JOURNAL OF GEOPHYSICAL RESEARCH, VOL. 114, C00B04, doi:10.1029/2008JC004929, 2009

Evaluation of a coastal ocean circulation model for the Columbia River plume in summer 2004

Yonggang Liu,[1,2] Parker MacCready,[1] Barbara M. Hickey,[1] Edward P. Dever,[3] P. Michael Kosro,[3] and Neil S. Banas[1]

Received 22 May 2008; revised 15 September 2008; accepted 8 December 2008; published 10 March 2009.

[1] Realistic hindcast of the Columbia River estuarine-plume-shelf circulation in summer 2004 using the Regional Ocean Modeling System nested within the Navy Coastal Ocean Model (NCOM) is quantitatively evaluated with an extensive set of observations. The model has about equal skill at tidal and subtidal properties. Tidal circulation and water properties are best simulated in the estuary, which is strongly forced and damped, but worst on the shelf. Subtidal currents are again best in the estuary. However, subtidal temperature and salinity are best simulated in the surface waters on the shelf, even inside the river plume. A comprehensive skill assessment method is proposed to evaluate the cross-scale modeling system with a focus on the plume. The model domain is divided into five dynamical regions: estuary, near- and far-field plume, near surface and deep layers. A skill score is obtained for each region by averaging the skills of different physical variables, and an overall skill is obtained by averaging the skills across the five regions. This weighting metric results in more skill weight per unit volume in the near surface layer where the plume is trapped and in the estuary. It is also demonstrated, through model/data comparison and skill assessment, that by nesting within NCOM, some important remote forcing, e.g., coastal trapped waves, are added to our model; on the other hand, some biases are also received. With a finer grid and more realistic forcing, our regional model improves skill over a larger-scale model in modeling the shelf-plume circulation.

Citation: Liu, Y., P. MacCready, B. M. Hickey, E. P. Dever, P. M. Kosro, and N. S. Banas (2009), Evaluation of a coastal ocean circulation model for the Columbia River plume in summer 2004, *J. Geophys. Res.*, *114*, C00B04, doi:10.1029/2008JC004929.

1. Introduction

[2] As the largest river on the U.S. west coast, the Columbia River (CR) has a significant influence on the hydrography, circulation and ecology along the northwest Pacific coast [e.g., *Hickey and Banas*, 2003]. Previous observational studies of the CR plume described the effects of wind and ambient flows on the plume, and demonstrated that the plume frequently is bidirectional [e.g., *Hickey et al.*, 1998, 2005]. However, the only comprehensive data set was limited to the winter season [*Hickey et al.*, 1998]. A recent process-oriented modeling study with idealized bathymetry and forcing used ECOM3d [e.g., *Garcia-Berdeal et al.*, 2002] to examine the Columbia plume response to mean and fluctuating wind conditions, ambient flow and a variety of outflow conditions. However, this model did not include tides, which may be responsible for much of the plume mixing in the nearshore region. Nor did this model have a realistic estuary. With the advent of larger-domain models that can be used to provide conditions on open boundaries, model studies using realistic bathymetry and forcing can now be used to extend our knowledge of plume dynamics on a regional basis.

[3] Three-dimensional circulation models on unstructured grids, ELCIRC [*Zhang et al.*, 2004] and SELFE [*Zhang and Baptista*, 2008], have been developed for the CR estuary-plume-shelf system with realistic forcing; the early hindcast was validated using only two variables (water level and salinity) mainly in the estuary [*Baptista et al.*, 2005]; the recent model results were validated using more significant amount of data, especially vessel-based flow-through data, but only on the shelf (Y. J. Zhang et al., Daily forecasts of Columbia River plume circulation: A tale of spring/summer cruises, submitted to *Journal of Geophysical Research*, 2009). More recently, the Regional Ocean Modeling System (ROMS) was used in the RISE (River Influence on Shelf Ecosystem, http://www.ocean.washington.edu/rise/) project to simulate the CR estuary-plume-shelf system with realistic atmospheric, tidal and river forcing, and the model results were validated quantitatively with time series at three moorings (temperature (T), salinity (S), and velocity) and qualitatively against hydrographic observations on the shelf [*MacCready et al.*, 2009]. Based on these model results, the energy budget of the wind- and tide-induced mixing in the

