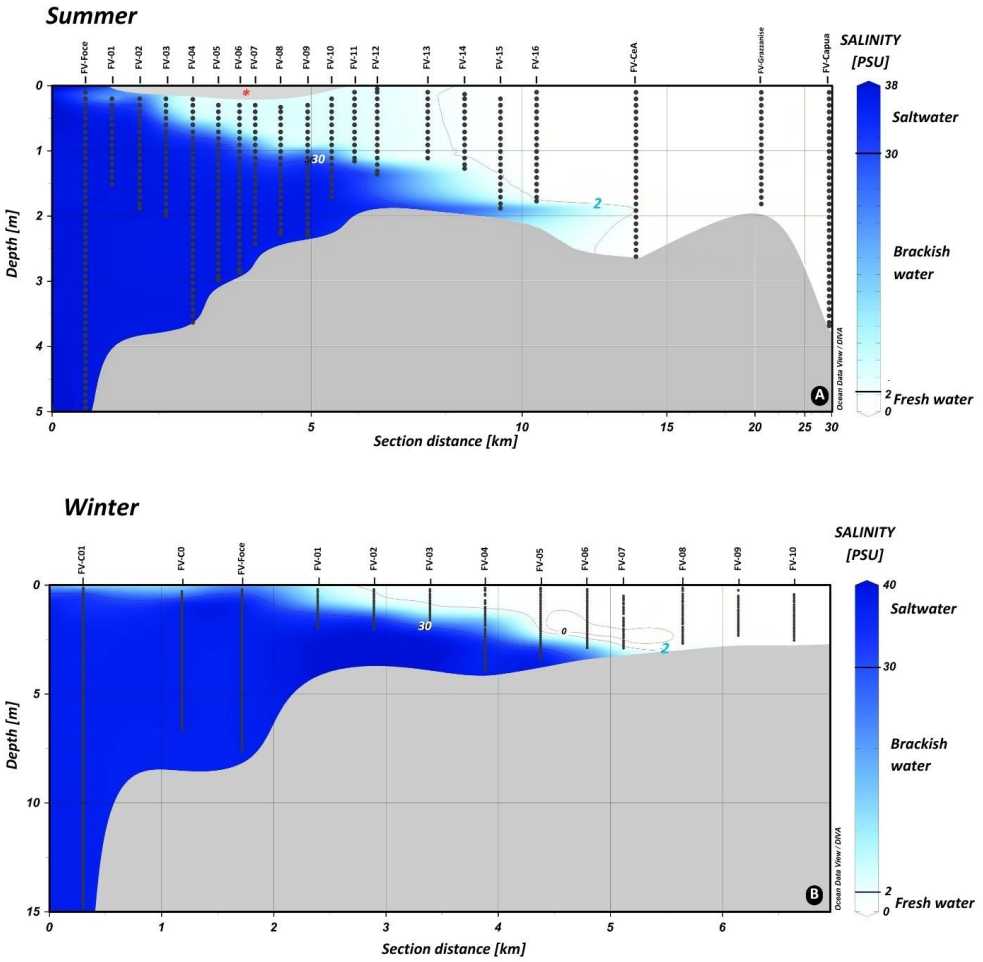


Fig. 4: Longitudinal salinity sections reconstructed from CTD profiles along the terminal reach of the Volturno River: (a) summer campaign, July 2025 (top); (b) winter campaign, January 2026 (bottom). Colours indicate salinity values; black dots mark CTD measurements along the water column. The red asterisk in panel (a) indicates a NoData area excluded from interpolation. Section distances are not in scale.

reach. Therefore, the comparison between Fig. 4a and Fig. 4b documents the seasonal displacement of the salt wedge and the contrasting salinity structure between low-flow summer and higher-discharge winter conditions. These longitudinal sections provide the field-based reference for the subsequent statistical analysis of salinity variability.

2.3 Depth-bin aggregation

Individual probe readings are aggregated into adaptive depth bins before modelling. Starting from a base resolution of 1 cm, bins are widened by merging adjacent intervals until at least three distinct probe stations contribute to each bin. The within-bin standard error is therefore well defined at every depth level, and shallower depths, where more probes overlap, receive narrower and more informative bins than deeper depths.

To be more precise, let $s = 1, \ldots, S$ denote the depth bins obtained after aggregation, where S is the number of available bins in a given campaign, and let $j = 1, \ldots, J_s$ index the probe readings contributing to bin s. Within each bin, the sample mean and standard error of salinity and of each covariate are computed across the contributing readings, yielding observed values $\bar{y}_s$, $\bar{x}_{sk}$ and corresponding standard errors $\tau_{s\bar{y}}$, $\tau_{sk\bar{x}}$. More specifically, let y_{sj} denote the salinity reading associated with reading j in depth bin s. The observed bin-level salinity is defined as

$$\bar{y}_s = \frac{1}{J_s} \sum_{j=1}^{J_s} y_{sj}, \tag{1}$$

and the corresponding bin-level uncertainty is estimated by the empirical standard error

$$\tau_{s\bar{y}} = \frac{1}{\sqrt{J_s}} \left[\frac{1}{J_s - 1} \sum_{j=1}^{J_s} (y_{sj} - \bar{y}_s)^2 \right]^{1/2}. \tag{2}$$

Similarly, for each covariate $k = 1, \ldots, K$, let x_{sjk} denote the corresponding covariate reading associated with reading j in depth bin s. The observed bin-mean covariate value is $\bar{x}_{sk} = \frac{1}{J_s} \sum_{j=1}^{J_s} x_{sjk}$, with empirical standard error $\tau_{sk\bar{x}} = \frac{1}{\sqrt{J_s}} \left[\frac{1}{J_s - 1} \sum_{j=1}^{J_s} (x_{sjk} - \bar{x}_{sk})^2 \right]^{1/2}$. The K observed bin-level covariates are collected in the vector $\bar{\mathbf{x}}_s = (\bar{x}_{s1}, \ldots, \bar{x}_{sK})^\top$.

The procedure yields 68 bins for Jan-26 (depth range 0.13 to 6.66 m) and 43 bins for Jul-25. The average coefficient of variation of bin-level salinity ($\tau_{s\bar{y}}/|\bar{y}_s| \times 100$) is about 16 per cent across Jan-26 bins, which confirms that within-bin variability is non-negligible and justifies explicit modelling of uncertainty.

3 Statistical model

The modelling objective is to estimate the association between depth-resolved salinity and the physico-chemical covariates introduced in Section 2.3, while propagating the uncertainty induced by the depth-bin aggregation. We assume that the observed bin-level quantities $\bar{y}_s$ and $\bar{x}_{sk}$ are noisy summaries of latent salinity and covariate values, denoted by y_s^* and x_{sk}^*, respectively. Let us denote the observation process as

$$\bar{x}_{sk} \mid x_{sk}^* \sim \mathcal{N}\big(x_{sk}^*, \tau_{sk\bar{x}}^2\big)\,, \qquad \bar{y}_s \mid y_s^* \sim \mathcal{N}\big(y_s^*, \tau_{s\bar{y}}^2\big)\,, \tag{3}$$

where $\tau_{sk\bar{x}}$ and $\tau_{s\bar{y}}$ are empirical bin-level uncertainty estimates. The regression relationship of scientific interest is specified at the latent level, that is,

$$y_s^* = \sum_{k=1}^{K} \beta_k\, x_{sk}^* + u_s, \tag{4}$$

where β_k is the conditional association between the covariate k, defined as

$$x_{sk}^* \mid \mu_k, \sigma_{x,k} \sim \mathcal{N}\big(\mu_k, \sigma_{x,k}^2\big)\,, \tag{5}$$

and the salinity, while u_s represents residual depth-dependent structure not explained by the covariates.

We then remark that salinity profiles display vertical correlation. Therefore, after conditioning on the measured physico-chemical covariates, we have that neighbouring

depth bins s and t may remain correlated as a consequence of the smooth transition between fresher surface water and saline bottom water. To account for this, we model the residual vector $\mathbf{u} = (u_1, \ldots, u_S)^\top$ as a Gaussian process,

$$\mathbf{u} \sim \mathcal{N}_S(\mathbf{0}, \Sigma(\boldsymbol{\theta})), \tag{6}$$

with covariance matrix $\Sigma(\boldsymbol{\theta})$ whose generic element is

$$\sigma_{ss'}(\boldsymbol{\theta}) = \sigma_u^2 \exp\left(-\frac{|d_s - d_{s'}|}{\ell}\right), \tag{7}$$

where $\boldsymbol{\theta} = (\sigma_u, \ell)$ and $d_s \in [0, 1]$ is the standardised midpoint depth of bin s, obtained by dividing the physical bin midpoint z_s (in metres) by the maximum sampled depth $z_{\max}$ of the corresponding campaign. The parameter σ_u controls the amplitude of the residual depth-dependent component, and ℓ is the depth correlation length-scale. We emphasise that ℓ is expressed in the standardised $[0, 1]$ depth coordinate and not in metres: to recover the physical correlation length, ℓ must be multiplied by $z_{\max}$, which differs between the two campaigns ($z_{\max} = 6.66$ m in Jan-26 and $z_{\max} = 10.00$ m in Jul-25, see Section 2.3). Consequently, a numerically similar posterior length-scale on the two campaigns (we report 0.24 and 0.29 in Section 5.1) corresponds to physical correlation lengths of roughly 1.6 and 2.9 m, respectively, a non-trivial difference that should be kept in mind when comparing the kernel across regimes. Residual salinity deviations at nearby depths are a priori more strongly correlated than deviations at distant depths, with the strength of this dependence decreasing exponentially as the vertical separation increases.

