Ocean Modelling 99 (2016) 1–14

Contents lists available at ScienceDirect

Ocean Modelling

journal homepage: www.elsevier.com/locate/ocemod

Evaluation of a long-term hindcast simulation for the Columbia River estuary

Tuomas Kärnä*, António M. Baptista

NSF Science and Technology Center for Coastal Margin Observation & Prediction, Oregon Health & Science University, Portland, OR, USA

ARTICLE INFO

Article history:
Received 14 May 2015
Revised 30 November 2015
Accepted 22 December 2015
Available online 31 December 2015

Keywords:
Estuarine circulation
Model validation
Autonomous underwater vehicle
Mixing processes

ABSTRACT

In order to simulate the biogeochemical function of estuaries across the land-ocean continuum, circulation models must represent a cascade of complex physical processes spanning several spatial and temporal scales. Furthermore, governing physical processes tend to vary under different flow regimes, in response to external forcings. Model validation must therefore cover all relevant flow regimes and span sufficiently long time to represent transient and slowly-varying phenomena. We focus in a multi-year hindcast simulation of the Columbia River estuary – a mesotidal, river-dominated estuary that is also influenced by coastal upwelling in an Eastern Boundary Current system. Model skill is assessed against long-term observational time series, covering the lower estuary (for salinity) as well as most of the tidal river (for water temperature and elevation). In addition, high-resolution profiles of velocity and salinity are used to study salt transport mechanisms at a single station. Results indicate that the model captures the estuarine dynamics of the system, but the skill depends on the flow regime: In general the model performs far better during spring tides (i.e., under partially mixed or time-dependent salt wedge regimes) than under neap tides (i.e., salt wedge and strongly stratified regimes). While the model accurately represents tidal salt transport mechanisms, it tends to underestimate gravitational transport which becomes more important under neap tide conditions. Furthermore, the skill decreases during high river discharge periods, because the model has difficulty capturing the extremely strong stratification characteristic to those periods.

1. Introduction

Numerical modeling of estuarine flows is challenging because of complex bathymetric features, energetic flows and sharp gradients between water masses. In addition, estuarine dynamics tend to vary significantly due to the physical forcings, e.g., tidal variability, seasonal changes in freshwater flow, and synoptic or seasonal weather conditions. Depending on the forcings, estuaries may therefore exhibit multiple flow regimes, that may substantially differ in terms of the dominant physical processes. Calibrating and validating circulation models to all relevant flow regimes is thus of crucial importance.

Circulation models are typically validated for specific, relatively short time periods, whose length is limited by the availability of observational data and computational resources. Such a short-term validation, however, lacks proper representation of slowly-varying phenomena and may miss certain combinations of physical forcings. In this paper we present a skill assessment for a single long-term, multi-year simulation for the Columbia River estuary (Fig. 1). Long-term simulations are necessary to represent slow, history-dependent, seasonal, or interannual aspects of estuarine flows, such as biochemical processes, sediment transport, and response to weather anomalies (e.g., El Niño Southern Oscillation). Assessing the skill of such simulations, however, requires long-term observational record in order to obtain reliable error metrics across the flow regimes. In this work we rely on the rich observational data set of the SATURN network (Science And Technology University Research Network, Baptista et al., 2015) in the Columbia River estuary.

In terms of the flow regimes, we quantify the model skill versus regimes defined by the classification scheme introduced by Geyer and MacCready (2014) (henceforth G–MC classification). The G–MC classification is based on the two main forcings of estuarine systems: tidal currents and river discharge. River discharge affects the

* Corresponding author.
E-mail address: karna@ohsu.edu (T. Kärnä).

http://dx.doi.org/10.1016/j.ocemod.2015.12.007

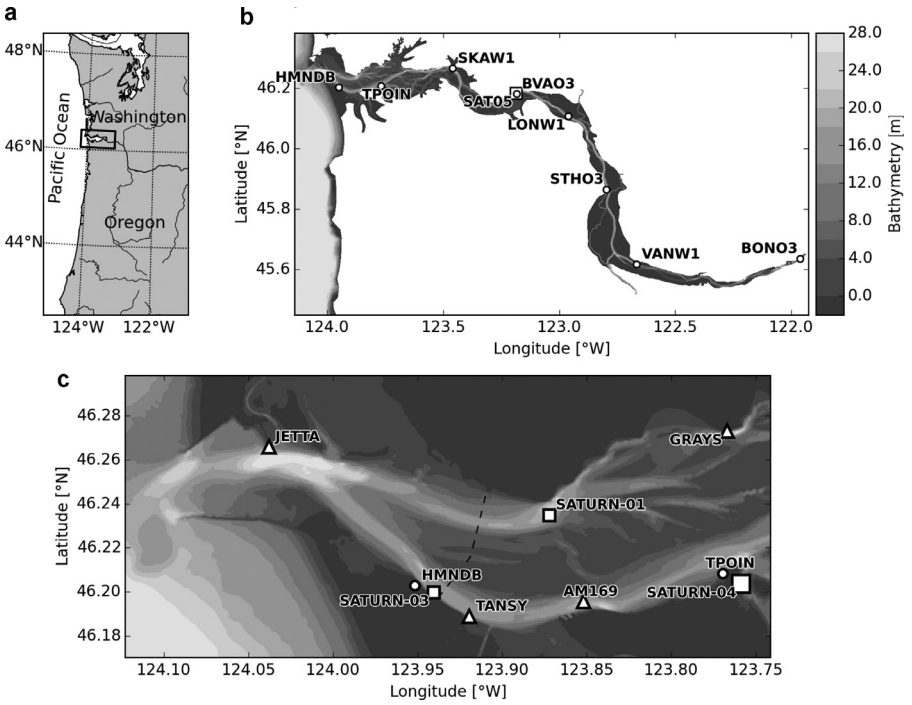

Fig. 1. Geographical location of the Columbia River estuary (a), bathymetry of the tidal river (b), and the lower estuary (c). The multi-disciplinary SATURN endurance stations are marked with squares. Triangles indicate stations that measure only physical quantities. Water level stations are marked with circles. Bathymetry color scale has been cropped at 28 m.

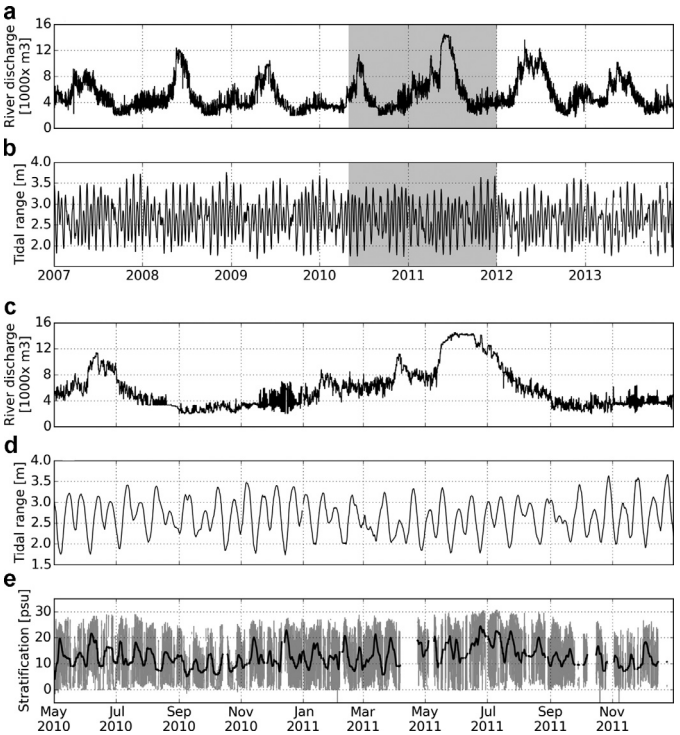

Fig. 2. Physical conditions for the simulation period; (a) river discharge at BONO3; (b) tidal range at TPOIN. Subsequent panels show correlation between river discharge (c), tidal range (d), and observed stratification at SATURN-03 (e) for a shorter time period. Stratification is computed as the salinity difference between the bottom (13.0 m) and surface (2.4 m) measurements. Instantaneous stratification is plotted in gray; the black line is the low-pass filtered signal.

freshwater Froude number Fr_f, that measures the hydraulic criticality of a stratified water column. The magnitude of tidal currents, on the other hand, affects the mixing parameter M, that is a proxy for mixing due to tidal currents and bottom friction. M is scaled to take into account the inhibitory effect of stratification on mixing: $M \approx 1$ indicates that tidal currents are strong enough to mix the entire water column in a half tidal cycle (Geyer and MacCready, 2014). In the context of the Columbia River estuary, the four relevant regimes in the G–MC parameter space are: strongly stratified (low flow, neap tides), partially mixed (low flow, spring tides), salt wedge (high flow, neap tides) and time-dependent salt wedge (high flow, spring tides) regimes.

In this work we analyze model skill for a multi-year hindcast simulation spanning years 2007–2013. River discharge and tidal range are presented in Fig. 2 for the analysis period. The river discharge is highest during the spring freshet period (typically May–June, Fig. 2a), its magnitude varying due to yearly snowmelt conditions and dam operations; for the study period the freshet flows range from 8000 to 15,000 $m^3\ s^{-1}$. During the dry season (July–October) discharge may fall below 2000 $m^3\ s^{-1}$. Tidal range varies from 1.7 m for the smallest neap tides to 3.8 m for the largest spring tides (Fig. 2b). The spring-neap progression is not stationary, however: There's a clear secondary modulation at roughly 190 day time scale, where spring-neap difference varies from the maximum 1.7–3.8 m to much smaller 2.1–3.0 m. This modulation is mostly due to tidal harmonics, namely the superposition of the five dominating tidal constituents (M2, 0.97 m amplitude; K1, 0.40 m; S2, 0.24 m; O1, 0.24 m; N2, 0.18 m). The magnitude of the tides is additionally affected by the river discharge, large discharge tending to decrease tidal range (e.g. during 2011 freshet, Fig. 2d). Both the annual variability of river discharge and the 190 day periodicity of tidal conditions further stress the importance of sufficiently long skill assessment studies.

River discharge and tidal range control stratification and circulation in the estuary (observed stratification is shown in Fig. 2e): stratification is anti-correlated with tidal range, being stronger during neaps; This is especially evident for the weakest neaps (less that 2.0 m tidal range). Stratification is further controlled by the river discharge, higher flows resulting in stronger stratification.