[1]School of Oceanography, University of Washington, Seattle, Washington, USA.
[2]Now at College of Marine Science, University of South Florida, St. Petersburg, Florida, USA.
[3]College of Oceanic and Atmospheric Sciences, Oregon State University, Corvallis, Oregon, USA.

0148-0227/09/2008JC004929$09.00

CR estuary and plume was discussed [*MacCready et al.*, 2009], a Lagrangian particle-tracking analysis of an upwelling event was described [*Banas et al.*, 2009], and characteristic patterns of the CR plume in summer are summarized [*Liu et al.*, 2009]. However, only a fraction of the available observations were used in this preliminary model validation, model sensitivity experiments are not reported, and a comprehensive evaluation of the model performance over both estuary and shelf is needed.

[4] With the advance of coastal ocean observation technologies, more and more observations are available for model validation and evaluation. There is a need to assess model skills in simulating different variables, and to give an overall skill score of model performance. *Oke et al.* [2002] provided an objective means for evaluating model skills based on mean squared error (*MSE*) reduction. This skill is relative to a control run, and is designed for identifying the best model simulation from a series of sensitivity experiments. The model skill of a single parameter may be quantified in terms of the *MSE* itself or root *MSE* (*RMSE*) [e.g., *Wilkin*, 2006]. Based on the *MSE*, a quantitative measure of model predictive skill was proposed by *Willmott* [1981]. As a nondimensional number, the Willmott skill (*WS*) scores for different physical parameters can be easily compared with one another. Recently, it was used to evaluate individual parameters simulated by estuarine [e.g., *Warner et al.*, 2005b] and coastal [e.g., *Wilkin*, 2006] circulation models, respectively. A comprehensive model skill for multiple physical parameters in the estuary-plume-shelf system has not been attempted.

[5] For a cross-scale model as presented by *MacCready et al.* [2009], evaluation of the model skill is particularly challenging. The model domain includes three dynamically distinct regions: estuary, plume and shelf. The model water depth extends from a few meters in the estuary to a few thousand meters in the open ocean; however, the main feature of interest, the CR plume, only appears in the top few meters. We expect model skill will vary with physical parameters, regions, depths and timescales. Moreover, the number of observations (and their degrees of freedom) will be very different in different regions, typically representing a compromise between cost, the interests of involved scientists, weather, and instrument loss. The skill averaging approach here is an attempt to make the average skill representative of our chosen scientific focus, by giving different weights (still admittedly arbitrary) to different regions.

[6] In this paper, the RISE model is qualitatively and quantitatively validated with an extensive set of observations (hydrography, moored ADCP current, bottom pressure, *T*, and *S*, sea level and HF radar current data) in a multiscale approach. Compared with *MacCready et al.* [2009], this paper offers the following new points: (1) Model/data comparison of the cross-scale system are enhanced by including moored *T*, *S*, and velocity data in the estuary and HF radar data on the shelf for the first time. (2) Detailed model sensitivity experiments are reported. (3) A new model skill assessment method is proposed and used to evaluate the model. Model skills for different physical parameters are quantitatively assessed over different spatial and timescales. An overall model skill is given by weighting the average skills of all the available parameters in different dynamic regions and timescales. (4) Relative skills of the ROMS and the NCOM models, and advantages and disadvantages of nesting within the NCOM model are discussed.

[7] The paper is arranged as follows: a brief description of the observations and definition of model skills is described in sections 2 and 3, respectively, followed by model configuration and sensitivity experiments in section 4. Model validation and evaluation are detailed in sections 5 and 6, respectively, followed by a summary in section 7.