Combining the components (3)–(6) gives the full model

$$
\begin{aligned}
&\bar{y}_s \mid y_s^* \sim \mathcal{N}(y_s^*, \tau_{s\bar{y}}^2), \\
&y_s^* = \sum_{k=1}^{K} \beta_k \, x_{sk}^* + u_s, \\
&x_{sk}^* \mid \mu_k, \sigma_{x,k} \sim \mathcal{N}(\mu_k, \sigma_{x,k}^2), \\
&\bar{x}_{sk} \mid x_{sk}^* \sim \mathcal{N}(x_{sk}^*, \tau_{sk\bar{x}}^2), \\
&\mathbf{u} \sim \mathcal{N}_S(\mathbf{0}, \Sigma(\boldsymbol{\theta})).
\end{aligned}
\tag{8}
$$

For computation we marginalise the Gaussian process residual. Let $D_y = \mathrm{diag}(\tau_{1\bar{y}}^2, \ldots, \tau_{S\bar{y}}^2)$ denote the diagonal matrix collecting the bin-level uncertainty variances for salinity. The observation equation for the salinity profile can be written in vector form as $\bar{\mathbf{y}} = \mathbf{y}^* + \boldsymbol{\varepsilon}$, where $\boldsymbol{\varepsilon} \sim \mathcal{N}_S(\mathbf{0}, D_y)$ represents the bin-level observational uncertainty. According to (8), the latent salinity profile satisfies $\mathbf{y}^* = \boldsymbol{m} + \mathbf{u}$, where $\boldsymbol{m}$ is the regression mean vector with generic element $m_s = \sum_{k=1}^{K} \beta_k x_{sk}^*$ and $\mathbf{u} \sim \mathcal{N}_S(\mathbf{0}, \Sigma(\boldsymbol{\theta}))$ is the Gaussian process residual component. Combining the two layers yields $\bar{\mathbf{y}} = \boldsymbol{m} + \mathbf{u} + \boldsymbol{\varepsilon}$. Assuming independence between the Gaussian process residual and the observational uncertainty term, the sum $\mathbf{u}+\boldsymbol{\varepsilon}$ is Gaussian with covariance matrix $\Sigma(\boldsymbol{\theta}) + D_y$. Therefore, integrating out the latent residual process $\mathbf{u}$ gives the marginal likelihood

$$
\bar{\mathbf{y}} \mid x_{sk}^*, \boldsymbol{\beta}, \sigma_u, \ell \sim \mathcal{N}_S\left(\boldsymbol{m}, \Sigma(\boldsymbol{\theta}) + D_y\right). \tag{9}
$$

This marginal representation separates two sources of residual variation, that is, the smooth depth-dependent component $\Sigma(\boldsymbol{\theta})$ capturing the residual vertical structure not explained by the covariates, and the diagonal heteroskedastic component D_y, propagating the bin-level uncertainty in salinity.

All variables are standardised to zero mean and unit variance before fitting. As a consequence, the intercept is absorbed by the standardisation and is set to zero by construction. Priors are specified on the standardised scale. Regression coefficients are assigned weakly regularising Gaussian priors, $\beta_k \sim \mathcal{N}(0, 0.5^2)$, so that large effects remain possible while extreme coefficient values are discouraged. The population means of the latent covariates are assigned $\mu_k \sim \mathcal{N}(0, 1)$, and the corresponding scale parameters are assigned half-normal priors, $\sigma_{x,k} \sim \mathcal{HN}(0, 1)$. The amplitude of the residual depth-dependent component is assigned a half-normal prior, $\sigma_u \sim \mathcal{HN}(0, 1)$. The Gaussian process length-scale is assigned a log-normal prior, $\ell \sim \text{LogNormal}(\log 0.2, 0.5^2)$, centred at 0.2 on the standardised $[0, 1]$ depth axis. Because ℓ is defined on the rescaled depth, this prior expresses the same a priori statement on the fraction of the sampled depth range over which residual correlation persists, in both campaigns. In physical units this corresponds to an exponential correlation of $e^{-1} \approx 0.37$ at a vertical separation of about 1.3 m for Jan-26 (where $z_{\max} = 6.66$ m) and about 2.0 m for Jul-25 (where $z_{\max} = 10.00$ m), values that bracket the thickness of the halocline transition observed in each profile. The same prior is therefore neutral with respect to the seasonal difference in physical correlation length, and is used unchanged across the two campaigns to preserve comparability.

Let $\boldsymbol{\Theta}$ denote the full set of unknown quantities, including the regression coefficients $\boldsymbol{\beta}$, the latent covariates x^*_{sk}, the covariate-level parameters $(\mu_k, \sigma_{x,k})$, and the Gaussian process hyperparameters (σ_u, ℓ). Bayesian inference is based on the joint posterior distribution

$$
\begin{aligned}
&p(\boldsymbol{\Theta} \mid \bar{\mathbf{y}}, \bar{\mathbf{X}}) \propto \\
&p(\bar{\mathbf{y}} \mid x^*_{sk}, \boldsymbol{\beta}, \sigma_u, \ell) \prod_{s=1}^{S} \prod_{k=1}^{K} p(\bar{x}_{sk} \mid x^*_{sk}) p(x^*_{sk} \mid \mu_k, \sigma_{x,k}) p(\boldsymbol{\beta}) p(\boldsymbol{\mu}) p(\boldsymbol{\sigma}_x) p(\sigma_u) p(\ell),
\end{aligned}
\tag{10}
$$

where $p(\bar{\mathbf{y}} \mid x^*_{sk}, \boldsymbol{\beta}, \sigma_u, \ell)$ is the marginal Gaussian likelihood in (9). Posterior samples from this distribution are used to obtain coefficient summaries, credible intervals, posterior probabilities of physically expected directions, Gaussian process hyperparameter estimates and posterior predictive distributions of salinity profiles.

We use the Cholesky factorisation in all multivariate Gaussian computations to avoid explicit matrix inversion and improve numerical stability. A non-centred parameterisation is adopted for the latent covariates to improve Hamiltonian Monte Carlo mixing. Posterior inference is carried out using the No-U-Turn sampler (Hoffman and Gelman 2014) implemented in Stan version 2.32.2 (Carpenter et al. 2017), accessed from R through rstan version 2.32.7. The statistical relevance of each covariate is assessed from the posterior distribution of its regression coefficient. On the standardised scale, β_k quantifies the conditional association between covariate k and salinity after accounting for the other covariates, bin-level observational uncertainty and residual depth dependence. For each coefficient we report the posterior median, the 95 per cent credible interval, and the posterior probability of the expected physical direction. We compute $\Pr(\beta_k > 0)$ for covariates expected to increase with salinity (temperature and pH) and $\Pr(\beta_k < 0)$ for covariates expected to decrease with salinity (dissolved oxygen and the phytoplankton-related fluorescence variables).

To evaluate the contribution of the two main modelling components, namely the errors-in-variables layer for the covariates and the Gaussian process depth-dependence layer, we consider a 2 by 2 factorial set of restricted models. Each reduced model is obtained by imposing a restriction on the full model $\mathrm{M}_{\mathrm{full}}$ in (8), so that differences in fit and predictive performance can be attributed to the inclusion or exclusion of a specific modelling assumption.The model $\mathrm{M}_{\mathrm{noGP}}$ removes the Gaussian process depth-dependence layer by imposing $\sigma_u \equiv 0$, $\Sigma(\boldsymbol{\theta}) \equiv \mathbf{0}$ and $u_s \equiv 0$. The covariate errors-in-variables layer is retained, but residual dependence across depth bins is excluded. The marginal distribution becomes $\bar{\mathbf{y}} \mid \{x^*_{sk}\}, \boldsymbol{\beta} \sim \mathcal{N}_S\big(\boldsymbol{m}^{\mathrm{noGP}}, D_y\big)$,

with $m_s^{\mathrm{noGP}} = \sum_{k=1}^{K} \beta_k \, x_{sk}^*$, so that conditional on the latent covariates the salinity observations are independent across depth bins. The model $\mathrm{M}_{\mathrm{noEIV}}$ removes the errors-in-variables layer by setting $x_{sk}^* \equiv \bar{x}_{sk}$, while retaining the Gaussian process residual. The baseline model $\mathrm{M}_{\mathrm{base}}$ removes both components and reduces to a weighted linear regression on the observed depth-bin averages, with bin-specific weights proportional to $\tau_{s\bar{y}}^{-2}$. It ignores observational uncertainty in the covariates and assumes no residual depth dependence after conditioning on the observed predictors. The same prior specifications are used across all models for the parameters that remain active in each specification. The priors on $\boldsymbol{\beta}$, $\boldsymbol{\mu}$ and $\boldsymbol{\sigma}_x$ are unchanged in $\mathrm{M}_{\mathrm{full}}$ and $\mathrm{M}_{\mathrm{noGP}}$, where the covariate errors-in-variables layer is retained. In $\mathrm{M}_{\mathrm{noEIV}}$ and $\mathrm{M}_{\mathrm{base}}$, the latent covariates and their population parameters are removed because the restriction $x_{sk}^* \equiv \bar{x}_{sk}$ is imposed. Similarly, the priors on σ_u and ℓ are used only in $\mathrm{M}_{\mathrm{full}}$ and $\mathrm{M}_{\mathrm{noEIV}}$, where the Gaussian process depth-dependence layer is active. This ensures that the restricted comparison reflects the inclusion or exclusion of the two modelling components, rather than changes in prior assumptions across specifications.

4 Simulation study

The data-generating process is calibrated on the real $\mathrm{M}_{\mathrm{full}}$ posteriors so that the synthetic regime matches the study system rather than an idealised one. The ground truth $(\boldsymbol{\beta}, \sigma_u, \ell, \boldsymbol{\mu}_x, \boldsymbol{\sigma}_x)$ is set to the posterior medians estimated on the corresponding campaign, and the depth grid d_s and the plug-in measurement-error scales $\tau_{s\bar{y}}, \tau_{sk\bar{x}}$ are taken from the real binned data. The latent covariates are drawn from a multivariate Gaussian with the empirical covariance of the standardised observed covariates, which reproduces the cross-covariate collinearity of the field data ($|r| > 0.99$). Because the fitted model assigns the latent covariates an independent population prior, as in (5), this construction embeds a realistic degree of prior misspecification into every scenario rather than generating data from the estimator's own assumptions.

Four primary scenarios are considered, and each one targets a specific inferential question. **S1** (January-like) uses the Jan-26 truth with a near-zero Gaussian process amplitude and tests coefficient recovery under the most extreme collinearity. **S2** (July-like) uses the Jul-25 truth with a substantial Gaussian process amplitude and tests whether the residual depth structure is identifiable when it is genuinely present. **S3** sets $\sigma_u \equiv 0$ exactly and tests whether the model manufactures spurious depth structure when none exists, which underlies the seasonal contrast interpretation. **S4** is a misspecified scenario in which the residual depth process follows a squared-exponential rather than an Ornstein-Uhlenbeck kernel and the measurement errors are heavy-tailed (Student-t with three degrees of freedom), testing robustness to kernel and error misspecification. Each scenario is replicated 200 times. We additionally implement a supplementary scenario **S1b**, which shares with S1 the structural misspecification already embedded in the primary design, that is, latent covariates generated from a multivariate Gaussian with the full empirical covariance, while the model assigns them an independent population prior. However, it goes further on two fronts. First, the eigenvalues of Σ_x are independently rescaled by a factor drawn uniformly in $[0.8, 1.2]$ in each replicate, so that the simulation truth is no longer a single fixed covariance but a family centred on the empirical one. Second, the conditioning number of Σ_x is itself randomised across replicates. The net effect is that the simulation no longer evaluates the estimator against a single misspecified Σ_x but against a small distribution of plausible misspecifications, which addresses the concern that the calibration result in S1 might be tied to one specific covariance structure. In every replicate the four specifications are fitted with the same sampler settings used for the field data. Performance, shown in Table 3, is summarised by the bias and root-mean-square error of the posterior-median coefficients, the empirical coverage of the nominal 95 per cent credible intervals, the sign-reversal frequency, the recovery of the Gaussian process hyperparameters and the Markov chain Monte Carlo diagnostics. For the two

Table 3: Simulation study: recovery of the two identified tracer coefficients under each scenario. Truth calibrated on the real M_{full} posteriors. Bias and RMSE on the standardised scale; Cov. is empirical coverage of the nominal 95 per cent credible interval; SR is the sign-reversal frequency. Each scenario uses 200 replicates.