Model results for the analysis period are obtained from our most recent hindcast simulation database, called database 33

Table 1
List of endurance stations and measured variables. Distance is measured from the mouth along the navigation channel. The reported depth stands for depth below datum, except at SATURN-04 where it is depth below instantaneous free surface. Variables marked with filled circles are used in this study; upper estuary stations that only have episodic salinity intrusion (hollow circles) were excluded from the skill analysis.

Station	Latitude	Longitude	Distance	Depth	Variable		
	[° N]	[° W]	[km]	[m]	Elevation	Salinity	Temperature
HMNDB	46.2027	123.9517	14	0.0	•		
TPOIN	46.2081	123.7691	27	0.0	•		
SKAW1	46.2661	123.4589	51	0.0	•		
BVAO3	46.1812	123.1835	81	0.0	•		
LONW1	46.1082	122.9603	101	0.0	•		•
STHO3	45.8636	122.7961	131	0.0	•		
VANW1	45.6167	122.6667	163	0.0	•		
BONO3	45.6336	121.9618	222	0.0	•		
JETTA	46.2660	124.0378	4	6.4		•	•
SATURN-03	46.1997	123.9407	14	2.4		•	•
				8.2		•	•
				13.0		•	•
TANSY	46.1888	123.9195	16	8.4		•	•
SATURN-01	46.2350	123.8719	18	19.5		•	•
AM169	46.1955	123.8516	21	14.3		•	•
SATURN-04	46.2036	123.7586	28	0.3		○	•
GRAYS	46.2732	123.7669	30	6.4		○	•
SATURN-05	46.1845	123.1874	81	2.5			•

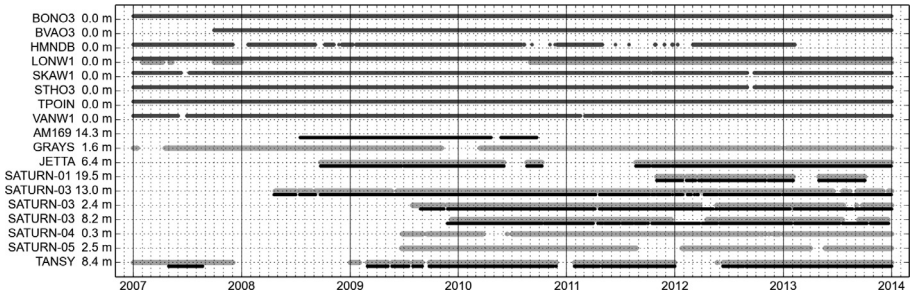

Fig. 3. Availability of SATURN/NOAA station data for selected stations for years 2007–2013. Illustrated variables are: water elevation (dark gray), water temperature (light gray), salinity (black).

(DB33). Model outputs are compared against endurance station data for water elevation, temperature and salinity in the estuary and tidal river. In addition to standard statistical metrics derived for the entire data set, we present skill metrics for each regime separately and show that the model skill significantly varies across regimes. Furthermore, we study salt transport mechanisms in detail for each regime using high-resolution vertical profiles of salinity and velocity at SATURN-01 station. In general, the model performs better during spring tides when tidal salt transport dominates.

Observational data sets and the circulation model are briefly introduced in Section 2. Analysis of model skill based on the long-term time series data is presented in Section 3.1, followed by SATURN-01 profiler comparison in Section 3.2. We conclude with discussion in Section 4.

2. Methods

2.1. Observations

2.1.1. Station observations

Model skill is evaluated against observational data originating from various stations in the estuary and tidal river. The stations are listed in Table 1 and their locations are illustrated in Fig. 1. Water elevation data originates from NOAA tide gauges spanning the entire tidal river (stations BONO3, VANW1, STHO3, LONW1, BVAO3, SKAW1, TPOIN, and HMNDB). Temperature records are from stations LONW1, SATURN-05, GRAYS, SATURN-04, SATURN-01, TANSY, SATURN-03, and JETTA. Salinity data are from lower estuary stations AM169, SATURN-01, TANSY, SATURN-03, and JETTA. The temporal coverage of these data sets is shown in Fig. 3 for the analysis years 2007–2013.

2.1.2. SATURN-01 profiler

In addition to fixed instruments, SATURN-01 station hosts a continuously operated profiler that travels from near-bed elevation (roughly 1 m above the bed) to the surface. Each vertical profile takes roughly 3 min. The profiler carries various instruments, including a composite CT measuring salinity (Falmouth Scientific FSI Digital OEM CT) and temperature (Seabird SBE 3F). The depth of the profiler relative to the free surface is deduced from pressure sensor measurements. The sampling rate of the CT is 1 Hz. For the purposes of the paper the data is binned to 0.33 Hz frequency. Since 2011 SATURN-01 has also had a bottom-mounted, upward-looking Acoustic Doppler Profiler (SonTek ADP 1500–3) for measuring velocity profiles. The ADP bin size is 0.5 m, and sampling period is 5 min. The availability of SATURN-01 profiler data, plotted against the G–MC estuary parameters, is shown in Fig. 4.

2.2. Circulation model

Numerical simulations were carried out with the unstructured grid, finite element model SELFE (Semi-implicit Eulerian–Lagrangian Finite Element model; Zhang and Baptista, 2008) version 4.0.1. SELFE solves the Navier–Stokes equations with hydrostatic and Boussinesq approximations. The horizontal grid consists of triangular elements that are extruded in vertical to form a 3D prismatic mesh. Vertical grid consists of a terrain following S grid (Song and Haidvogel, 1994) near the surface, and equipotential z

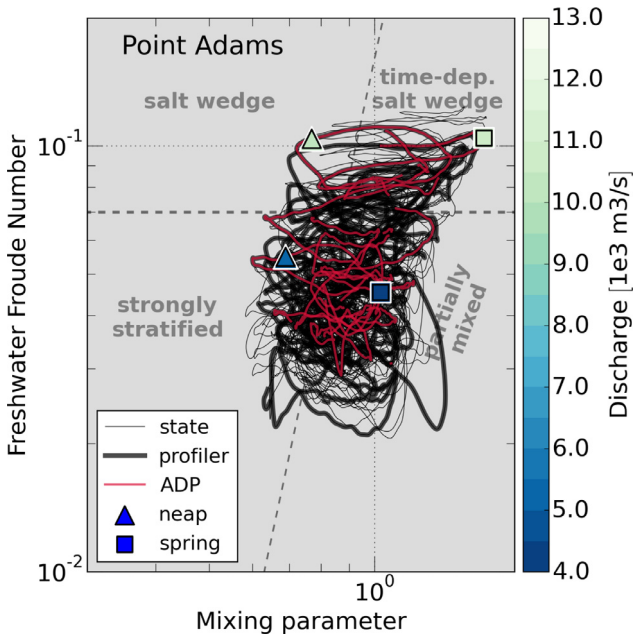

Fig. 4. Availability of SATURN-01 profiler data plotted in the G-MC parameter space for years 2008–2013 (thick black line). The overlaid red line indicates coinciding ADP data. The thin black line indicates the state of the estuary for the entire time span 2008–2013, as estimated from the circulation model (see 2.3). The dashed lines are used to classify the state of the estuary into one of the four flow regimes. Four days chosen to compare different estuary regimes are marked with triangles (neap tides) and squares (spring tides). The date symbols are colored by the residual river discharge observed at BVAO3. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

grid below. The equations are solved in Cartesian space instead of traditional vertical coordinate space.

SELFE uses a semi-implicit formulation to march the equations in time. The evolution of the free surface and vertical diffusion are treated implicitly. Advection of momentum is marched in time with an Eulerian-Lagrangian (ELM) method. This formulation allows relatively long time steps without affecting numerical stability. Tracer transport, on the other hand, uses an explicit, mass conservative upwind method, where shorter time steps are used to retain conservation and monotonicity properties.

Vertical subgrid-scale mixing is described by the Generic Ocean Turbulence Model (GOTM; Burchard et al., 1999), which provides vertical eddy viscosity and diffusivity to the circulation model. In this work a standard k–ε turbulence closure model with Canuto A stability functions (Canuto et al., 2001) is used. No explicit horizontal diffusion is applied.

2.2.1. Model setup

The model domain extends from latitude 39°N to 50°N and roughly 300 km in the offshore direction. Horizontal mesh resolution (triangle edge length) ranges from tens of meters in the estuary and river to 3 km in the ocean. Resolution in the main channels of the lower estuary is roughly 180 m.

The vertical grid consists of 37 S levels, and maximum 17 z levels. The transition from S- to z-levels occurs at 100 m below datum. The S grid is defined such that in shallow areas (below 30 m, i.e. most of the estuary) the coordinates revert to conventional sigma layers.

Bathymetry is a composite of multiple National Geophysical Data Center data sets: ETOPO2v2 (NGDC, 2006), 3 arc second Coastal Relief Model (NGDC, 2011), as well as 1/3 arc second rasters for selected coastal regions. Data compiled by Oregon Department of Geology and Mineral Industries are used in the estuary. Bathymetry in the lower estuary and navigation channel is corrected using recent US Army Corps of Engineers survey data.

Temperature, salinity and water elevations are imposed at the Pacific boundary from global models: Navy Coastal Ocean Model (NCOM, for years 1999–2012; Barron et al., 2006) and Hybrid Coordinate Ocean Model (HYCOM) Global 1/12 Analysis (for year 2013–2014). These models provide only subtidal water elevation, on top of which eight dominant tidal constituents (O1, K1, Q1, P1, K2, N2, M2, S2) are superimposed. The tidal constituents are obtained from a regional inverse model (Myers and Baptista, 2001). Near the open boundary, temperature and salinity values are nudged towards NCOM/HYCOM values using a relaxation time of 2 days.

In the riverine end boundary conditions are imposed at Bonneville Dam, and Willamette, Lewis, and Cowlitz rivers based on observational data. The data originates from USGS and USACE (see Table 2). Gaps in the river data sets were filled with nearest available station data or combining data from other boundaries.

Atmospheric forcing originates from the NOAA/NCEP North American Regional Reanalysis. Wind speed 10 m above surface, air pressure, and heat radiation fluxes are used. Evaporation and precipitation were neglected as they are insignificant compared to the riverine freshwater flux.

The hindcast simulation covers years 1999–2014. We focus our analysis on a representative seven year period (2007–2013). The model was spun up for two weeks, starting in December 17, 1998. Initial conditions for salinity and temperature were obtained from the global models in the shelf, while a linear ramp from global model to constant values (0 psu for salinity and 7.5 °C for temperature) was used in the estuary. Simulations were carried out with 36.0 s time step, storing the model state every 15 min.