2. Observations

[8] An array of three moorings were deployed along the 72 m isobaths on the shelf near the CR mouth (Figure 1), each equipped with a downward looking ADCP measuring current velocity every 0.5 m throughout the major part of the water column. On each mooring chain, there were three Seacat *T* and *S* sensors at 1, 5, and 20 m nominal depths, and a T logger chain measuring *T* every 5 m throughout the water column. Pressure sensors were also installed near the bottom of each site. All of these moored data were successfully collected from 22 June through 1 September 2004. In addition to the moorings, HF radars were deployed near the CR mouth to measure surface currents in the plume area. Detailed information about the mooring and HF radar data is given by E. P. Dever and P. M. Kosro (Observed tidal variability in the region of the Columbia River plume, submitted to *Journal of Geophysical Research*, 2009).

[9] Data from CTD casts along three hydrographic sections were collected from R/V *Wecoma* (Figure 1). Both the north and the south sections, which are located offshore of Grays Harbor (GH) and Cape Meares (CM), respectively, were sampled twice. A third section (RP) extended southwestward from the river mouth approximately along the plume axis as determined by drifters.

[10] Data from a number of moorings measuring *T* and *S* at selected depths are available in the estuary (Figure 1). Among these, three stations (red26, tansy and am169) were equipped with upward looking ADCP measuring current velocity throughout most of the water column. These data were obtained courtesy of A. Baptista and the CORIE team. CORIE (http://www.ccalmr.ogi.edu/CORIE) is a multipurpose observation and forecasting system for the Columbia River estuary and plume [*Baptista et al.*, 1998, 1999, 2005; *Baptista*, 2002]. Coastal sea level time series were downloaded from the CO-OPS/NOAA website http://tidesandcurrents.noaa.gov/station_retrieve.shtml?type = Historic+Tide+Data for tide gauge stations La Push, Toke Point in Grays Harbor, and Astoria in the CR estuary.

3. Defining Model Skills

[11] A widely used statistical measure of the agreement between the modeled and observed variables is the correlation coefficient (*CC*) or its square, coefficient of determination. It describes colinearity between the two time series. Often it is used in conjunction with a linear regression coefficient (*R*) [e.g., *Liu and Weisberg*, 2005]. The *MSE* is an alternative, commonly used measure of accuracy in numerical weather prediction [e.g., *Murphy*, 1992] and