Scenario	Coef.	Model	Bias	RMSE	Cov.	SR
S1 (Jan-like, $\sigma_u \approx 0.02$)	Temp.	M_{full}	−0.032	0.076	0.96	0.00
		M_{noEIV}	−0.036	0.096	0.82	0.00
		M_{noGP}	−0.027	0.085	0.93	0.00
		M_{base}	−0.042	0.118	0.12	0.00
S1	O_2	M_{full}	−0.023	0.068	0.95	0.00
		M_{noEIV}	−0.020	0.089	0.81	0.00
		M_{noGP}	−0.021	0.076	0.94	0.00
		M_{base}	−0.015	0.107	0.06	0.00
S2 (Jul-like, $\sigma_u \approx 0.52$)	Temp.	M_{full}	−0.151	0.204	0.95	0.00
		M_{noEIV}	−0.139	0.203	0.82	0.00
		M_{noGP}	−0.158	0.261	0.85	0.04
		M_{base}	0.024	1.369	0.02	0.24
S2	O_2	M_{full}	0.108	0.181	0.98	0.07
		M_{noEIV}	0.109	0.196	0.89	0.08
		M_{noGP}	0.103	0.248	0.91	0.16
		M_{base}	0.268	1.776	0.02	0.45
S3 ($\sigma_u = 0$)	Temp.	M_{full}	−0.056	0.082	0.91	0.00
		M_{noEIV}	−0.041	0.099	0.74	0.00
		M_{noGP}	−0.056	0.083	0.90	0.00
		M_{base}	−0.057	0.106	0.03	0.00
S3	O_2	M_{full}	−0.045	0.074	0.93	0.00
		M_{noEIV}	−0.032	0.102	0.74	0.00
		M_{noGP}	−0.045	0.075	0.93	0.00
		M_{base}	−0.023	0.086	0.03	0.00
S4 (misspecified)	Temp.	M_{full}	−0.126	0.188	0.95	0.00
		M_{noEIV}	−0.126	0.208	0.75	0.00
		M_{noGP}	−0.131	0.261	0.89	0.02
		M_{base}	−0.011	1.999	0.04	0.25
S4	O_2	M_{full}	0.084	0.152	0.98	0.03
		M_{noEIV}	0.093	0.183	0.89	0.09
		M_{noGP}	0.054	0.197	0.94	0.07
		M_{base}	−0.056	2.230	0.03	0.38

identified tracers, M_{full} attains an empirical 95 per cent coverage of 0.95 to 0.98 in every scenario, whereas M_{noEIV} systematically under-covers (0.74 to 0.89) and M_{base}

collapses to 0.02 to 0.12. The collapse of the naive estimator is accompanied by an explosion of its RMSE under the more variable regimes (up to 2.2 on the standardised scale in S2 and S4, against 0.15 to 0.20 for M_{full}) and by sign-reversal frequencies of 0.24 to 0.45 for the tracer coefficients.

The Gaussian process component behaves as required by the seasonal-contrast interpretation (Table 4). When the residual depth structure is genuinely present (S2,

Table 4: Simulation study: recovery of the Gaussian process hyperparameters by M_{full} (posterior-median averaged over replicates).

Scenario	σ_u true	$\hat{\sigma}_u$	ℓ true	$\hat{\ell}$
S1 (Jan-like, $\sigma_u \approx 0.02$)	0.024	0.025	0.24	0.21
S2 (Jul-like, $\sigma_u \approx 0.52$)	0.522	0.441	0.29	0.23
S3 ($\sigma_u = 0$)	0.000	0.015	0.24	0.20
S4 (misspecified)	0.522	0.401	0.29	0.26

true $\sigma_u = 0.52$) it is recovered ($\hat{\sigma}_u = 0.44$); when it is absent by construction (S3, true $\sigma_u = 0$) the posterior median collapses to $\hat{\sigma}_u = 0.015$, which confirms that the model does not manufacture spurious vertical structure and that the near-zero January estimate on the field data is a genuine null rather than an identifiability artefact. In the near-zero scenario S1 the truth (0.024) is recovered almost exactly (0.025).

Robustness to misspecification is established by S4, in which the residual process follows a squared-exponential rather than an Ornstein-Uhlenbeck kernel and the measurement errors are heavy-tailed. The qualitative conclusions are unchanged: M_{full} retains near-nominal tracer coverage (0.95 to 0.98), recovers a substantial Gaussian process amplitude ($\hat{\sigma}_u = 0.40$) and continues to dominate the reduced specifications. The validation is therefore not an artefact of generating data from the estimator's own assumptions. Visualization of these results, in terms of both RMSE and coverage, can be found in Figure 5.

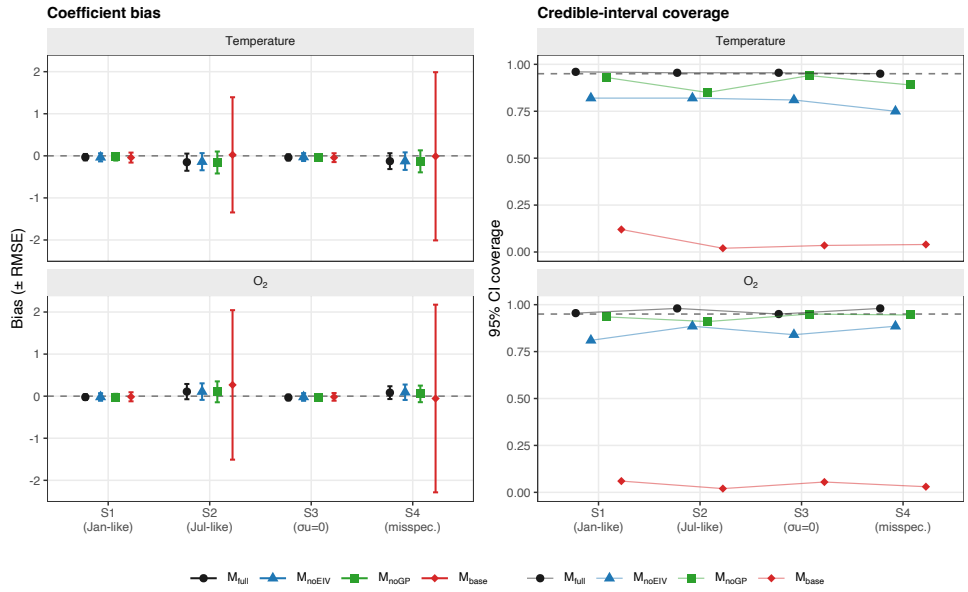


Fig. 5: Simulation study: bias (± RMSE) and empirical 95 per cent credible-interval coverage of the two tracer coefficients, by scenario and specification. The dashed line in the right panel marks nominal coverage.

5 Results

This Section reports the empirical results for the four model specifications introduced in Section 3. In all analyses, the full errors-in-variables plus Gaussian-process specification (M_{full}) is systematically compared against the restricted versions (M_{noEIV}, M_{noGP}, M_{base}) in order to isolate the contribution of the measurement-error correction and of the residual depth-dependence component. Posterior inference is based on four Markov chains with 24000 iterations each and 4000 warmup iterations. We distinguish between two complementary result categories: (i) inferential results, concerning the estimated covariate–salinity associations and the residual depth structure, and (ii) predictive results, concerning cross-month prediction performances. We finally provide a discussion of the main findings.

5.1 Inferential results

Table 5 reports the posterior median and 95 per cent credible interval of the standardised regression coefficients under all four specifications, for both campaigns. Under

Table 5: Posterior median (95 per cent CI) of β_k for the four model specifications. Jan-26 (top) and Jul-25 (bottom). Values on the standardised scale. Coefficients whose 95 per cent credible interval does not include zero in bold.

	M_{full}	M_{noEIV}	M_{noGP}	M_{base}
January 2026				
Temp.	**0.37** [0.18, 0.54]	**0.73** [0.48, 0.95]	**0.40** [0.20, 0.57]	**1.35** [1.26, 1.45]
O_2	**−0.49** [−0.66, −0.31]	−0.20 [−0.44, 0.02]	**−0.46** [−0.62, −0.30]	**0.46** [0.37, 0.55]
pH	0.10 [0.00, 0.31]	**0.11** [0.01, 0.20]	**0.11** [0.02, 0.33]	**0.14** [0.08, 0.19]
Chl-a	0.03 [−0.49, 0.55]	−0.05 [−0.16, 0.06]	0.05 [−0.47, 0.55]	**0.10** [0.02, 0.17]
Phyc. cy.	−0.03 [−0.46, 0.41]	0.03 [−0.06, 0.11]	−0.02 [−0.43, 0.42]	**−0.22** [−0.26, −0.19]
Phyc. ery.	−0.05 [−0.25, 0.15]	0.00 [−0.03, 0.02]	−0.05 [−0.25, 0.15]	0.00 [−0.02, 0.01]
Distance	0.00 [−0.04, 0.04]	0.05 [−0.02, 0.12]	0.00 [−0.04, 0.04]	**0.09** [0.05, 0.13]
July 2025				
Temp.	**0.59** [0.32, 0.85]	−0.14 [−0.73, 0.45]	**0.70** [0.45, 0.91]	**−0.13** [−0.19, −0.07]
O_2	−0.32 [−0.88, 0.24]	**−0.60** [−1.00, −0.20]	**−0.83** [−1.23, −0.40]	**−0.62** [−0.72, −0.51]
pH	−0.15 [−0.86, 0.58]	−0.26 [−0.62, 0.09]	−0.18 [−0.81, 0.45]	**−0.56** [−0.74, −0.38]
Chl-a	−0.31 [−0.71, 0.10]	−0.18 [−0.61, 0.22]	−0.42 [−0.85, 0.04]	0.02 [0.00, 0.04]
Phyc. cy.	0.16 [−0.34, 0.67]	0.12 [−0.10, 0.34]	**0.53** [0.12, 0.94]	0.02 [0.01, 0.03]
Phyc. ery.	−0.22 [−0.66, 0.19]	−0.06 [−0.41, 0.28]	−0.35 [−0.80, 0.06]	**−0.11** [−0.13, −0.09]
Distance	−0.27 [−0.73, 0.21]	**−1.12** [−1.55, −0.69]	−0.27 [−0.66, 0.10]	**−1.44** [−1.57, −1.30]