Bottom stress is parametrized by imposing either drag coefficient C_d or using the conventional law-of-the-wall condition with bottom roughness length z_0. These parameters were tuned to match salinity and water elevation observations. The law-of-the-wall condition is used in the shelf and estuary, where z_0 is 10^{-4} m, except the main channels downstream of Astoria–Megler Bridge where 10^{-6} m is used. In the river the drag coefficient was applied, with $C_d = 0.006$ between Longview and Willamette confluence, and $C_d = 0.009$ upstream.

The presented model configuration corresponds to the latest Columbia River estuary hindcast simulation, called database 33

Table 2
Data sources for river boundary conditions.

River	Boundary location	Variable	Data source
Columbia	Bonneville Dam	Discharge	Bonneville Dam outflow (USGS 14128870)
		Temp.	Bonneville Dam outflow (USGS 14128870) and Warrendale, OR (USACE SHEF code TWIRGZZAZD)
Willamette	Morrison Bridge	Discharge	Willamette River Below Falls (USGS 14207770)
		Temp.	Willamette River at Portland, OR (USGS 14211720)
		Elevation	Willamette River at Portland, OR (USGS 14211720)
Lewis	Woodland, WA	Discharge	Lewis River at Ariel, WA (USGS 14220500)
		Temp.	Weighted average of Bonneville and Willamette
Cowlitz	Lexington, WA	Discharge	Cowlitz River at Castle Rock, WA (USGS 14243000)
		Temp.	Same as Bonneville

(DB33). For comparison we will also present results for an earlier hindcast simulation, DB22. Compared to DB33, DB22 uses coarser mesh resolution (roughly 300 m in the main channels) and longer time step (90 s; see Kärnä et al., 2015). In addition DB22 simulations were carried out with an older version of SELFE (version 3.0c), including a different generic length-scale turbulence closure implementation.

2.3. G-MC classification parameters

The estuary classification system introduced by Geyer and MacCready (2014) is based on two dimensionless parameters, the freshwater Froude number Fr_f, and mixing number M, given by

$$Fr_f = \frac{U_R}{\sqrt{\beta g S_{oce} H}},$$
$$M^2 = \frac{C_D U_T^2}{\omega N_0 H^2}, \quad (1)$$

with,

$$N_0^2 = \frac{\beta g S_{oce}}{H},$$

where U_R is the river flow velocity, i.e. the river volume flux divided by the cross-section of the estuary, g is the gravitational acceleration, β is the haline contraction coefficient, S_{oce} is the maximal ocean salinity, H is the characteristic depth of the estuary, C_D is bottom drag coefficient, U_T is the amplitude of depth averaged tidal velocity, and $\omega = 2\pi/T_{M2}$ is the tidal frequency.

Fr_f is directly proportional to the river discharge, scaled by the maximal speed of internal waves in the system. M, on the other hand, is proportional to tidal forcing, specifically the bottom friction velocity $u_*^2 = C_D U_T^2$ induced by the tidal currents. N_0 is the buoyancy frequency assuming linear stratification over the water column. M^2 is therefore a ratio of tidal and mixing time scales: $M \approx 1$ implies that tidal mixing is strong enough to mix the entire water column in a half tidal cycle.

These parameters were computed from a cross-section extracted from the circulation model (shown with a dashed line in Fig. 1c). U_R was taken as the sectionally and tidally averaged velocity normal to the cross-section. The bottom friction velocity u_* was computed with the same method as in the circulation model, i.e. the law-of-the-wall boundary condition. M was computed with the maximal u_* along the cross-section and tidal day, because the friction term in M represents the amplitude of tidally induced bed stress. The effective depth H was taken as the mean depth of the cross-section. Tidal averages were computed with a Butterworth low-pass filter (passband $T = 8T_{M2}$, amplification $\pm$ 3 dB; stopband $T = T_{M2}$, attenuation $\pm$ 30 dB) to ensure smooth transition between regimes. In addition the following constant were used: $g = 9.81$ m s^{-2}, $\beta = 7.7 \times 10^{-4}$ psu^{-1}, $S_{oce} = 34$ psu, $T_{M2} = 44714.0$ s.

Following Geyer and MacCready (2014) the state of the estuary was finally classified to one of the four regimes based on the curves: $Fr_f = 7.0 \times 10^{-2}$ and $M = Fr_f^{1/6}\alpha^{1/4}$, where $\alpha = 3.4$ (dashed lines in Fig. 4). The value of α is based on an empirical fit over a wide range of estuaries (Geyer, 2010).

2.4. Error metrics

Model skill is quantified with a number of statistical measures. Let o_i and m_i, $i = 1, \ldots, N$ be the observed and modeled time series, respectively. Denoting the mean of the time series by $\overline{m}$, the bias, Root Mean Square Error (RMSE) and Centered Root Mean Square Error (CRMSE) are defined as

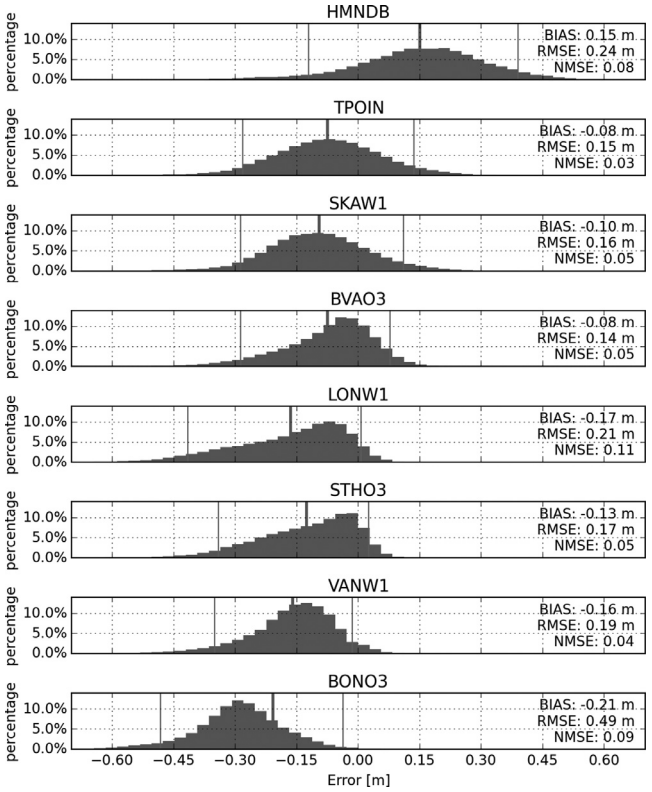

Fig. 5. Error histograms of simulated water levels in selected stations ranging from the lower estuary (top) to Bonneville Dam (bottom) for years 2007–2013. Vertical lines indicate the bias (thick line), 5% and 95% percentiles (thin lines).

$$\mathrm{BIAS} = \overline{m} - \overline{o},$$
$$\mathrm{RMSE}^2 = \frac{1}{N}\sum_{i=1}^{N}(m_i - o_i)^2,$$
$$\mathrm{CRMSE}^2 = \frac{1}{N}\sum_{i=1}^{N}((m_i - \overline{m}) - (o_i - \overline{o}))^2.$$

Standard deviation (σ_m) and correlation coefficient (R) are given by

$$\sigma_m^2 = \frac{1}{N}\sum_{i=1}^{N}(m_i - \overline{m})^2,$$
$$R = \frac{1}{\sigma_o \sigma_m}\frac{1}{N}\sum_{i=1}^{N}(m_i - \overline{m})(o_i - \overline{o}).$$

CRMSE is related to σ_m and R though equation (Taylor, 2001):

$$\mathrm{CRMSE}^2 = \sigma_o^2 + \sigma_m^2 - \sigma_o \sigma_m R. \quad (2)$$

Making use of the Law of Cosines, the latter relation can be visualized in a Taylor diagram (Taylor, 2001), a polar coordinate plot where the radial coordinate is $r = \sigma_m$ and the angle is $\theta = \arccos(R)$. CRMSE then appears as radial distance from the position of a perfect model ($r = \sigma_o, \theta = 0$).

Eq. (2) has the units of the principal variable squared. Scaling (2) by σ_o^2 leads to dimensionless quantities and the normalized Taylor diagram,

$$\mathrm{CRMSE}'^2 = 1 + \sigma_r'^2 - \sigma_m R, \quad (3)$$

with

$$\mathrm{CRMSE}' = \frac{1}{\sigma_o}\mathrm{CRMSE},$$
$$\sigma_m' = \frac{\sigma_m}{\sigma_o}.$$

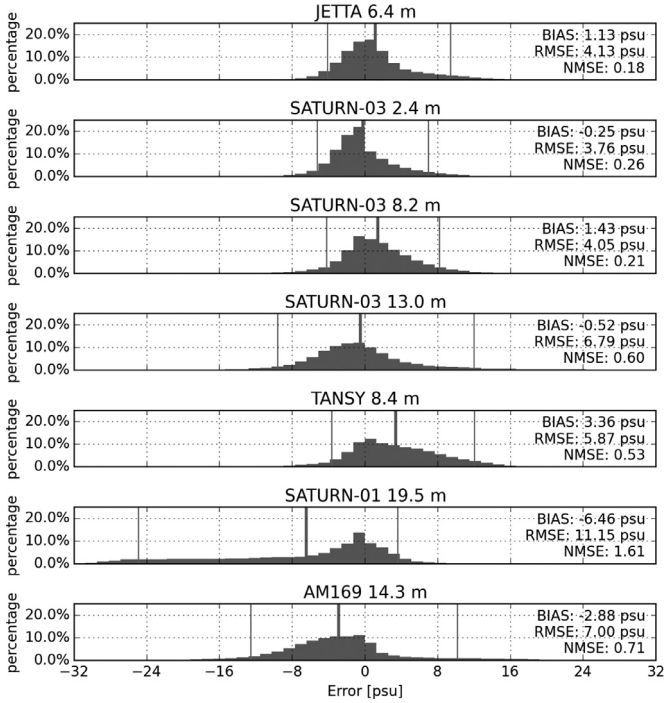

Fig. 6. Error histograms of simulated salinity in selected stations for years 2007–2013. Vertical lines indicate the bias (thick line), 5% and 95% percentiles (thin lines).

In the normalized diagram the perfect model always lies at ($r = 1, \theta = 0$). This allows comparison of data sets with different variances or units in the same diagram.