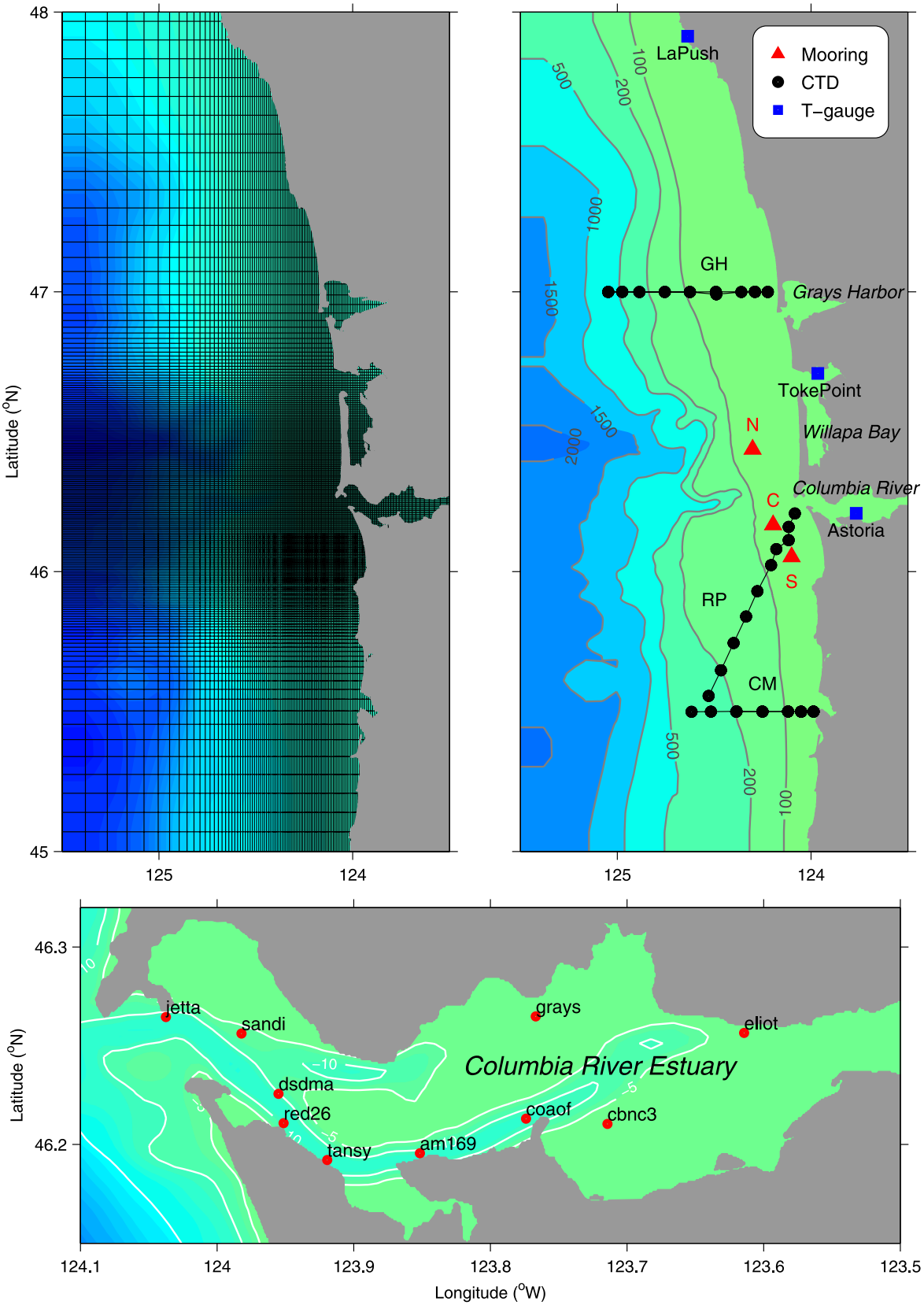

Figure 1. Maps of (top left) model grid, (top right) mooring, CTD and tide gauge stations on the shelf and (bottom) mooring locations in the Columbia River estuary, superimposed on bathymetric contours (units in m). Note that the complete grid also includes a "river" extending about 300 km to the east [*MacCready et al.*, 2009].

ocean modeling [e.g., *Oke et al.*, 2002; *Wilkin*, 2006]. $MSE = \langle (m - o)^2 \rangle$, where m and o are time series of the modeled and observed variables respectively, and $\langle\ \rangle$ denotes a mean. The *MSE* can also be written as

$$MSE = MB^2 + SDE^2 + CCE^2, \quad (1)$$

i.e., the *MSE* is comprised of three parts: the mean bias, $MB = \langle m \rangle - \langle o \rangle$, the standard deviation error, $SDE = S_m - S_o$, and the cross-correlation error, $CCE = [2S_mS_o(1\text{-}CC)]^{1/2}$, where $\langle m \rangle$ and $\langle o \rangle$ are the respective means; S_m and S_o are the respective standard deviations [e.g., *Oke et al.*, 2002]. The skill score of an experiment is defined as a reduction of the MSE with respect to a reference experiment [e.g., *Murphy*, 1992; *Oke et al.*, 2002]:

$$SS = 1 - MSE/MSE_R \quad (2)$$

where the subscript R denotes a reference experiment. Positive skill ($SS > 0$) indicates an improvement, while negative skill ($SS < 0$) means a worse result; $SS = 0$ means no improvement, and $SS = 1$ means perfect skill. *SS* is a relative number, and can be used to choose the best model run among a series of sensitivity experiments. *Oke et al.* [2002] applied this skill to a multiple parameter case (velocity, T and S), where each *MSE* is normalized by its variance so that an average skill can be obtained from all the parameters.