M_{full} in Jan-26, temperature and dissolved oxygen are the only variables with posterior probability of the expected direction exceeding 0.97. Temperature has a positive coefficient of 0.37 (95 per cent CI: 0.18 to 0.54), and dissolved oxygen a negative coefficient of −0.49 (95 per cent CI: −0.66 to −0.31), both in the direction expected from the known thermohaline properties of the salt wedge: marine water at the bottom of the river mouth is warmer, saltier and richer in dissolved oxygen than the cold, fresh river water above. The pH coefficient is weakly positive (0.10, 95 per cent CI: 0.00 to 0.31; $\Pr(\beta > 0) = 0.97$), consistent with the slightly higher alkalinity of marine water. Chlorophyll, phycocyanin, phycoerythrin and distance from the sea all have credible intervals that span zero, which indicates that after conditioning on temperature and oxygen these variables carry no additional information about salinity at the depth-bin level. In Jul-25, we again find a positive temperature coefficient (0.59, 95 per cent CI:

0.32 to 0.85; $\Pr(\beta > 0) = 1.00$) but uncertain oxygen (-0.32, 95 per cent CI: -0.88 to 0.24; $\Pr(\beta < 0) = 0.87$). This characterization captures the global cold-fresh to warm-saline end-member contrast between the upstream freshwater reference stations and the marine intrusion. Conistently with Jan-26 results, the effect of all other covariates remain uncertain.

The contrast between $\mathrm{M_{full}}$ and $\mathrm{M_{noEIV}}$ is also informative. Without the errors-in-variables correction, the January temperature coefficient inflates from 0.37 to 0.73, a factor of nearly two, while the oxygen coefficient contracts from -0.49 to -0.20 and loses statistical precision ($\Pr(\beta < 0) = 0.96$ under $\mathrm{M_{full}}$ against 0.04 under $\mathrm{M_{noEIV}}$). This is the expected signature of measurement-error bias under multicollinearity (Carroll et al. 2006): the covariate with the largest error-to-signal ratio (dissolved oxygen, coefficient of variation of about 3.1 per cent) is attenuation-biased toward zero, and the correlated variable (temperature) absorbs the missing association. The distance-from-sea variable acquires an apparent positive effect ($\Pr(\beta > 0) = 0.91$) under $\mathrm{M_{noEIV}}$ that vanishes in $\mathrm{M_{full}}$.

The behaviour of $\mathrm{M_{base}}$ in both campaigns illustrates the consequences of ignoring all three modelling challenges at once. In Jan-26, dissolved oxygen has a positive coefficient (0.46) under $\mathrm{M_{base}}$, the opposite sign from every other specification, the opposite of the physical expectation and the opposite of the raw bivariate correlation ($r = -0.99$). In Jul-25, temperature becomes negative under $\mathrm{M_{base}}$. Both are classical collinearity-induced sign reversals: when redundant predictors are entered jointly without regularisation, the least-squares estimator becomes unreliable and the resulting coefficients are uninterpretable. Additional evidence appears in the credible interval of the distance-from-sea coefficient in Jul-25: $\mathrm{M_{base}}$ reports -1.44 with a 95 per cent CI of $[-1.57, -1.30]$, an interval roughly two-and-a-half times narrower than the corresponding interval under $\mathrm{M_{full}}$ ($[-0.73, 0.21]$). This implausibly tight interval is however not a sign that the baseline has identified a strong distance effect. Rather, the

errors-in-variables specifications widen the interval to match the genuine underlying uncertainty, which is precisely the diagnostic value of the comparison.

Table 6 shows the posterior summaries of the Gaussian process hyperparameters under M_{full} for both campaigns. In Jan-26 the posterior of σ_u is concentrated very

Table 6: Posterior median (95 per cent CI) of the Gaussian process hyperparameters σ_u (residual amplitude) and ℓ (length-scale on the $[0, 1]$ standardised depth axis) under M_{full}.

Campaign	σ_u	ℓ
Jan-26	0.024 (0.003, 0.078)	0.24 (0.09, 0.64)
Jul-25	0.52 (0.20, 0.97)	0.29 (0.12, 0.68)

close to zero (median 0.024), which indicates that temperature and dissolved oxygen together account for essentially all depth-resolved variation in salinity. The Gaussian process component is active but effectively dormant: the covariate structure is sufficient to explain the halocline. In July the posterior of σ_u is centred at 0.52 with a 95 per cent credible interval of 0.20 to 0.97, which establishes that a substantial fraction of the depth to salinity relationship is not captured by the linear regression on the observed covariates. The deeper, more graduated summer halocline involves vertical mixing and density gradients that the simple linear combination of near-surface probe averages cannot represent: some of the depth signal is genuinely orthogonal to the measured physico-chemical variables. The length-scale ℓ is similar in both campaigns (median 0.24 in Jan-26 and 0.29 in Jul-25), which corresponds to a physical correlation length of roughly 1.5 to 1.9 m in the depth domain. This is consistent with the thickness of the halocline observed in the profiles.

The qualitative seasonal contrast is also visually apparent in the posterior of the residual depth process u_s itself (see Figure 6). In Jan-26 the posterior median of u_s is essentially zero across the entire depth range and the 95 per cent band is tight around

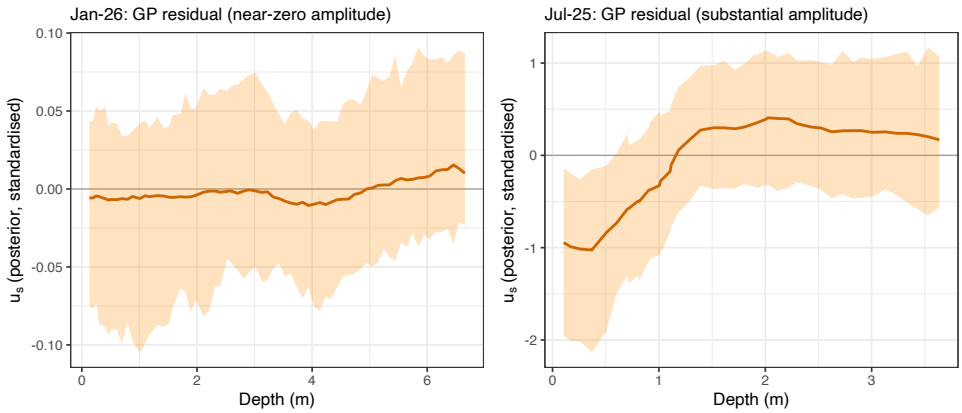


Fig. 6: Posterior of the residual depth process u_s under $\mathrm{M}_{\mathrm{full}}$, on the standardised salinity scale. Solid line: posterior median. Shaded band: 95 per cent credible interval. Left: Jan-26 (near-zero amplitude). Right: Jul-25 (substantial amplitude, with visible depth structure). The pointwise behaviour mirrors the posterior of the amplitude hyperparameter σ_u in Table 6.

zero, which is the pointwise counterpart of the near-zero $\hat{\sigma}_u$ in Table 6. In Jul-25 the posterior median has a clear depth structure with positive and negative excursions of order 0.5 on the standardised scale, and the credible band is correspondingly broad. This is the depth structure that the covariate regression cannot explain and that the Gaussian process absorbs in the summer fit.

5.2 Predictive accuracy

Cross-month predictive performance is assessed by training each model on one campaign and predicting the salinity depth profile of the other campaign. For the Gaussian process based models, the predictive distribution at the new depth locations is computed analytically by conditioning the trained posterior on the new covariate values, integrating over the posterior uncertainty in $\boldsymbol{\beta}$, σ_u and ℓ. For the errors-in-variables layer, the posterior predictive latent covariate x^*_{sk} at each new bin is approximated by the conjugate posterior under the trained hyperparameters $(\hat{\mu}_k, \hat{\sigma}_{x,k})$ and the new

measurement uncertainty $\tau_{sk\bar{x}}$. Performance is measured by root-mean-square prediction error (RMSPE) and continuous ranked probability score (CRPS, Gneiting and Raftery 2007), the latter rewarding both calibration and sharpness of the predictive distribution. The results are shown in Table 7. The baseline model $\mathrm{M_{base}}$ fails considerably in the Jan-26 to Jul-25 direction (RMSPE of 23.1 PSU), roughly 2.3 times worse than $\mathrm{M_{full}}$. Also, the full model consistently outperforms $\mathrm{M_{noEIV}}$ in both directions. In the Jan-26 to Jul-25 direction the difference is 2.2 PSU in RMSPE and 0.9 PSU in CRPS. This is attributable to the more reliable coefficient estimates under errors-in-variables. The attenuated oxygen coefficient and the inflated temperature coefficient under $\mathrm{M_{noEIV}}$ produce a prediction that systematically misses the intermediate salinity values in the transition zone of July. The model without $\mathrm{M_{noGP}}$ achieves slightly better RMSPE than $\mathrm{M_{full}}$ in the Jul-25 to Jan-26 direction (6.7 against 7.2 PSU) but it is worse in the reverse direction (10.1 against 9.9). The modest advantage of $\mathrm{M_{noGP}}$ in the first direction is consistent with the near-zero σ_u found for Jan-26. Indeed, predicting January data from a July-trained model benefits little from the Gaussian process component because the January residuals carry almost no depth autocorrelation beyond what the covariates explain. Conversely, when predicting July from January the Gaussian process is needed to capture the summer depth structure, and

Table 7: Cross-month prediction performance. RMSPE: root-mean-square prediction error (PSU). CRPS: mean continuous ranked probability score (PSU). Lower values are better. Best value in each column in bold.