The Taylor diagram, i.e. Eqs. (2) and (3), is based on centered signals $m_i - \overline{m}$ and thus it does not take into account any model bias. We therefore complement the Taylor diagram with a BIAS-RMSE plot.

To compare skill of different variables, we use Normalized Mean Square Error (NMSE):

$$\mathrm{NMSE} = \frac{1}{\sigma_o^2}\frac{1}{N}\sum_{i=1}^{N}(m_i - o_i)^2. \tag{4}$$

NMSE is 0 for a perfect model, and 1 for a model that is equivalent to the mean of the observations, $m_i = \bar{o}$, which can be seen as the simplest possible predictive model. NMSE greater than one therefore indicates poor predictive skill.

When comparing two model versions, we additionally use the Murphy Score (MS, Murphy, 1988), defined as

$$\mathrm{MS} = 1 - \frac{\frac{1}{N}\sum_{i=1}^{N}(m_i - o_i)^2}{\frac{1}{N}\sum_{i=1}^{N}(r_i - o_i)^2}, \tag{5}$$

where r_i denotes the outputs of a reference model. The Murphy Score is unity for a perfect model, zero for a model equivalent to the reference, and negative for a model worse than the reference.

3. Results

3.1. Station time series

3.1.1. Error histograms

Error histograms for simulated water levels are presented in Fig. 5. In general elevation skill is high: for most stations NMSE is below 0.1 (except at LONW1). Excluding BONO3, the root mean square error is below 0.24 m throughout the domain. The model

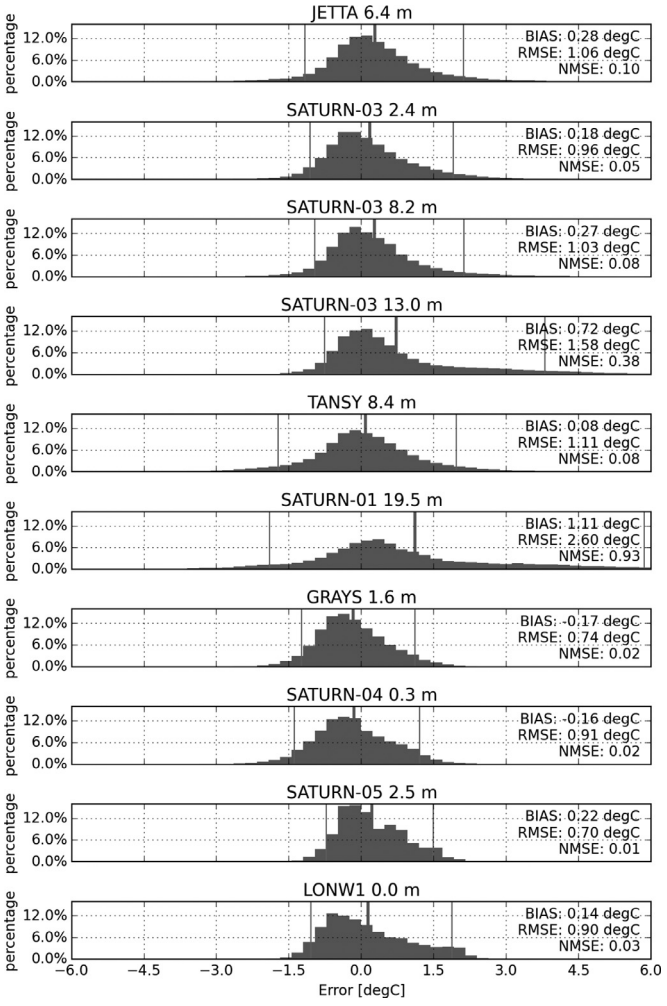

Fig. 7. Error histograms of simulated water temperature in selected stations for years 2007–2013. Vertical lines indicate the bias (thick line), 5% and 95% percentiles (thin lines).

tends to overestimate water elevation near the mouth (HMNDB) but underestimates it in the upstream river. The skill is poorest at BONO3 where the mean level is underestimated by 0.21 m on average.

The error histograms are markedly skewed at certain stations, namely BVAO3, LONW1 and STHO3. The skewness arises from underestimated high water levels (not shown); At these stations the model regularly fails to capture the highest elevation, while low waters are predicted accurately. Underestimation of high waters could arise from various sources. For example, it could be related to the wetting-drying procedure or misrepresentation of small-scale topographic features (e.g., missing dikes). Detailed analysis of this discrepancy however remains a topic for future research.

Regardless of the station the model also exhibits a subtidal bias that varies on monthly time scales (not shown). The bias tends to be stronger during winter months when there are episodes of strong negative bias (up to 0.3 m). This is possibly related to winter storms that may lead to additional water influx through runoff and inflow from smaller tributaries, neither of which is properly represented in the model. The bias also varies slightly during the high flow season and summer months, but the magnitude is considerably smaller (0.1 m).

Error histograms for salinity are presented in Fig. 6. In the deepest part of the channels salinity tends to be underestimated (SATURN-03 13.0 m, SATURN-01 19.5 m and AM169 14.3 m); the

Table 3
Comparison of skill of DB33 and DB22 for years 2007–2012. The Murphy Score (MS) is computed using DB22 as the reference model; positive values indicate improved skill.

Variable	Station	Simulation	RMSE	BIAS	NMSE	MS
Elevation [m]	TPOIN	DB22	0.18	**−0.06**	0.05	0.00
		DB33	**0.15**	−0.08	**0.03**	**0.24**
Elevation [m]	HMNDB	DB22	0.30	0.23	0.13	0.00
		DB33	**0.24**	**0.16**	**0.08**	**0.35**
Elevation [m]	BVAO3	DB22	0.38	0.22	0.36	0.00
		DB33	**0.14**	**−0.07**	**0.05**	**0.86**
Salinity [psu]	JETTA 6.4 m	DB22	4.90	**−0.09**	0.27	0.00
		DB33	**4.18**	1.11	**0.19**	**0.30**
Salinity [psu]	SATURN-01 19.5 m	DB22	15.11	−12.21	2.91	0.00
		DB33	**11.21**	**−6.34**	**1.50**	**0.46**
Salinity [psu]	SATURN-03 13.0 m	DB22	8.64	−3.02	0.92	0.00
		DB33	**7.34**	**−0.75**	**0.67**	**0.10**
Temperature [°C]	JETTA 6.4 m	DB22	1.23	0.36	0.14	0.00
		DB33	**1.00**	**0.22**	**0.10**	**0.31**
Temperature [°C]	SATURN-01 19.5 m	DB22	2.42	**0.78**	**0.88**	0.00
		DB33	**2.33**	0.81	0.93	**0.15**
Temperature [°C]	SATURN-03 13.0 m	DB22	2.30	1.13	0.83	0.00
		DB33	**1.65**	**0.76**	**0.43**	**0.44**

histogram is also considerably wider in these cases. This is because the model underestimates salt retention near the bed during ebb tides (Kärnä et al., 2015), which leads to large errors (close to 30 psu at SATURN-01). Comparing the bottom (13.0 m) and middle (8.3 m) depths at SATURN-03, however, indicates that the poor skill is only confined to the near bed elevations. In general, RMSE is around 4 psu or less for surface and mid-depths; it reaches 11 psu at SATURN-01.

Temperature skill in general is high (Fig. 7). The skill is again poorest near the bed, at SATURN-03 (13.0 m) and SATURN-01 (19.5 m), where NMSE reaches 0.4 and 0.9, respectively. Excluding the bottom stations, however, NMSE is below 0.10, and in the upper estuary and river (GRAYS, SATURN-04, SATURN-05, LONW1) it is 0.03 or less. RMSE is below 2.6 °C in all cases.

The skill for temperature is generally better than that for salinity: At SATURN-01, for example, temperature NMSE is 0.9 whereas it is 1.7 for salinity. The difference is due to the changing characteristics of the end-member water masses. Water temperature varies between roughly 8 °C (summer) and 12 °C (winter) in the shelf, and 5 °C (winter) and 18 °C (summer) in the river. Consequently there are times during spring and fall when both the oceanic and fresh water masses have roughly the same temperature. Under these conditions, the model reproduces temperatures nearly perfectly, regardless of the distribution of oceanic and riverine water masses in the estuary. Salinity of the end-member water masses, on the other hand, remains more or less constant throughout the year, and thus leads to larger errors.

3.1.2. Comparison against previous hindcast simulation

Skill metrics contrasting DB33 and DB22 simulations are presented in Table 3 for selected stations. DB33 yields better skill (positive *MS*) in all cases. The improvement is highest for elevations at BVAO3 ($MS = 0.86$), but it is significant in most stations, *MS* being around 0.3 or greater. The exceptions are SATURN-03 (13.0 m) for salinity and SATURN-01 (19.5 m) for temperature, where the improvement was more modest.

Improved skill is mainly due to higher mesh resolution, better mesh quality (e.g., fitting local bathymetric features) and inreased temporal resolution. During the development of DB33 we have experimented with finer grids and higher temporal resolution, but the skill does not significantly improve to justify higher computational cost and disk usage (Kärnä et al., 2015). It is therefore likely that improving the model skill further requires changes in the discretization, as discussed in Section 4.

Table 4
Frequency of each regime for years 2007-2013, based on the circulation model classifier.

Regime	Days	%
Partially mixed	1303	51.2
Time dependent salt wedge	840	33.1
Strongly stratified	275	10.8
Salt wedge	125	4.9

3.1.3. Skill for each regime

Based on the circulation model outputs, the state of the estuary was classified with the G–MC classification, assigning it to one of the four flow regimes, as detailed in Section 2.3.

The frequency of occurrence of each regime is presented in Table 4. For the majority of the time (84%) the estuary belongs to either the partially mixed or time dependent salt wedge regime. Both of these regimes are associated with spring tides and hence strong tidal mixing. The highly stratified neap tide regimes (strongly stratified and salt wedge), on the other hand, cover only 16% of the time span.

Model skill was computed separately for each regime. The results are presented in a normalized Taylor diagram combined with a BIAS-RMSE plot (Figs. 8–10).

Elevation skill is generally high: most data points fall within the 0.5 CRMSE′ circle (Fig. 8). The difference between regimes is small, except at BONO3 where the skill varies significantly. The stations tend to cluster in the same area in the bias plot, except BONO3 and HMNDB which appear as outliers. BONO3 is an outlier because of its location near the upstream boundary (Bonneville Dam). At HMNDB elevations are constantly overestimated in contrast to other stations. This could be due to local bathymetric features near the station (Hammond boat basin).