[12] Also based on the *MSE*, a quantitative model skill was presented by *Willmott* [1981],

$$WS = 1 - MSE \Big/ \left\langle (|m - \langle o \rangle| + |o - \langle o \rangle|)^2 \right\rangle. \quad (3)$$

The highest value, $WS = 1$, means perfect agreement between model and observation, while the lowest value, $WS = 0$, indicates complete disagreement. Recently, this was used to evaluate ROMS in the simulation of multiple parameters in the Hudson River estuary [*Warner et al.*, 2005b] and on the southeast New England Shelf [*Wilkin*, 2006]. The Willmott skill will be used to quantify model performance in simulating different parameters from the best model run.

4. Model Configuration and Sensitivity Experiments

4.1. Model Configuration

[13] The ocean circulation hindcasts of the CR estuary-plume-shelf system were performed using ROMS (Rutgers version 2.2), a free surface hydrostatic, primitive equation model in widespread use for estuarine and coastal studies [e.g., *Haidvogel et al.*, 2000, 2008; *MacCready et al.*, 2002; *Garcia-Berdeal et al.*, 2002; *Marchesiello et al.*, 2003; *Lutjeharms et al.*, 2003; *Peliz et al.*, 2003; *Li et al.*, 2005; *Warner et al.*, 2005b; *Wilkin*, 2006; *He and Wilkin*, 2006; *Dong and McWilliams*, 2007; *Wilkin and Zhang*, 2006; *Barth et al.*, 2008]. The major (shelf and estuary) part of the model grids and the bathymetry are shown in Figure 1; the actual model domain also includes a long "river" from the east, which allows tidal energy to propagate freely upstream [see *MacCready et al.*, 2009, Figure 1]. Horizontal resolution of the grid is ~400 m in the estuary and near-plume region where the grid is most dense, telescoping to ~9 km in the far corners. In the vertical direction, 20 sigma layers are chosen between the free surface and the bottom. Maximum depth in the domain is 2155 m, and the minimum depth is set to 3 m, so wetting and drying of the intertidal regions is not included. Bathymetry data are from the GEODAS [*Divins and Metzger*, 2002] data set. The model bathymetry is smoothed for numerical accuracy so that the total fractional change in depth over a grid cell is less than 0.8. This still fills in the deeper channels in the CR estuary. To remedy this, the channel along the southern thalweg is artificially deepened to match depths from navigational charts. The baroclinic time step is 51.75 s, and the barotropic time step is 20 times shorter. The simulation period is about 3 months in summer 2004 (from 3 June to 1 September).

[14] Model forcing includes river (daily river flow and T from the USGS gauging station at Beaver Army Station, http://waterdata.usgs.gov/or/nwis/dv/?site_no = 14246900&agency_cd = USGS), tides (surface height and depth averaged velocity using 10 tidal constituents from the TPXO6.0 analysis [*Egbert and Erofeeva*, 2002]), hourly winds, surface pressure, air temperature, humidity, and shortwave and downward longwave radiation (from Northwest Modeling Consortium MM5 regional forecast model [*Mass et al.*, 2003]). An assessment of MM5 model winds over the region are given by *Tinis et al.* [2006]. Surface fluxes of momentum and heat are calculated in ROMS using bulk formulations [*Fairall et al.*, 1996a, 1996b; *Liu et al.*, 1979]. Net shortwave radiation and downward longwave radiation are given as external forcing, and then upward longwave, sensible, and latent heat fluxes are calculated in the model.