	Jul-25 to Jan-26		Jan-26 to Jul-25	
Model	RMSPE	CRPS	RMSPE	CRPS
$\mathrm{M_{full}}$	7.2	4.3	**9.9**	**6.3**
$\mathrm{M_{noEIV}}$	8.2	5.7	12.0	7.2
$\mathrm{M_{noGP}}$	**6.7**	**3.9**	10.1	6.5
$\mathrm{M_{base}}$	8.5	6.4	23.1	14.6

M_{full} has lower CRPS even when its RMSPE advantage is modest. Figure 7 visualises the cross-month posterior predictive distribution for the full model in both directions. The observed bin means lie within the 95 per cent predictive band over most of the

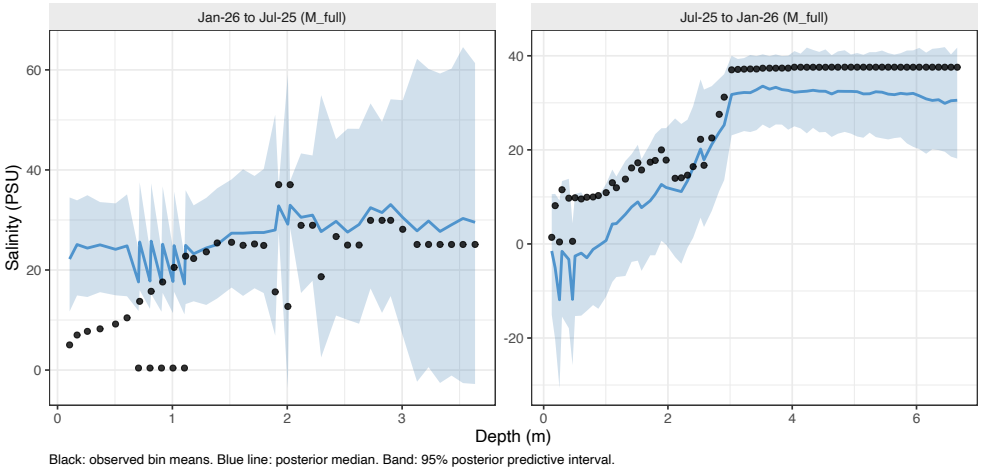


Fig. 7: Cross-month posterior predictive distribution under M_{full}, in both directions. Black points: observed bin-mean salinity. Blue line: posterior median predicted salinity at each depth. Shaded band: 95 per cent posterior predictive interval, propagating uncertainty in $\boldsymbol{\beta}$ and in the Gaussian process hyperparameters from the training campaign. Left: a July-trained model predicting the January profile (RMSPE 7.2 PSU). Right: a January-trained model predicting the July profile (RMSPE 9.9 PSU).

depth range; the larger residuals occur in the transition zone where the halocline is steepest, which is consistent with the role of the Gaussian process residual in absorbing depth structure not captured by the linear covariate combination.

5.3 Discussion

5.3.1 Temperature and dissolved oxygen as salt-wedge tracers

The identification of temperature and dissolved oxygen as the dominant covariates across both seasons and all robust model specifications reflects a well-understood thermohaline mechanism. The salt wedge consists of Mediterranean-origin water that is

warmer and richer in dissolved oxygen than the freshwater column above; along the river axis, the transition from fresh to saline water is therefore a transition between distinct water masses that differ systematically in all measured properties (Savenije 2005; Sorgente et al. 2020). The consistent association of temperature and dissolved oxygen with salinity across the two campaigns should be interpreted as evidence that these variables partly reflect the contrast between water masses within the river mouth, rather than as direct drivers of saltwater intrusion. In a stratified river mouth system, the transition from freshwater to saline water is commonly accompanied by changes in temperature, density, oxygen concentration and other physico-chemical parameters, because the water column integrates the effects of river discharge, marine inflow, vertical mixing, residence time and seasonal atmospheric forcing (Savenije 2005). The errors-in-variables model adds the correct characterisation of the relative strength of each association. Without measurement-error correction, the temperature coefficient is inflated by a factor of two and the oxygen coefficient shrinks toward zero, an attenuation pattern that reverses the apparent dominance between the two tracers. The physical interpretation of the corrected estimates is that temperature and dissolved oxygen carry roughly comparable information about salinity at the bin level, after accounting for the noise in each measurement. This strong association is consistent with the presence of distinct water masses and with the vertical and longitudinal organisation of the salt wedge. However, it should not be interpreted as evidence that temperature or dissolved oxygen directly control salinity, nor that they can be used as universal indicators of saltwater intrusion under all hydrological conditions. Dissolved oxygen may respond to several processes besides water-mass origin, including temperature-dependent solubility, turbulence, stratification, biological production, respiration and organic matter degradation (Lin et al. 2006; Miao et al. 2023). Similarly, temperature may help distinguish water masses during a synoptic survey, but

it is also controlled by seasonal heating, air–water exchange, sampling time and local hydrodynamic conditions.

5.3.2 pH and the secondary tracers

The posterior probability of a positive pH coefficient ($\Pr(\beta > 0) = 0.97$) in Jan-26 under M_{full} is consistent with the known alkalinity difference between oceanic water (pH of about 8.2) and river water (pH of about 7.9 at the freshwater end of the Volturno river mouth), but the effect is small and its credible interval just reaches zero. Its recurrence across M_{full} and M_{noGP} (but not M_{noEIV} or M_{base}) suggests that it is partially masked by measurement imprecision in the raw data and that the errors-in-variables correction is necessary to reveal even this modest signal. Chlorophyll-a and the phytoplankton pigments (phycocyanin, phycoerythrin) are consistently unidentified across all robust specifications. Their bivariate correlations with salinity are in the expected negative direction (marine water is oligotrophic, river water carries more suspended organic matter and pigments), but the effect disappears once temperature and dissolved oxygen are conditioned upon. The pigment signals are therefore collinear proxies for water-mass origin rather than independent contributors to salinity variability.

The weak positive pH coefficient estimated for Jan-26 should be interpreted with particular caution. Although marine waters generally have higher alkalinity and a stronger carbonate buffering capacity than many river waters, pH in estuarine environments is not controlled by salinity alone. It may vary as a function of freshwater and seawater end-member composition, carbonate equilibria, dissolved inorganic carbon, organic matter degradation, biological activity and gas exchange with the atmosphere (Mosley et al. 2010). Therefore, pH should not be described as a direct indicator of saltwater intrusion. A more appropriate interpretation is that, in the January dataset, pH records a modest change along the freshwater–marine mixing gradient, possibly reflecting differences in carbonate chemistry between riverine and marine waters.

The fact that the pH signal is weak and that its credible interval approaches zero is consistent with this interpretation. The parameter may capture part of the chemical transition between freshwater and saline water, but its independent contribution becomes limited once temperature, dissolved oxygen and other correlated variables are included in the model. This is also coherent with the hydrochemical complexity of coastal plains such as the Volturno, where salinity patterns may be influenced not only by present-day river–sea interaction but also by groundwater exchange, stratigraphic heterogeneity, residence time and local geochemical processes.

Chlorophyll-a and the phytoplankton-related pigments should also be treated as secondary tracers. Their relationship with salinity may reflect the distribution of suspended biomass, turbidity, nutrient availability and freshwater inputs, but these variables are not expected to exert a direct control on salinity. Their loss of statistical relevance after conditioning on temperature and dissolved oxygen suggests that, at the scale of the present depth-bin analysis, they mainly record environmental gradients that are already captured by the dominant water-mass structure. Thus, they may help describe ecological or biogeochemical variability along the river mouth, but they should not be interpreted as independent predictors of the salt wedge.

5.3.3 Depth autocorrelation and the seasonal contrast

The near-zero Gaussian process amplitude in January ($\sigma_u = 0.024$) and the substantial amplitude in July ($\sigma_u = 0.52$) point to a qualitative difference in the depth structure of the two salt wedges. In January the halocline is steep and the transition is sharp enough that temperature and dissolved oxygen, which track the same physical gradient, capture essentially all depth-resolved salinity variability. In July the intrusion is deeper, the stratification extends over a larger vertical range and intermediate mixing layers may exist that the observed probe averages do not fully characterise. The Gaussian process absorbs this residual depth signal and prevents it from biasing the regression coefficients. Removing the Gaussian process in July ($\mathrm{M_{noGP}}$) produces

instability in the oxygen and phycocyanin coefficients (both acquiring anomalously large absolute values), which confirms that the Gaussian process is performing a regularising function in the summer analysis. The simulation study (Section 4) confirms that σ_u is genuinely identified: the posterior median tracks the truth in scenarios where the residual depth process is present (S2, S4) and collapses to zero in the scenario where it is absent by construction (S3).

The contrast between the near-zero GP amplitude in January and the substantially higher GP amplitude in July suggests a seasonal difference in the vertical organisation of the salinity field. In January, higher freshwater discharge may have limited the inland extent of the salt wedge and produced a sharper transition between saline and freshwater conditions. Under these circumstances, the measured physico-chemical variables appear to capture most of the depth-resolved salinity variability.

In July, by contrast, reduced freshwater discharge likely favoured a more developed and vertically persistent saline intrusion, as commonly observed in microtidal or low-gradient river mouth reaches during low-flow conditions (Savenije 2005; Gaiolini et al. 2023). The larger GP amplitude in the summer campaign may therefore reflect residual vertical structure that is not fully described by the linear combination of the measured covariates. This residual structure could be associated with vertical density stratification, incomplete mixing, local bathymetric effects, residence-time differences, or the presence of intermediate mixing layers along the halocline.

From a geological and hydrogeological point of view, this result is consistent with the sensitivity of the Volturno coastal plain to salinisation processes. Previous studies have highlighted the role of the Volturno lowland as a vulnerable alluvial-coastal system, where subsidence, stratigraphic heterogeneity and low hydraulic gradients may enhance susceptibility to marine ingression and salinisation (Buffardi et al. 2021; Colombani et al. 2024). Although the present study focuses on surface-water salinity profiles rather than groundwater salinisation, the observed seasonal behaviour of the

salt wedge is coherent with a broader coastal-plain setting in which hydrological forcing and geomorphological conditions jointly influence the distribution of saline waters.

The GP component should therefore be interpreted not only as a statistical correction for residual depth autocorrelation, but also to account for unmeasured physical structure in the water column. In the summer campaign, this structure appears more relevant, possibly because the salt wedge is more vertically developed and less completely described by pointwise covariate measurements.

5.3.4 Predictive performance across two contrasting hydrological regimes

The cross-month predictions establish that the physico-chemical associations are stable enough between the two surveyed regimes to support out-of-sample forecasting. An RMSPE of 9.9 PSU in the Jan-26 to Jul-25 direction, against a salinity range of 0.4 to 37.8 PSU, represents about 26 per cent of the full range. Given that the two campaigns differ substantially in spatial extent, temperature regime and salt-wedge depth, this level of accuracy is consistent with the interpretation that $\boldsymbol{\beta}$ captures physical relationships that hold across the high-discharge winter regime and the low-discharge summer regime sampled here. We stress, however, that the cross-month exercise is a single-fold test based on $N = 2$ synoptic surveys and is therefore informative about transferability between these two specific hydrological regimes, not about seasonal generalisation in the statistical sense, which would require a multi-year programme with several campaigns per season. We discuss this limitation explicitly in Section 6.