The Taylor diagram for salinity is shown in Fig. 9. The station closest to the mouth, JETTA, has one of the highest skill with little variation over the regimes. For the other stations, the skill varies significantly: skill tends to improve as one moves from high flow to low flow conditions and also from neap tides to springs. Skill tends to be worst during the salt wedge regime; best skill is obtained during either partially mixed or time dependent regime. Skill at SATURN-01 is clearly the worst, RMSE exceeding 17 psu during the salt wedge regime.

These results suggest that salinity does not propagate optimally landward in the estuary: Near the mouth (JETTA), where the salt

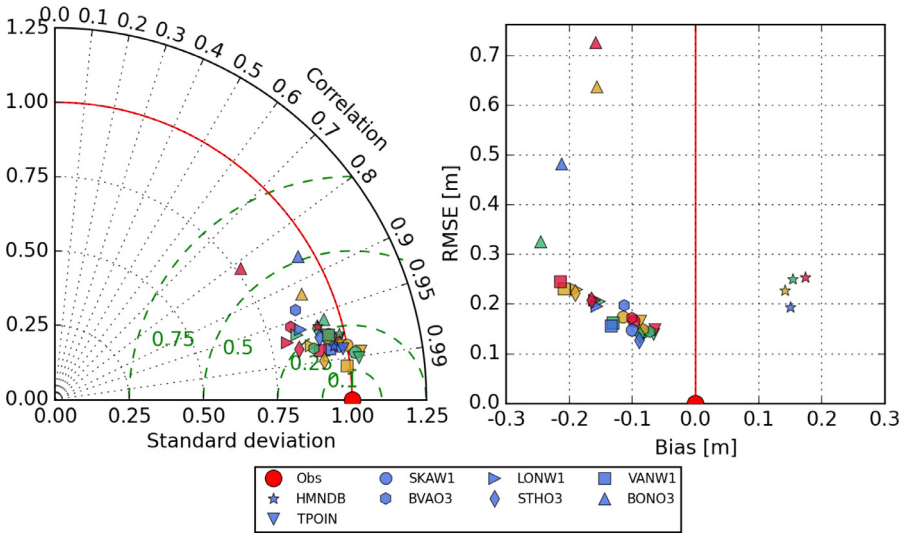

Fig. 8. Normalized Taylor diagram (left) and RMSE versus bias (right) for simulated water elevation for years 2007–2013. The colors indicate different G–MC regimes: blue, strongly stratified; green, partially mixed; red, salt wedge; yellow, time-dependent salt wedge. In the Taylor diagram normalized standard deviation is on the radial axis; correlation coefficient is on the angular axis; green dashed lines indicate CRMSE′. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

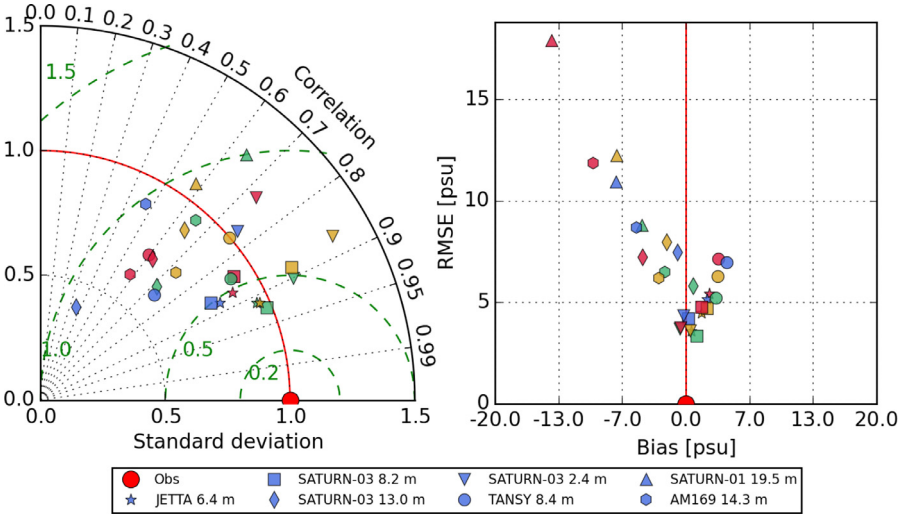

Fig. 9. Normalized Taylor diagram (left) and RMSE versus bias (right) for simulated salinity for years 2007–2013. The colors indicate different G–MC regimes: blue, strongly stratified; green, partially mixed; red, salt wedge; yellow, time-dependent salt wedge. In the Taylor diagram normalized standard deviation is on the radial axis; correlation coefficient is on the angular axis; green dashed lines indicate CRMSE′. Some data points may lie outside the normalized Taylor diagram. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

wedge is being formed in the beginning of each flood tide, salinity is predicted well regardless of the regime. Upper in the estuary, however, the skill is poorer, especially near the bed during neap tides.

Kärnä et al. (2015) showed that the model underestimates gravitational circulation, which affects salinity intrusion and salt retention near the bed. Gravitational circulation is an important mechanism of landward salt transport in the estuary, and becomes dominant during neap tides when tidal mixing (and tidal salt transport) diminishes. In addition, skill tends to be poorer under high flow conditions when the water column is strongly stratified, because the model does not capture sharp density gradients (which also affects gravitational circulation, Kärnä et al., 2015).

As before, the skill for temperature is better than that of salinity (Fig. 10): most of the data points cluster within the 0.5 CRMSE′ circle in the Taylor diagram. The two exceptions are the bottom stations, SATURN-01 and SATURN-03 13.0 m. For these stations, the skill is poorest during the neap tides (salt wedge regime followed by strongly stratified regime). Temperature skill therefore tends to behave similarly as salinity skill although the magnitude of the error is different.

3.2. SATURN-01 profiler

The statistical measures presented above give an overall view of the model skill, but provide little insight into the physical processes associated with each regime. In this section we use short-term, high-resolution data sets from the SATURN-01 profiler to compare the flow characteristics in detail. It should be noted that, as the time series comparison indicates, SATURN-01 is a challenging station to model, most notably due to the local bathymetry in the North Channel (Fig. 1).

The availability of SATURN-01 profiler data is shown in Fig. 4. From this data set we chose 4 tidal days (24.84 h) for each flow

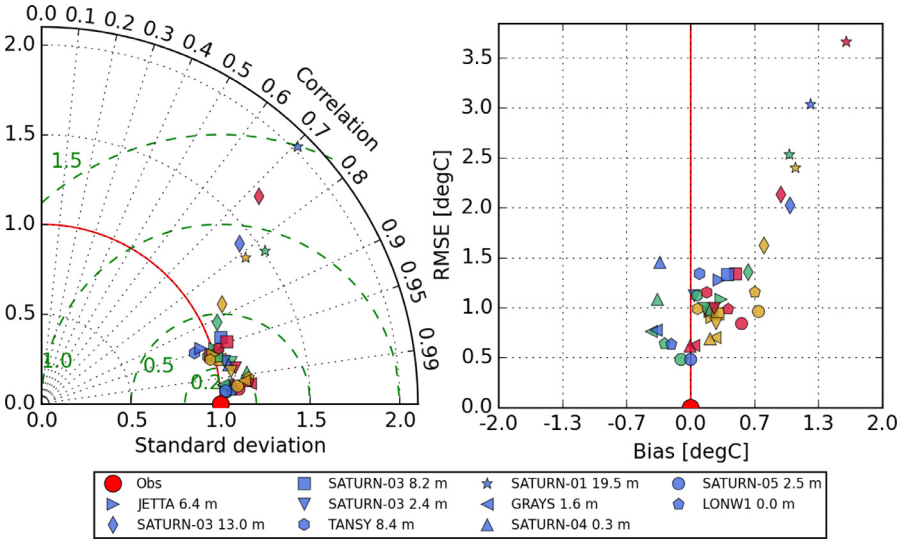

Fig. 10. Normalized Taylor diagram (left) and RMSE versus bias (right) for simulated water temperature for years 2007–2013. The colors indicate different G-MC regimes: blue, strongly stratified; green, partially mixed; red, salt wedge; yellow, time-dependent salt wedge. In the Taylor diagram normalized standard deviation is on the radial axis; Correlation coefficient is on the angular axis; green dashed lines indicate CRMSE′. Some data points may lie outside the normalized Taylor diagram. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

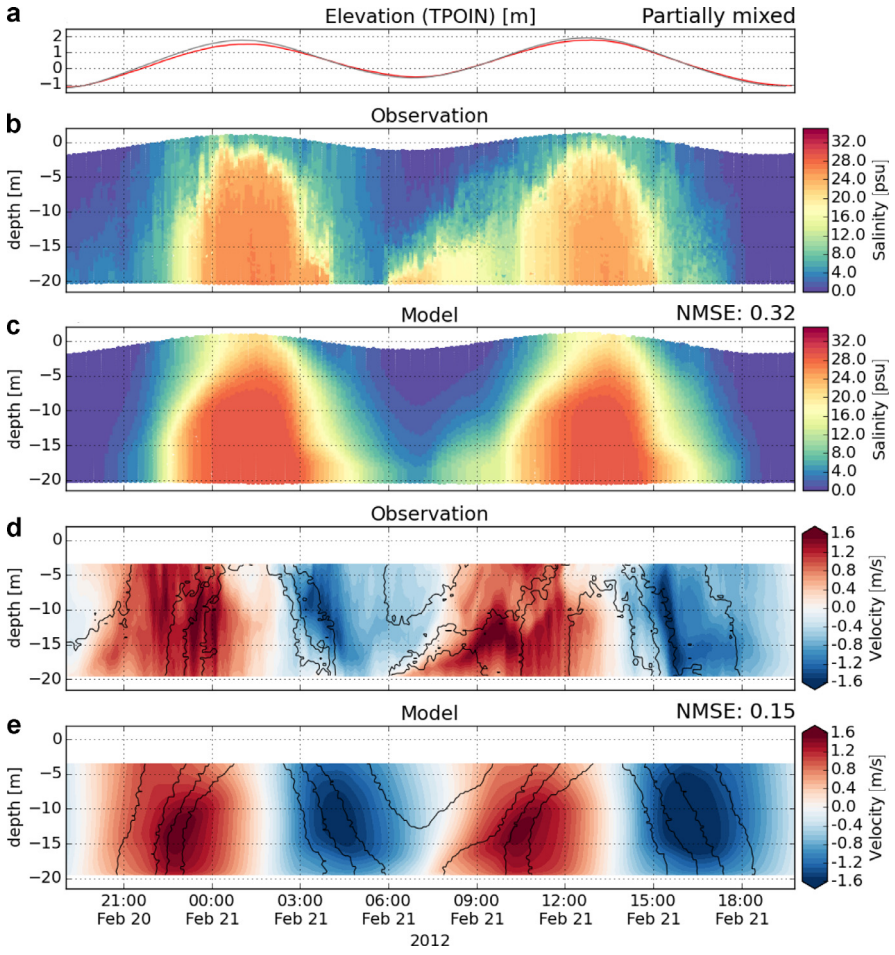

Fig. 11. Comparison of SATURN-01 profiler data for partially mixed conditions (low flow, spring tides). (a) Water elevation at TPOIN (red, observed; black, simulated), (b) observed salinity, (c) modeled salinity, (d) observed along-channel velocity, (e) modeled along-channel velocity. The direction of along-channel velocity is defined as the first principal component of each data set. Positive values indicate landward flow. Thin black lines in panels (d) and (e) indicate 2, 9, 16, and 23 psu isohalines. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