4.2. Open Boundary Forcing: NCOM Versus HYCOM

[15] The model is initialized with, and one-way nested within, the Navy Coastal Ocean Model—California Current System (NCOM-CCS) regional model [*Shulman et al.*, 2004]. The NCOM-CCS assimilates ~14 km daily Multi-Channel Sea Surface Temperature (MCSST); it is forced by atmospheric fluxes from a high-resolution Coupled Ocean Atmospheric Mesoscale Prediction System (COAMPS) reanalysis product [*Kindle et al.*, 2002], and one-way nested in the 1/8° global NCOM [e.g., *Barron et al.*, 2006; *Kara et al.*, 2005]. As in the global NCOM, the NCOM-CCS assimilates three-dimensional T and S observations [*Shulman et al.*, 2007] derived from the Modular Ocean Data Assimilation System (MODAS [*Fox et al.*, 2002]). The NCOM-CCS model domain extends from 30°N to 49°N with a horizontal resolution of about 9 km. We use output from NCOM-CCS version h002 _expt_20.6, which does not have the CR plume. Such open boundary forcing provides useful information that is otherwise not represented by our limited domain model, for example, offshore eddy activity, coastal trapped waves, and remote forcing of the seasonal baroclinic coastal jet [*Hickey et al.*, 2006]. Coastal trapped waves have been shown to be particularly important in the study area [*Battisti and Hickey*, 1984]. Dynamical open boundary conditions for the free surface and depth-averaged momentum are given by the *Chapman* [1985] and *Flather*

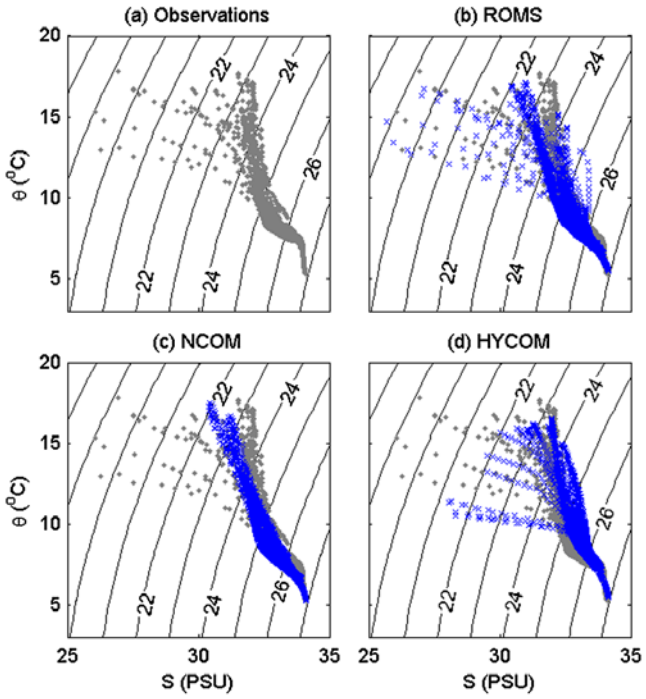

Figure 2. T-S diagrams for CTD data at the five hydrographic sections from observations (gray) with model values overlain (black): (a) Observations, (b) ROMS, (c) NCOM, and (d) HYCOM.

[1976] formulations. Three-dimensional fields of velocity and tracers are treated with a radiation boundary condition (all as per *Marchesiello et al.* [2001]). Over a six grid point wide region on the open ocean boundaries the *T*, *S* and momentum fields are relaxed to their NCOM-CCS values: over 10 days at the boundary and ramping to 60 days at six grid points in. The 12-hourly NCOM-CCS data (*T*, *S*, velocity and sea level) were low-pass-filtered (2.5-day running mean) and 2-day subsampled before they were interpolated onto the ROMS grid for open boundary forcing.