The cross-month predictions suggest that the relationships estimated by the model contain information that is partly transferable between seasonal conditions. However, this transferability should be interpreted cautiously. The two campaigns differ in river discharge conditions, spatial coverage, temperature regime and salt-wedge extent. Therefore, the ability to predict one campaign from the other does not imply

that the river mouth behaves identically in winter and summer, but rather that some first-order relationships between salinity and the measured physico-chemical variables remain recognisable across contrasting hydrological settings.

The RMSPE of 9.9 PSU in the Jan-26 to Jul-25 prediction, relative to a salinity range of approximately 0.4 to 37.8 PSU, indicates encouraging predictive skill given the strong seasonal contrast. This level of accuracy supports the interpretation that the regression coefficients $\boldsymbol{\beta}$ may capture first-order physical associations between salinity and the measured covariates that are not purely campaign-specific but are at least partly transferable across seasonal forcing conditions. Nevertheless, the residual uncertainty remains hydrologically meaningful. From an operational perspective, this means that the model can support the interpretation of saltwater intrusion patterns, but it should not replace direct salinity monitoring, especially during low-flow periods when the salt wedge may become more extended and vertically structured.

The better performance of the full model relative to the baseline specification is important from a hydrogeological perspective. It indicates that accounting for measurement uncertainty and depth dependence improves the robustness of the inferred relationships. This is particularly relevant in river mouth monitoring, where vertical gradients may be sharp, sampling depth may affect the measured value, and small measurement errors can lead to misleading interpretations of the position or thickness of the mixing zone.

5.3.5 Implications for environmental risk assessment

Uncertainty quantification is increasingly recognised as a prerequisite for translating statistical model output into actionable environmental risk products (Ciriello et al. 2021). The two simulation findings most relevant to environmental risk assessment are the credible-interval coverage and the sign-reversal frequency of the naive specification. Under near-perfect collinearity, M_{base} achieves an empirical coverage of 0.02 to 0.12 for the two tracer coefficients (Table 3), against a nominal 0.95. A risk product

computed from $\mathrm{M}_{\mathrm{base}}$ would therefore report credible intervals roughly an order of magnitude too tight, and a hazard map based on those intervals would systematically underestimate the probability of high-salinity events at depth. The sign-reversal frequency adds a second, more qualitative failure mode: in the misspecified S4 scenario the naive specification produces a sign-reversed posterior median for dissolved oxygen in 38 per cent of replicates. In risk-product language, this is the rate at which a downstream user, asking the model whether oxygen-rich water is associated with marine intrusion, would be told the wrong answer with high confidence. The full EIV+GP specification reduces this rate to 3 to 7 per cent and restores nominal coverage, at the modest cost of a single posterior fit per campaign with the tuning of Section B.

This calibration result has a direct operational counterpart. The two identified tracers, temperature and dissolved oxygen, are the empirical basis for the low-cost monitoring network introduced in Section 1. A network of dissolved-oxygen sensors deployed along the lower 10 km of the river mouth would, on the corrected coefficients, deliver salinity estimates whose physical scale is roughly -4.3 PSU per mg kg^{-1} (Appendix A). On the naive coefficients in Jan-26, the same network would be configured against a positive oxygen-salinity association, the opposite of physical reality, and would generate alarms in exactly the wrong direction. The errors-in-variables layer is what prevents this failure mode, and the simulation evidence is what allows the recommendation to be made on more than analogical grounds.

The two complementary characterizations of the Jul-25 fit (Section C.2) also map onto two risk products. The full-axis fit informs a basin-scale vulnerability picture that combines the freshwater reach and the marine intrusion within a single posterior, suitable for catchment-level planning of the kind reported by Busico et al. (2021), Gaiolini et al. (2023), Tarallo et al. (2023) and Abd-Elaty and Polemio (2023) for coastal plain settings. The salt-wedge fit informs an operational risk picture restricted to the surveyed lower river reach, suitable for siting sensors and triggering alarms inside

the wedge itself, where the temperature-salinity relationship has the opposite sign and the residual depth process is much weaker. Treating the two as alternative versions of the same quantity, or ranking one as headline against the other as sensitivity, would obscure both and would in a risk-assessment workflow produce inconsistent downstream products: a vulnerability map that under-represents the local thermal stratification of the wedge, and an operational alarm that under-represents the basin-scale gradient. The model proposed in this paper makes both products available from a single estimation pipeline, and the comparison between them is itself a diagnostic of how strongly a given risk question depends on the surveyed spatial extent.

6 Conclusions

This paper addressed the statistical modelling of depth-resolved saltwater intrusion in the Volturno River mouth under bin-level uncertainty and residual vertical dependence. The objective was to identify which physico-chemical variables are most consistently associated with salinity and to provide an uncertainty-aware representation of salinity profiles across contrasting hydrological conditions. The main methodological contribution is a Bayesian errors-in-variables framework with a Gaussian process component capturing residual dependence along the water column. This formulation separates three sources of variation that are often confounded in depth-resolved environmental profiles, that is, the covariate–salinity association, the uncertainty induced by depth-bin aggregation, and the smooth residual structure along depth.

The empirical results indicate that temperature and dissolved oxygen are the most robust covariates associated with salinity. The Gaussian process component revealed a clear seasonal contrast, as residual vertical structure is negligible during the winter high-discharge campaign, whereas it becomes substantial during the summer low-discharge campaign. This suggests that the summer salt wedge is vertically more

structured and less completely described by the measured physico-chemical covariates. The predictive analysis further showed that accounting for both bin-level uncertainty and residual depth dependence improves the probabilistic reconstruction and cross-campaign transferability of salinity profiles relative to simpler specifications.

Several limitations define natural extensions of the present work. First, the model is intentionally depth-centred. It represents salinity as a function of depth and measured covariates, but it does not explicitly reconstruct the full depth–distance geometry of the salt wedge. This is relevant because saltwater intrusion is intrinsically organised along both the river axis and the water column, with the position, thickness and landward extent of the saline wedge varying longitudinally and vertically (Savenije 2005; Gaiolini et al. 2023). The domain-restricted July analysis showed that the estimated temperature–salinity association changes when the analysis is restricted to the lower river reach, suggesting that the spatial extent of the sampled transect can affect the inferred relationships. Second, the two available campaigns provide an informative contrast between winter high-discharge and summer low-discharge conditions, but they do not support a population-level characterisation of seasonal variability. The cross-campaign prediction should therefore be interpreted as a transferability check between two observed hydrological regimes, not as evidence of general seasonal robustness. Repeated multi-year surveys would be required to quantify inter-annual variability, distinguish persistent salt-wedge features from short-term hydrodynamic fluctuations and assess the stability of the estimated covariate effects. Third, the bin-level uncertainty estimates are treated as plug-in quantities. This choice allows uncertainty from the aggregation step to be propagated into the regression, but it does not account for the uncertainty of the uncertainty estimates themselves. Future extensions could model these scales directly within the Bayesian hierarchy, especially when the number of readings contributing to each depth bin is limited. Finally, from a hydrogeological perspective, the model could also be linked with information on coastal-plain stratigraphy,

subsidence and groundwater salinisation. Previous work on the Volturno coastal plain has shown that the area is sensitive to subsidence and salinisation processes, suggesting that surface-water intrusion and groundwater salinity should be interpreted within a coupled coastal-plain framework rather than as completely separate phenomena (Bufardi et al. 2021; Gaiolini et al. 2023; Colombani et al. 2024). In this perspective, the present statistical framework could contribute to a future monitoring and modelling system aimed at supporting risk assessment, water-resource management and the development of a digital twin for the Volturno River coastal plain.

Acknowledgements

The field campaigns were conducted as part of the SAL.WE. project funded by CNR. We thank the staff of CNR ISMAR Naples for instrument operation and data processing, and Autorità di Bacino Distrettuale Appennino Meridionale.

Competing interests

The authors declare no competing interests.

Data and code availability

The R and Stan code, simulation scripts, and all analysis outputs are publicly available on GitHub at https://github.com/agostinognasso/bms_volturno and archived on Zenodo at https://doi.org/10.5281/zenodo.20444545. The raw CTD campaign data are archived on Zenodo at https://doi.org/10.5281/zenodo.20447565.

Author contributions

A.G.: statistical methodology, Stan computational models implementation, design and execution of the simulation study, formal data analysis, and manuscript drafting. **L.R.P.**: field investigation, data collection, data curation, physical interpretation of

the results, manuscript writing and editing. **R.M.**: statistical methodology, Stan computational models implementation, and manuscript revision. **R.D.**: field investigation and data collection. **F.M.**: design of the field campaigns, physical interpretation of the results, manuscript editing, project super-vision and manuscript revision. **G.S.**: project supervision, manuscript revision, and final approval.
All authors read and approved the final version of the manuscript.

References

Abd-Elaty, I., Polemio, M.: Saltwater intrusion management at different coastal aquifers bed slopes considering sea level rise and reduction in fresh groundwater storage. Stochastic Environmental Research and Risk Assessment **37**(6), 2083–2098 (2023) https://doi.org/10.1007/s00477-023-02381-9

Busico, G., Buffardi, C., Ntona, M.M., Vigliotti, M., Colombani, N., Mastrocicco, M., Ruberti, D.: Actual and forecasted vulnerability assessment to seawater intrusion via GALDIT-SUSI in the Volturno River mouth (Italy). Remote Sensing **13**(18), 3632 (2021) https://doi.org/10.3390/rs13183632

Buffardi, C., Barbato, R., Vigliotti, M., Mandolini, A., Ruberti, D.: The Holocene evolution of the Volturno coastal plain (Northern Campania, Southern Italy): implications for the understanding of subsidence patterns. Water **13**(19), 2692 (2021) https://doi.org/10.3390/w13192692

Berry, S.M., Carroll, R.J., Ruppert, D.: Bayesian smoothing and regression splines for measurement error problems. Journal of the American Statistical Association **97**(457), 160–169 (2002) https://doi.org/10.1198/016214502753479301

Banerjee, S., Gelfand, A.E., Finley, A.O., Sang, H.: Gaussian predictive process models for large spatial data sets. Journal of the Royal Statistical Society: Series B **70**(4),

825–848 (2008) https://doi.org/10.1111/j.1467-9868.2008.00663.x

Corrado, G., Amodio, S., Aucelli, P.P.C., Pappone, G., Schiattarella, M.: The subsurface geology and landscape evolution of the Volturno coastal plain, Italy: interplay between tectonics and sea-level changes during the Quaternary. Water **12**(12), 3386 (2020) https://doi.org/10.3390/w12123386