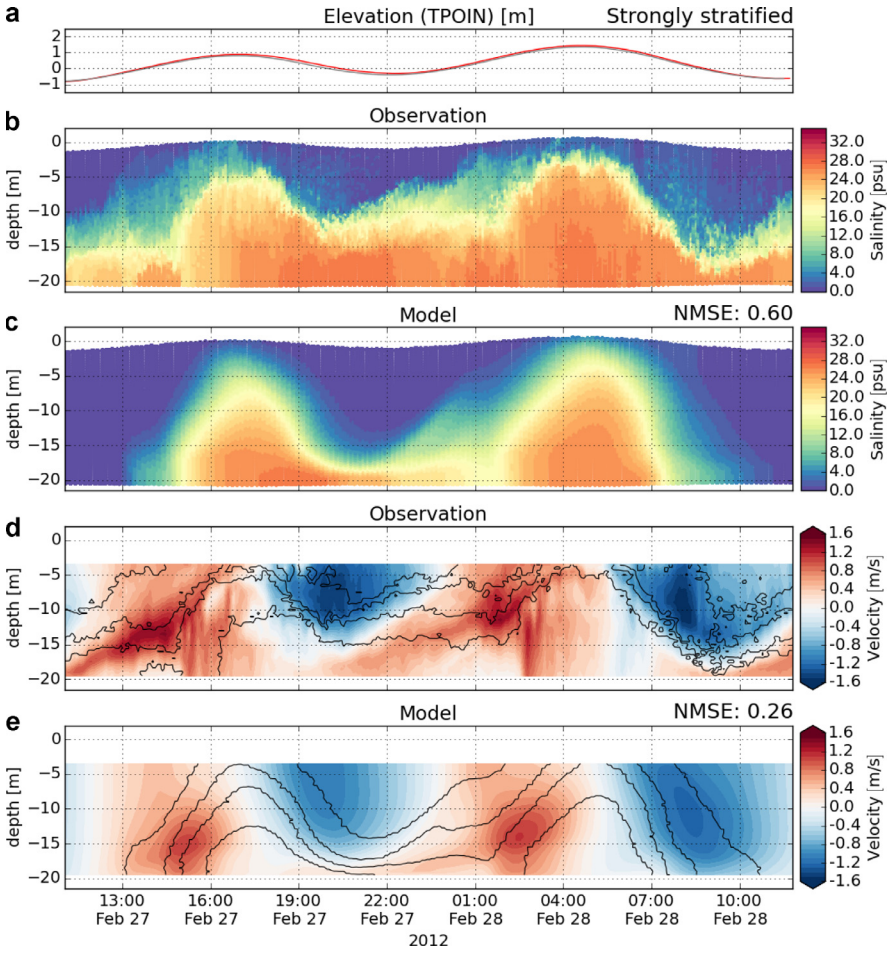

Fig. 12. Comparison of SATURN-01 profiler data for strongly stratified conditions (low flow, neap tides). (a) Water elevation at TPOIN (red, observed; black, simulated), (b) observed salinity, (c) modeled salinity, (d) observed along-channel velocity, (e) modeled along-channel velocity. The direction of along-channel velocity is defined as the first principal component of each data set. Positive values indicate landward flow. Thin black lines in panels (d) and (e) indicate 2, 9, 16, and 23 psu isohalines. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

regime, when both the profiler and ADP data were available. The days were chosen to be as distinct in the G–MC parameter space as possible while having high data quality. Furthermore, in order to facilitate meaningful comparisons the windows were selected so that tidal evolution was similar in all cases. The chosen time periods are (beginning of each window): partially mixed, 2012–02–20 19:00 PST; strongly stratified, 2012–02–27 11:00 PST; salt wedge, 2012–03–30 13:15 PST; time dependent salt wedge, 2012–05–06 07:45 PST.

For brevity we are only focusing on velocity and salinity observations. As estuarine circulation in this system is dominated by advection and diffusion with two end member water masses, temperature fields tend to follow very similar pattern as salinity. Moreover, temperature has only a minor effect on the density structure: In the case of the shown SATURN-01 data sets, temperature accounts to roughly 1% of the density difference between the oceanic and riverine water masses.

Comparison of salinity and along channel velocity fields is presented in Figs. 11–14. Modeled fields were evaluated at the same space-time points as the SATURN-01 profiler data. The along channel velocity was determined as the first principal component of the horizontal velocity field in each data set.

3.2.1. Partially mixed regime

Fig. 11 presents salinity and velocity for the partially mixed regime. Observed salinity intrusion shows two distinct salt pulses associated with the flood tides (around 00:00 and 12:00 in Fig. 11b). In addition, high salinity is observed at the bed during early minor flood (after 06:00). Under these conditions, the presence of salt in the bottom layer during low water depends on the tidal asymmetry: In this particular case the minor ebb (between 2:00 and 7:00) is fairly weak, and it does not flush the salt wedge very far downstream, hence it quickly returns to the station location as the tide turns. The SATURN-01 data set nonetheless contains several examples where the minor ebb is stronger, resulting in two clearly separated salt “pulses”, but none of such examples had simultaneous ADP data.

The velocity field (Fig. 11d) shows that the currents are relatively homogeneous over the vertical. The flood currents encompass the entire water column during both major and minor floods, although strong shear is observed at the halocline, especially at the early phase of the flood (after 06:00).

The model captures the flood-induced salinity intrusion with good accuracy (Fig. 11b, NMSE is 0.32), but tends to overestimate salinity. The high near-bed salinity observed at 06:00 is not captured, possibly related to the overestimated ebb currents (Fig. 11e).

To summarize, under the partially mixed regime the flow is tidally dominated, featuring two salt pulses associated with each flood. Some retention of salt may appear in the lower layer depending on the strength of the minor ebb. Due to weak stratification, velocity tends to be homogeneous over the water column, strongest shear occurring at early floods (when bottom layer

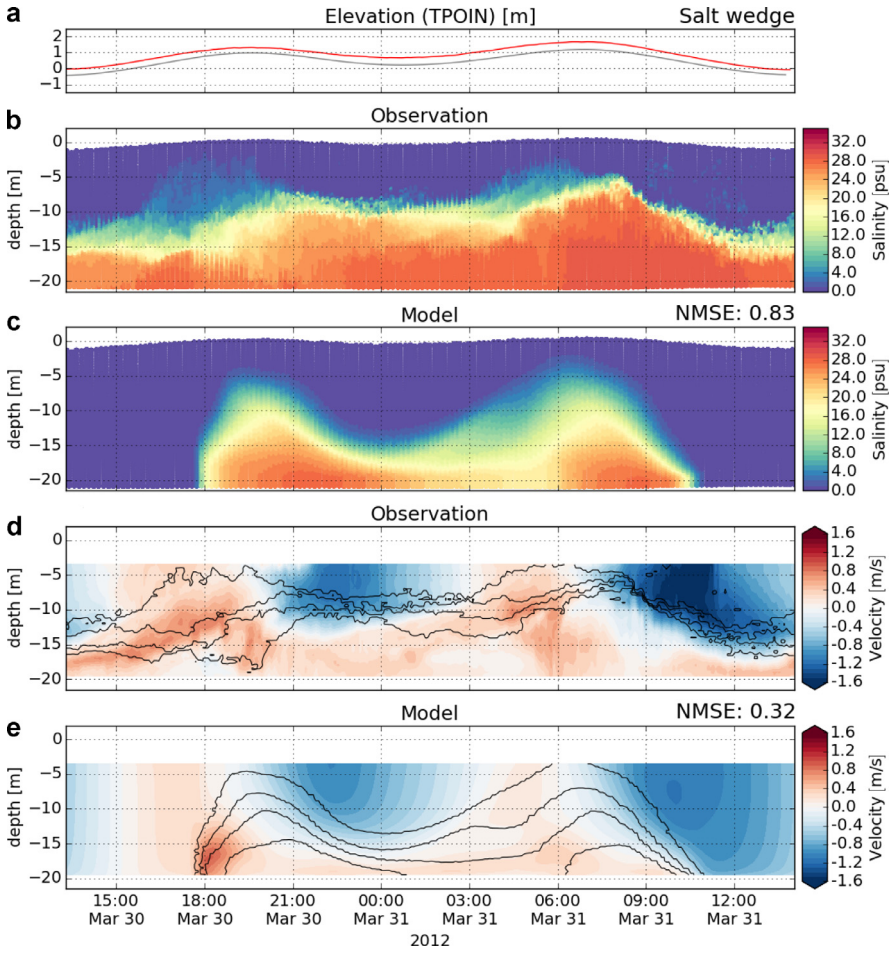

Fig. 13. Comparison of SATURN-01 profiler data for salt wedge conditions (high flow, neap tides). (a) Water elevation at TPOIN (red, observed; black, simulated), (b) observed salinity, (c) modeled salinity, (d) observed along-channel velocity, (e) modeled along-channel velocity. The direction of along-channel velocity is defined as the first principal component of each data set. Positive values indicate landward flow. Thin black lines in panels (d) and (e) indicate 2, 9, 16, and 23 psu isohalines. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

begins to flow landward). Model is able to capture the main salinity intrusion events, but underestimates stratification and tends to overestimate currents.