Colombani, N., Alessandrino, L., Gaiolini, M., Gervasio, M.P., Ruberti, D., Mastrocicco, M.: Unravelling the salinity origins in the coastal aquifer/aquitard system of the Volturno River (Italy). Water Research **263**, 122145 (2024) https://doi.org/10.1016/j.watres.2024.122145

Carpenter, B., Gelman, A., Hoffman, M.D., Lee, D., Goodrich, B., Betancourt, M., Brubaker, M., Guo, J., Li, P., Riddell, A.: Stan: a probabilistic programming language. Journal of Statistical Software **76**(1), 1–32 (2017) https://doi.org/10.18637/jss.v076.i01

Ciriello, V., Lee, J., Tartakovsky, D.M.: Advances in uncertainty quantification for water resources applications. Stochastic Environmental Research and Risk Assessment **35**(5), 955–957 (2021) https://doi.org/10.1007/s00477-021-01998-y

Colombo, P., Miller, C., Yang, X., O'Donnell, R., Maranzano, P.: Warped multifidelity gaussian processes for data fusion of skewed environmental data. Journal of the Royal Statistical Society Series C: Applied Statistics **74**(3), 844–865 (2025)

Carroll, R.J., Ruppert, D., Stefanski, L.A., Crainiceanu, C.M.: Measurement Error in Nonlinear Models: A Modern Perspective, 2nd edn. Chapman & Hall/CRC, Boca Raton, FL (2006)

European Parliament and Council: Directive 2000/60/EC of the European Parliament and of the Council of 23 October 2000 establishing a framework for Community

action in the field of water policy. Technical Report L327, Official Journal of the European Communities (2000)

Fofonoff, N.P., Millard, J. Robert C.: Algorithms for computation of fundamental properties of seawater. Technical Report 44, UNESCO Technical Papers in Marine Science (1983)

Gaiolini, M., Colombani, N., Mastrocicco, M., Postacchini, M.: Seawater intrusion assessment along the Volturno River (Italy) via numerical modeling and spectral analysis. Journal of Hydrology **626**, 130289 (2023) https://doi.org/10.1016/j.jhydrol.2023.130289

Gneiting, T., Raftery, A.E.: Strictly proper scoring rules, prediction, and estimation. Journal of the American Statistical Association **102**(477), 359–378 (2007) https://doi.org/10.1198/016214506000001437

Hoffman, M.D., Gelman, A.: The No-U-Turn sampler: adaptively setting path lengths in Hamiltonian Monte Carlo. Journal of Machine Learning Research **15**, 1593–1623 (2014)

Kim, S.-H., Aral, M.M., Eun, Y., Park, J.J., Park, C.: Impact of sensor measurement error on sensor positioning in water quality monitoring networks. Stochastic Environmental Research and Risk Assessment **31**(3), 743–756 (2016) https://doi.org/10.1007/s00477-016-1210-1

Kuusela, M., Stein, M.L.: Locally stationary spatio-temporal interpolation of Argo profiling float data. Proceedings of the Royal Society A **474**(2220), 20180400 (2018) https://doi.org/10.1098/rspa.2018.0400

Lin, J., Xie, L., Pietrafesa, L.J., Shen, J., Mallin, M.A., Durako, M.J.: Dissolved oxygen stratification in two micro-tidal partially-mixed estuaries. Estuarine, Coastal and

Shelf Science **70**(3), 423–437 (2006) https://doi.org/10.1016/j.ecss.2006.06.032

Mosley, L.M., Husheer, S.W., Hunter, K.A.: Spatial variability of pH and inorganic carbon speciation in estuaries using the composition of the river and seawater end members. Environmental Modelling & Software **25**(12), 1658–1663 (2010) https://doi.org/10.1016/j.envsoft.2010.06.014

Muff, S., Riebler, A., Held, L., Rue, H., Saner, P.: Bayesian analysis of measurement error models using integrated nested Laplace approximations. Journal of the Royal Statistical Society: Series C (Applied Statistics) **64**(2), 231–252 (2015) https://doi.org/10.1111/rssc.12069

Miao, Y., Wang, B., Li, D., Sun, X., Xu, Z., Sun, Q., Jiang, Z., Ma, X., Jin, H., Chen, J.: Observational studies of the effects of wind mixing and biological process on the vertical distribution of dissolved oxygen off the Changjiang Estuary. Frontiers in Marine Science **10**, 1081688 (2023) https://doi.org/10.3389/fmars.2023.1081688

Richardson, S., Gilks, W.R.: A Bayesian approach to measurement error problems in epidemiology using conditional independence models. American Journal of Epidemiology **138**(6), 430–442 (1993) https://doi.org/10.1093/oxfordjournals.aje.a116875

Rasmussen, C.E., Williams, C.K.I.: Gaussian Processes for Machine Learning. MIT Press, Cambridge, MA (2006)

Savenije, H.H.G.: Salinity and Tides in Alluvial Estuaries. Elsevier, Amsterdam (2005)

Schlitzer, R.: Ocean Data View. Alfred Wegener Institute. https://odv.awi.de (2022)

Sorgente, R., Di Maio, A., Pessini, F., Ribotti, A., Bonomo, S., Perilli, A., Alberico, I., Lirer, F., Cascella, A., Ferraro, L.: Impact of freshwater inflow from the Volturno

River on coastal circulation. Frontiers in Marine Science **7**, 293 (2020) https://doi.org/10.3389/fmars.2020.00293

Tarallo, D., Alberico, I., Cavuoto, G., Pelosi, N., Punzo, M., Di Fiore, V.: Geophysical assessment of seawater intrusion: the Volturno Coastal Plain case study. Applied Water Science **13**(12), 234 (2023) https://doi.org/10.1007/s13201-023-02033-x

Zheng, Y., Han, F.: Markov Chain Monte Carlo (MCMC) uncertainty analysis for watershed water quality modeling and management. Stochastic Environmental Research and Risk Assessment **30**(1), 293–308 (2016) https://doi.org/10.1007/s00477-015-1091-8

Zhou, X., Yang, T., Shi, P., Yu, Z., Wang, X., Li, Z.: 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 (2017) https://doi.org/10.1007/s00477-017-1399-7

Appendix A Scale factors for back-conversion to physical units

All variables are standardised to zero mean and unit variance on the bin-averaged scale before fitting, which absorbs the intercept by construction and yields regression coefficients on the standardised scale. To convert a standardised coefficient β_k to a physical coefficient $\tilde{\beta}_k$ (change in salinity per unit change in the original covariate x_k), the back-conversion formula is

$$\tilde{\beta}_k = \beta_k \, \frac{\sigma_y}{\sigma_{x_k}},$$

where σ_y and σ_{x_k} are the bin-averaged standard deviations reported in Table A1. The bin-averaged means μ_y, μ_{x_k} are also reported because they are required to reconstruct the intercept-free linear predictor on the physical scale, $\hat{S}(x) = \mu_y + \sum_k \tilde{\beta}_k (x_k - \mu_{x_k})$, and to convert posterior predictive draws back to PSU for salinity. We emphasise that these are the means and standard deviations *across depth bins* (one observation per bin), not across raw probe readings.

Table A1: Bin-averaged mean (μ) and standard deviation (σ) for the response and the seven covariates in each campaign, used to standardise variables before fitting and required to back-transform standardised coefficients to physical units.

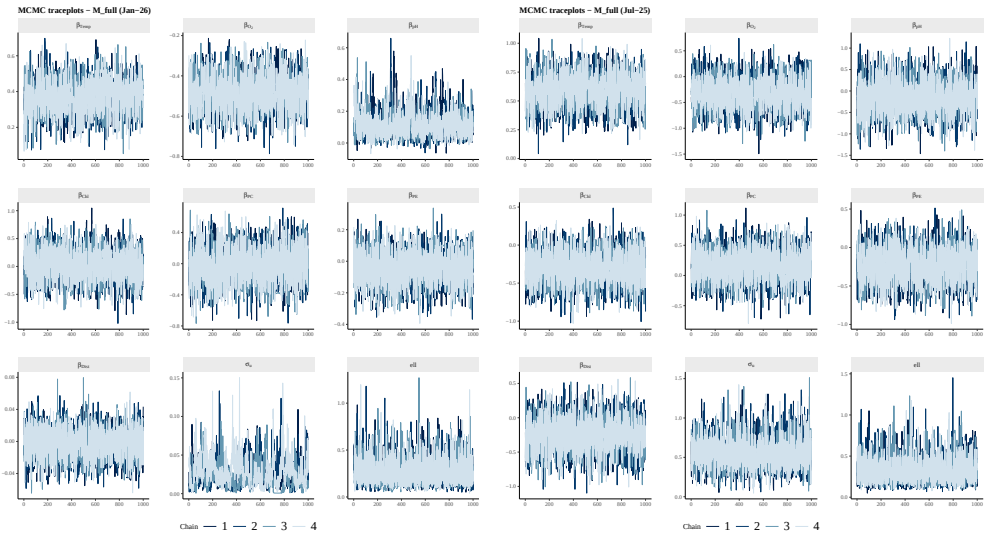
	Jan-26		Jul-25	
Variable	μ	σ	μ	σ
Salinity (PSU)	26.756	12.671	19.443	10.293
Temperature (°C)	13.356	2.091	26.180	1.022
O_2 (mg kg^{-1})	9.568	1.451	6.982	1.873
pH	8.226	0.201	7.776	0.063
Trx-Chl-a	0.745	2.768	3.123	1.640
Phycocyanin	0.496	1.363	2.877	1.594
Phycoerythrin	1.720	2.251	3.875	1.676
Distance (km)	3.375	1.408	12.004	7.126

As an illustration, the standardised temperature coefficient under M_{full} in Jan-26 is $\beta_T = 0.37$, which translates to a physical coefficient of $\tilde{\beta}_T = 0.37 \times (12.671/2.091) \approx$

2.24 PSU per°C; the standardised oxygen coefficient $\beta_{O_2} = -0.49$ translates to $-0.49\times$ $(12.671/1.451) \approx -4.28$ PSU permg kg^{-1}.

Appendix B Sampling behaviour of the proposed model

The sampling behaviour of the proposed model, M_{full}, was assessed using standard MCMC diagnostics and visual inspection of the chains. Posterior inference was based on four chains, with 24000 iterations per chain, including 4000 warmup iterations. The main scalar parameters considered in the inferential analysis are the seven regression coefficients, the residual Gaussian-process amplitude σ_u, and the length-scale parameter ℓ. Figure B1 reports the corresponding traceplots for the Jan-26 and Jul-25 campaigns. Overall, the chains show satisfactory mixing and stationary behaviour for



(a) Jan-26 (b) Jul-25

Fig. B1: Traceplots of the main scalar parameters of M_{full} for the two field campaigns.

the parameters used in the main posterior summaries. The Jul-25 fit shows a mildly

more persistent behaviour for σ_u, reflecting the stronger residual depth component estimated in that campaign, but without evidence of pathological non-convergence. Numerical diagnostics were consistent with the visual inspection: the maximum $\hat{R}$ values were close to one, and the effective sample sizes were adequate for the posterior summaries reported in the main text.