3.2.2. Strongly stratified regime

As tides become weaker, the observations show very strong salt retention near the bed throughout the tidal cycle (strongly stratified regime, Fig. 12). Tides modulate the position of the halocline; the thickness of the surface layer varies between 2 and 15 m, being smallest at high water. In addition, effect of tidal straining on the salinity field (internal tidal asymmetry, Jay and Musiak, 1994) is visible: Stratification increases during ebbs (e.g. at 07:00) whereas floods (e.g. at 01:00) are less stratified due to increased mixing.

The model performs poorly, NMSE is 0.6 for the salinity field. In general the model underestimates salinity intrusion and does not reproduce the sharp halocline. The retention of salt during minor ebb (around 22:00) is captured, but the major ebb (10:00) is too fresh.

The observed velocity data (Fig. 12d) indicate that the salt wedge is advected back and forth with the tides: For the majority of the tidal cycle, the bottom layer flows landward. The only exception is during the major ebb (around 06:00) when the entire water column flows seaward. During the minor ebb (around 19:00) the top 5–7 m of the salt wedge moves seaward, while the currents near the bed remain landward. The model reproduces a similar velocity field, but underestimates shear.

To summarize, under the strongly stratified regime, salinity is always present at the bed at SATURN-01 and the position of the halocline is controlled by the tides. Stratification and shear are strong. Bottom layer tends to flow landward, except during major ebbs. The model skill is poor: salt intrusion in general is underestimated and the salinity field is overly diffused.

3.2.3. Salt wedge regime

Under high flow and neap tide conditions, both stratification and salt retention are very strong (salt wedge regime, Fig. 13b). Halocline is extremely sharp, especially during major ebbs (at 09:00). Water column becomes less stratified during floods and restratifies during ebbs, again due to internal tidal asymmetry. Compared to the strongly stratified regime, the surface layer is completely fresh and thicker (5–12 m), due to stronger freshwater flow. Tides still modulate the position of the halocline, but the effect is smaller.

The observed currents (Fig. 13d) are weaker compared to other flow regimes, except during major ebb. The bottom layer flows landward for the entire day; ebbing currents are only confined to the surface layer. This indicates that salinity must be transported seaward higher in the water column or through lateral circulation. Indeed, ebbing currents do carry highly saline waters seaward in mid-depths (e.g., between 21:00 and 00:00). The net salt transport at this location is, however, landward indicating that lateral circulation is significant.

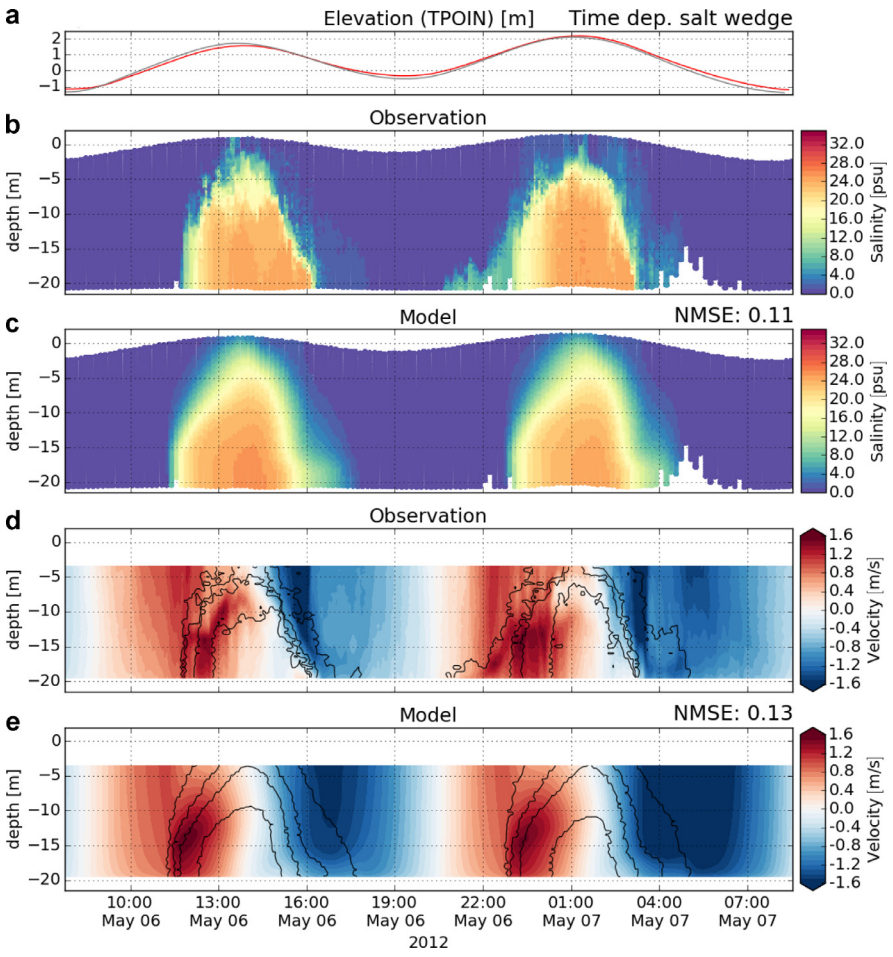

Fig. 14. Comparison of SATURN-01 profiler data for time dependent salt wedge conditions (high flow, spring tides). (a) Water elevation at TPOIN (red, observed; black, simulated), (b) observed salinity, (c) modeled salinity, (d) observed along-channel velocity, (e) modeled along-channel velocity. The direction of along-channel velocity is defined as the first principal component of each data set. Positive values indicate landward flow. Thin black lines in panels (d) and (e) indicate 2, 9, 16, and 23 psu isohalines. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Under these conditions, the modeled salt intrusion is significantly underestimated, leading to poor skill (NMSE 0.8, Fig. 13c). Modeled salinity and velocity fields are overly diffused. Currents and shear are underestimated due to lack of stratification.

To summarize, the salt wedge regime is characterized by a fresh surface layer and extremely strong stratification and shear. Bottom layer is always highly saline and flows landwards for the entire tidal day. This regime is the most difficult to model: both salinity intrusion and stratification are significantly underestimated.

3.2.4. Time dependent salt wedge regime

As tides become stronger under high flow conditions, characteristics of the flow change dramatically (time dependent salt wedge regime, Fig. 14). Tidal excursion of salinity is large, and saline waters reach SATURN-01 only during floods (panel b). The water column is completely fresh during low water. Stratification is strong. Maximal salinity in the bottom layer is somewhat smaller than under other regimes, indicating stronger dilution of salt.

The currents are mainly barotropic (panel d): flood currents are nearly homogeneous in the vertical; some shear is visible once the salt wedge is fully developed (e.g. 13:00). Shear is strong during ebbs.

The model performs clearly better than under the salt wedge regime; it captures the two pulses of salt with good skill (NMSE 0.11, Fig. 14c). Halocline is however smoother than in the observations. The velocity field is likewise well reproduced (NMSE 0.13, Fig. 14e), although the model tends to overestimate ebb currents especially during the main ebb.

To summarize, the time dependent salt wedge regime is tidally dominated (barotropic): salinity intrusion only occurs during floods, and velocity shear remains small except during ebbs. The flow characteristics therefore resemble the partially mixed regime. In contrast to the partially mixed regime, however, stratification is stronger and the water column is entirely fresh at low water. The model skill is very good for both salinity and velocity fields.

3.2.5. Residual salt transport

Residual salt transport profiles were computed for the four SATURN-01 data sets, following the procedure presented in Kärnä et al. (2015). Vertical profiles of along-channel velocity, $u = u(z,t)$, and salinity, $S = S(z,t)$, were estimated from the presented ADP and profiler data, respectively. Residual salt transport is defined as $\langle uS \rangle$, $\langle \cdot \rangle$ denoting the tidal average. The total transport can then be decomposed into mean and tidal (dispersive) components: $\langle uS \rangle = \langle u \rangle \langle S \rangle + \langle u'S' \rangle$, respectively, where $u' = u - \langle u \rangle$ denotes deviation from the mean profile.

Total, mean and tidal salt transport profiles are presented in Fig. 15 for the observations (gray lines) and the model (black).

Concentrating first only on the observed total transport (solid gray line) it is evident that the transport varies significantly under the four regimes. Salinity intrusion in the bottom layer is stronger

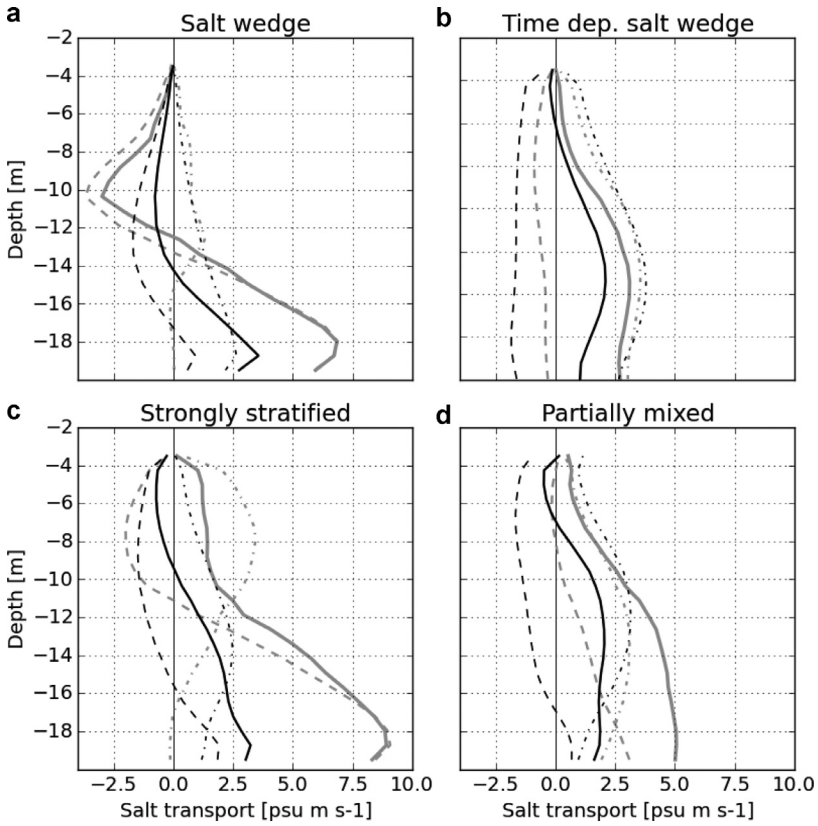

Fig. 15. Vertical profiles of residual salt transport at SATURN-01 for different flow regimes. Solid line, total salt transport $\langle uS \rangle$; dashed line, mean transport $\langle u \rangle \langle S \rangle$; dotted line, tidal transport $\langle u'S' \rangle$. Gray, observations; Black, model.

during neap tides (panels a and c) when gravitational effects are significant. During spring tides (panels b and d) intrusion still occurs in the bottom layer but it is weaker and nearly uniform in the bottom half of the water column. Salt transport is net landward in all cases; significant seaward transport only occurs during the salt wedge regime in the upper half of the water column.

The observed tidal transport (gray dash-dotted line) is always landward (or close to zero), while the mean transport (gray dashed line) may change signs. The influence of tides on these two components is clear: During neap tides, the mean transport is significant, especially in the bottom layer where tidal transport vanishes. For the salt wedge regime tidal transport is generally very small. In contrast, the strongly stratified regime shows significant tidal import of salt in the surface layer.

During spring tides, however, the roles of these two components switch: Mean transport is weak and tidal effects tend to dominate. For the time dependent salt wedge regime (panel b), the mean component is seaward for the entire water column while the tidal component is landward and roughly 3 times as high. During the partially mixed regime, mean transport is landward but still weaker than tidal transport, except at the bottom most 2 m.

This analysis suggests that spring-neap variability is the primary control of the salt transport: during spring tides the flow is mostly barotropic and tidal transport dominates. Under neap tides the flow is more baroclinic and gravitational effects dominate, especially near the bed.

The model behavior is in agreement with the earler skill assessment: The model tends to underestimate gravitational (mean) salinity intrusion in all cases. During spring tides, the transport profiles are in good agreement with the observations (panels b and d), especially for the tidal component. In panel (b) all the profiles are within 2 psu m s^{-1}. The error is larger for the partially mixed regime when the total transport is underestimated by 3 psu m s^{-1} in the bottom layer.

During neap tides, in contrast, the agreement is poor, mean transport being significantly underestimated (by nearly 80% in the bottom layer under the strongly stratified regime), leading into weak total salinity intrusion. It should be noted, however, that during neap tides the lack of salt in the model is so significant that comparing salt transport profiles is less meaningful: The modeled salt wedge is shorter, and hence SATURN-01 is closer to the tip of the salt wedge, resulting in very different dynamics.

4. Discussion and conclusions

We describe a multi-year hindcast simulation of the Columbia River estuary. The simulation captures the main characteristics of the system across the four estuarine regimes, and can be used to support further process studies.

The skill of the simulations was assessed using long-term observations. Water elevations are predicted with good accuracy across all regimes: RMSE for simulated water elevations is below 0.24 m at all stations (excluding BONO3). Temperature RMSE is below 2.6 °C throughout the domain. Salinity is predicted with good accuracy (RMSE below 4.2 psu) for surface and mid-depths, but skill is poorer near the bed (RMSE is 11.0 psu at SATURN-01). Model skill is poorest during the salt wedge regime, i.e. high river flow and neap tides. Outside these conditions, i.e. most of the year (estimated 84% of the time), the model performs satisfactorily.

The observations indicate that the dynamics of the Columbia River estuary vary greatly across the flow regimes. The regimes are controlled by river discharge and tidal range which fluctuate over long time scales (seasonal or annual scales for river discharge, and fortnightly to 190 day period for tidal conditions), which makes

comprehensive model skill assessment difficult. In this context, the endurance stations of the SATURN network and the NOAA tidal gauges are distinctively valuable and were essential for this study. Even with these assets, however, there are important estuarine parameters that are currently impossible to estimate reliably as a long-term time series. Of particular note is the salinity intrusion length, a fundamental parameter of estuarine circulation, whose long-term measurement would require a dense array of stations in the lower estuary.

High-resolution observations from the SATURN-01 profiler show that the roles of tidal and mean salinity transport switch between neap and spring tides: Under spring tides the flow at this location is tidally dominated (barotropic): Salinity intrusion coincides with floods, and most of saline waters are pushed back downstream during ebbs. Under neap tides, however, gravitational circulation becomes more important and the flow is more baroclinic: Salinity intrusion is stronger, and the water column remains stratified over the entire tidal cycle. River discharge provides a secondary control for the salt dynamics at SATURN-01: High discharge decreases salinity intrusion, increases stratification and velocity shear, and inhibits mixing. During high discharge, the surface layer is entirely fresh in contrast to the low flow conditions when some residual salinity appears in the surface layer as well.

Our analysis shows that the model skill strongly depends on the flow regime, and in particular tends to be better for spring rather than neap tides. The SATURN-01 profiler data suggest that poorer skill during neap tides is due to underestimated gravitational circulation. During spring tides, on the other hand, the model is able to capture the dominant barotropic salt transport. Some caution is however needed when generalizing the SATURN-01 profiler results to the entire estuary: due to the local bathymetry, SATURN-01 is an exceptional station where (based on available observational data) the near-bed salinity retention is stronger than anywhere else in the estuary; these results therefore exaggerate the discrepancy between the model and observations.

A salient feature of the circulation model is its diffusivity: the model does not capture the sharp density gradients frequently seen in the observations, which leads into underestimated gravitational circulation. The spurious diffusivity of the model has been discussed in Kärnä et al. (2015); analysis of high-resolution observations suggests that the effective vertical mixing of salinity in the SELFE-based simulation is one order of magnitude higher than what is observed. This spurious diffusion reduces the overall skill of the model, especially during neap tides, and should be addressed. Identifying and addressing the cause of the spurious diffusion is however non-trivial.

Chua and Fringer (2011) showed that using a TVD tracer transport scheme significantly improves the representation of stratification and salinity intrusion in San Francisco Bay simulations compared to an upwind scheme. Our calibration runs for the Columbia River estuary, however, suggest that using TVD tracer advection yields only marginal improvement in stratification and salinity intrusion (not shown). It has also been suggested that poor salinity intrusion may be alleviated by tuning turbulence closure parameters in cases where model's numerical diffusion becomes limiting (David Ralston, pers. comm.). Such results, however, are particular for the studied system and used circulation model. Assessing the accuracy of estuarine models, and identifying the strengths of different numerical schemes therefore remains an open research question, and necessitates systematic comparison of different schemes over a wide range of estuary regimes.

Acknowledgments

The National Science Foundation partially supported this research through cooperative agreement OCE-0424602. The National Oceanic and Atmospheric Administration (NA11NOS0120036 and AB-133F-12-SE-2046), Bonneville Power Administration (00062251) and Corps of Engineers (W9127N-12-2-007 and G13PX01212) provided partial motivation and additional support. This work used the Extreme Science and Engineering Discovery Environment (XSEDE), which is supported by National Science Foundation grant number ACI-1053575. The authors acknowledge the Texas Advanced Computing Center (TACC) at The University of Texas at Austin for providing HPC resources that have contributed to the research results reported within this paper.

References

Baptista, A.M., Seaton, C., Wilkin, M.P., Riseman, S.F., Needoba, J.A., Maier, D., Turner, P.J., Kärnä, T., Lopez, J.E., Herfort, L., Megler, V., Mcneil, C., Crump, B.C., Peterson, T.D., Spitz, Y.H., Simon, H.M., 2015. Infrastructure for collaborative science and societal applications in the Columbia River estuary. Front. Earth Sci. 1–24. doi:10.1007/s11707-015-0540-5.

Barron, C.N., Kara, A.B., Martin, P.J., Rhodes, R.C., Smedstad, L.F., 2006. Formulation, implementation and examination of vertical coordinate choices in the Global Navy Coastal Ocean Model (NCOM). Ocean Model. 11 (3–4), 347–375. doi:10.1016/j.ocemod.2005.01.004.

Burchard, H., Bolding, K., Villarreal, M.R., 1999. GOTM, A General Ocean Turbulence Model. Theory, Implementation and Test Cases. Technical Report EUR 18745. European Commission.

Canuto, V.M., Howard, A., Cheng, Y., Dubovikov, M.S., 2001. Ocean turbulence. part I: one-point closure model – momentum and heat vertical diffusivities. J. Phys. Oceanogr. 31 (6), 1413–1426. doi:10.1175/1520-0485(2001)031.

Chua, V.P., Fringer, O.B., 2011. Sensitivity analysis of three-dimensional salinity simulations in North San Francisco Bay using the unstructured-grid {SUNTANS} model. Ocean Model. 39 (3–4), 332–350. doi:10.1016/j.ocemod.2011.05.007.

Geyer, W., 2010. Estuarine salinity structure and circulation. In: Valle-Levinson, A. (Ed.), Contemporary Issues in Estuarine Physics. Cambridge University Press, pp. 12–26. doi:10.1017/CBO9780511676567.Chapter 2

Geyer, W.R., MacCready, P., 2014. The estuarine circulation. Ann. Rev. Fluid Mech. 46, 175–197. doi:10.1146/annurev-fluid-010313-141302.

Jay, D.A., Musiak, J.D., 1994. Particle trapping in estuarine tidal flows. J. Geophys. Res.: Oceans 99 (C10), 20445–20461. doi:10.1029/94JC00971.

Kärnä, T., Baptista, A.M., Lopez, J.E., Turner, P.J., McNeil, C., Sanford, T.B., 2015. Numerical modeling of circulation in high-energy estuaries: A Columbia River estuary benchmark. Ocean Model. 88 (0), 54–71. doi:10.1016/j.ocemod.2015.01.001.

Murphy, A.H., 1988. Skill Scores Based on the Mean Square Error and Their Relationships to the Correlation Coefficient. Mon. Weather Rev. 116 (12), 2417–2424.

Myers, E.P., Baptista, A.M., 2001. Inversion for tides in the Eastern North Pacific Ocean. Adv. Water Res. 24 (5), 505–519. doi:10.1016/S0309-1708(00)00041-5.

NGDC, 2006. 2-minute Gridded Global Relief Data (ETOPO2v2).

NGDC, 2011. U.S. Coastal Relief Model.

Song, Y., Haidvogel, D., 1994. A semi-implicit ocean circulation model using a generalized topography-following coordinate system. J. Comput. Phys. 115 (1), 228–244. doi:10.1006/jcph.1994.1189.

Taylor, K.E., 2001. Summarizing multiple aspects of model performance in a single diagram. J. Geophys. Res.: Atmos. 106 (D7), 7183–7192. doi:10.1029/2000JD900719.

Zhang, Y., Baptista, A.M., 2008. SELFE: a semi-implicit Eulerian-Lagrangian finite-element model for cross-scale ocean circulation. Ocean Model. 21 (3–4), 71–96. doi:10.1016/j.ocemod.2007.11.005.