Appendix C Robustness analysis

In what follows, we report three robustness analyses. The first examines the sensitivity to alternative prior specifications. The second evaluates the impact of restricting the Jul-25 dataset to the lower river reach, in order to assess how the estimated associations depend on the sampled spatial extent. The third investigates the sensitivity of the results to misspecification of the within-bin uncertainty estimates.

C.1 Prior sensitivity

We assess the sensitivity of the posterior to the prior specification by re-fitting $\mathrm{M}_{\mathrm{full}}$ under a grid of alternative priors on the two parameters most relevant to the conclusions, namely the regression coefficient prior $\beta_k \sim \mathcal{N}(0, s_\beta^2)$ and the residual amplitude prior $\sigma_u \sim \mathrm{HN}(0, s_{\sigma_u})$. The reference prior used elsewhere in the paper has $s_\beta = 0.5$ and $s_{\sigma_u} = 1$, both weakly informative on the standardised scale. We compare three values of each scale parameter spanning a factor of four ($s_\beta \in \{0.25, 0.5, 1\}$ and $s_{\sigma_u} \in \{0.5, 1, 2\}$). Table C2 reports the posterior median and 95 per cent credible interval of the two identified tracer coefficients (β_{T}, $\beta_{\mathrm{O_2}}$) and of the residual amplitude σ_u under each prior, for Jan-26. The qualitative conclusions are stable across the prior grid. The posterior median of the temperature coefficient β_{T} is essentially flat at 0.37 across the three β prior settings (range 0.37 to 0.37 to two decimal places) and varies between 0.37 and 0.39 across all five settings, a shift smaller than 0.05 on the standardised scale. The oxygen coefficient $\beta_{\mathrm{O_2}}$ varies between -0.46 and -0.50, a span

Table C2: Prior sensitivity for M_{full} on Jan-26. Posterior median (95 per cent CI) of the identified tracer coefficients (β_T, β_{O_2}) and of the residual Gaussian process amplitude σ_u under a grid of weakly informative priors. Reference row in bold.

Prior on β	Prior on σ_u	β_T	β_{O_2}	σ_u
$\mathcal{N}(0, 0.25^2)$	HN(0, 1.0)	0.37 [0.20, 0.53]	−0.48 [−0.63, −0.33]	0.022 [0.003, 0.070]
$\mathbf{\mathcal{N}(0, 0.50^2)}$	**HN(0, 1.0)**	**0.37 [0.18, 0.54]**	**−0.49 [−0.66, −0.32]**	**0.025 [0.004, 0.078]**
$\mathcal{N}(0, 1.00^2)$	HN(0, 1.0)	0.37 [0.17, 0.55]	−0.50 [−0.67, −0.32]	0.025 [0.004, 0.072]
$\mathcal{N}(0, 0.50^2)$	HN(0, 0.5)	0.38 [0.19, 0.55]	−0.48 [−0.66, −0.31]	0.024 [0.003, 0.071]
$\mathcal{N}(0, 0.50^2)$	HN(0, 2.0)	0.39 [0.19, 0.66]	−0.46 [−0.65, −0.18]	0.019 [0.002, 0.069]

of 0.04. Both shifts are well within a posterior standard deviation. The width of the credible interval responds to the prior scale as expected, contracting under $s_\beta = 0.25$ and widening under $s_\beta = 1$ (and similarly for σ_u), but the 95 per cent CIs exclude zero in every prior setting and the posterior probability of the expected physical direction remains above 0.97. The residual Gaussian process amplitude σ_u is insensitive to its own prior scale within the range explored, with posterior medians between 0.019 and 0.025; this confirms that the near-zero Gaussian process estimate in Jan-26 is a feature of the data, not of the prior. The unidentified coefficients (chlorophyll, the pigments, distance) widen their credible intervals under the looser prior, which is the expected behaviour for parameters that the likelihood does not constrain (not shown in Table C2). An analogous analysis on Jul-25 returns the same qualitative pattern.

C.2 Two complementary characterizations of the July fit

The Jul-25 dataset spans 44 km of river axis and combines two physically distinct subdomains. The lower 10 km contain the salt wedge proper, with the thermohaline structure that is the primary object of our methodological apparatus. The three far-upstream stations (FV-CeA at 13 km, FV-Grazzanise at 27 km, FV-Capua at 44 km) lie beyond the marine influence (mean salinity below 0.42 PSU) and act as freshwater

end-members. Fitting M_{full} on the combined dataset versus on the salt-wedge subdomain alone targets two different scientific questions, and we report both rather than ranking one over the other. The Jan-26 dataset has no analogous far-upstream stations and the same distinction does not arise. Table C3 summarises the two posterior fits.

Table C3: Two complementary fits of M_{full} to the Jul-25 dataset. The full-axis fit (top) characterises the global cold-fresh to warm-saline end-member contrast across 44 km of the river axis. The salt-wedge fit (bottom) restricts to the lower 10 km and characterises the local thermohaline structure of the wedge itself. S is the number of depth bins.

Inferential target	S	β_T	β_{O_2}	σ_u
Full-axis (river mouth to 44 km)	43	$0.59\ [0.32, 0.85]$	$-0.32\ [-0.88, 0.24]$	$0.52\ [0.20, 0.97]$
Salt-wedge (lower 10 km only)	30	$-0.65\ [-1.00, -0.28]$	$-0.15\ [-0.46, 0.15]$	$0.08\ [0.00, 0.34]$

The full-axis fit returns a positive β_T because the freshwater reference stations supply the cold-fresh end-member that anchors a global temperature-salinity correlation along the entire 44 km axis: along this axis, salinity and temperature both increase from the upstream freshwater reference toward the marine intrusion, regardless of vertical position. The accompanying $\sigma_u = 0.52$ reflects the residual depth structure that a single linear covariate combination cannot reconcile between the upstream and lower-river regimes merged into one dataset.

The salt-wedge fit returns a negative β_T because, within the lower 10 km in midsummer, marine bottom water is in fact slightly cooler than the warmer freshwater that flows seaward at the surface. This is the opposite of the global cross-axis correlation, and it is the local thermohaline signature of the wedge that a user interested in salt-wedge dynamics within the surveyed lower river reach would be looking for. The collapse of σ_u to 0.08 in this fit is also expected: once the cold-fresh end-member is removed, a single linear covariate combination is sufficient to describe depth structure in the remaining, more homogeneous subdomain, and there is little residual depth signal for the Gaussian process to absorb.

The two targets correspond to two distinct risk products, as developed in Section 5.3.5. A spatial vulnerability assessment that asks how the upstream freshwater reach is connected to the marine intrusion across the full sampled axis is well served by the full-axis fit. An operational risk assessment that asks about the position and thermal structure of the salt wedge within the lower river reach, for example to size a network of oxygen sensors along the lower 10 km, is well served by the salt-wedge fit. Treating the two as alternative versions of the same quantity, or ranking one as headline against the other as sensitivity, would obscure both.

This dual characterization also exposes a structural limitation of the depth-only specification: a one-dimensional $S(z)$ model cannot represent both regimes simultaneously, and the simplest way to do so is to extend the framework to a two-dimensional $S(z, r)$ representation, where the longitudinal coordinate enters the depth structure explicitly. We identify this as the most direct line of methodological extension (Section 6).

C.3 Sensitivity to plug-in measurement-error scales

The model treats the bin-level standard errors $\tau_{s\bar{y}}$ and $\tau_{sk\bar{x}}$ as known, plugged in from the within-bin probe replication (Section 3). When the number of replicates per bin is small, the plug-in scales themselves are imprecise, and the resulting EIV correction may overcorrect or undercorrect. To bound the impact of this misspecification, we refit $\mathrm{M}_{\mathrm{full}}$ on each campaign with the full vector of τ values uniformly multiplied by 0.75 and by 1.25, so that the plug-in measurement-error scales are reduced or inflated by a quarter relative to the reference value. Table C4 reports the posterior summaries.

The identified tracer coefficients move modestly across the three scalings. In Jan-26, β_{T} varies from 0.34 to 0.41 (range 0.07) and $\beta_{\mathrm{O_2}}$ from -0.49 to -0.47 (range 0.02); both 95 per cent credible intervals continue to exclude zero in every setting. In Jul-25, β_{T} varies from 0.51 to 0.61 (range 0.10) and $\beta_{\mathrm{O_2}}$ from -0.44 to -0.25, with the 95 per

Table C4: Sensitivity of M_{full} to a uniform ± 25 per cent rescaling of the plug-in measurement-error scales τ_y and τ_x. Posterior median and 95 per cent credible interval. The reference row in each campaign corresponds to the unscaled τ values used elsewhere in the paper.

Setting	β_T	β_{O_2}	σ_u
Jan-26, $\tau \times 0.75$	0.41 [0.24, 0.59]	−0.47 [−0.64, −0.31]	0.025 [0.004, 0.063]
Jan-26, $\tau \times 1.00$	0.37 [0.19, 0.55]	−0.49 [−0.66, −0.31]	0.024 [0.004, 0.077]
Jan-26, $\tau \times 1.25$	0.34 [0.13, 0.52]	−0.49 [−0.66, −0.29]	0.021 [0.002, 0.075]
Jul-25, $\tau \times 0.75$	0.51 [0.20, 0.79]	−0.44 [−0.98, 0.10]	0.44 [0.15, 0.84]
Jul-25, $\tau \times 1.00$	0.59 [0.32, 0.85]	−0.32 [−0.88, 0.24]	0.52 [0.20, 0.97]
Jul-25, $\tau \times 1.25$	0.61 [0.33, 0.85]	−0.25 [−0.82, 0.33]	0.59 [0.25, 1.06]

cent credible interval of β_{O_2} touching zero only for the inflated $\tau \times 1.25$ rescaling, consistent with the originally weak July oxygen signal already noted in Section 5.1. The displacement of every tracer coefficient remains within one posterior standard deviation of the reference fit. The residual amplitude σ_u inherits the qualitative seasonal contrast under every rescaling, remaining near zero in Jan-26 (posterior medians between 0.021 and 0.025) and substantial in Jul-25 (between 0.44 and 0.59). The conclusions of the paper are therefore robust to a ± 25 per cent misspecification of the plug-in measurement-error scales.