[16] Another candidate for open boundary forcing is the global Hybrid Coordinate Ocean Model (HYCOM) output, which is available through the multiinstitutional HYCOM Consortium data server http://hycom.coaps.fsu.edu/thredds/dodsC/glb_analysis.html. The global HYCOM [e.g., *Bleck*, 2002; *Halliwell*, 2004; *Chassignet et al.*, 2007] hindcast uses the Navy Coupled Ocean Data Assimilation (NCODA) system [*Cummings*, 2005], on a Mercator grid between 78°S and 47°N (1/12° equatorial resolution, a horizontal resolution of about 7 km on average), with a bipolar patch for regions north of 47°N. The global HYCOM includes freshwater forcing for the CR plume. Similar to that for NCOM-CCS data, the daily HYCOM data were 3-day low-pass-filtered and 2-day subsampled before they were mapped onto the ROM grid.

[17] Here we make some comparisons of the two candidate models for open boundary forcing to observations. Both NCOM-CCS (hereafter NCOM) and HYCOM original *T* and *S* data are extracted for the times of CTD sections, and the *T*-*S* diagrams are compared with observations (Figures 2c and 2d). HYCOM captures some low salinity (<30) signature of the plume, while NCOM does not. The NCOM forcing does not include river input, which will be provided by our ROMS model. For high-density water ($\sigma_\theta > 26$ kg m^{-3}), HYCOM better represents the observations than NCOM; however, for the density range 24 kg m^{-3} < σ_θ < 26 kg m^{-3}, the main water type on the shelf, NCOM is marginally better than HYCOM. Section GH (47°N) is "upstream" of the plume in the California Current which flows southward on the shelf. Comparison of the *T* and *S* distributions at Section GH (Figure 3) also indicates that NCOM has overall better agreement with the observations (both surface freshness and deep salt) than HYCOM. However, subsurface salinity from both models is generally lower than observed and NCOM salinity is lower than HYCOM salinity. Velocity comparisons of NCOM and HYCOM to observations at the three moorings do not indicate that one model is obviously better than the other (figures not shown).

[18] Model results near their open boundaries are generally less reliable than those in the interior. Our model domain extends from 45°N to 48°N, and thus lies within the NCOM domain (which goes to 49°N). However, our model domain crosses the northern boundary of the global HYCOM (47°N) and thus requires data from both the global HYCOM and its north polar patch. Data quality may not be the same for both model components; for example, data assimilation is not performed for the polar patch. In sensitivity experiments detailed below (Table 1, cases *g* versus *h*) we find that NCOM open boundary forcing gives markedly better results than HYCOM. This indicates NCOM is better suited for our applications than HYCOM. Therefore, NCOM was chosen for open boundary forcing.

4.3. Sensitivity Experiments

[19] The model is tested for its sensitivity to different parameter choices of light penetration depth, vertical mixing schemes, nudging terms, and open boundary forcing. Details of the experiments are summarized in Table 1. Model relative skill (*SS*) defined in equation (2) is used to assess the sensitivity in three areas: estuary, near- and far-field plume (the near-field plume refers to the region within ~1 tidal excursion of the river mouth, and the far-field refers to the outer part of plume [*MacCready et al.*, 2009]). These two shelf areas are further divided into two vertical layers using 20 m as the dividing level, making a total of five analysis regions. Although both the type and quantity of the observations are different in the five regions for the modeling period, the choice of the five groups is based more on scientific interest than on data availability. The main purpose of this estuary-plume-shelf system is to hindcast the coastal circulation and water property variations with a focus on the CR plume, a feature that is frequently only a few meters thick. Thus, more weight is put on the near surface layer in assessing model skill. We do this by separating the shallower surface layer (e.g., top 20 m) from deep layers on the shelf and assigning equal weights to their model skills. Similarly, both the estuary and the near-field should have more weight than the far-field plume for model evaluation. These criteria are roughly satisfied when the average model skills in these three regions are equally weighted in calculating the overall model skill, despite the fact that the volumes of the estuary and the near-field plume are much smaller than that of the far-field plume. Also note that the model may have different skills for different

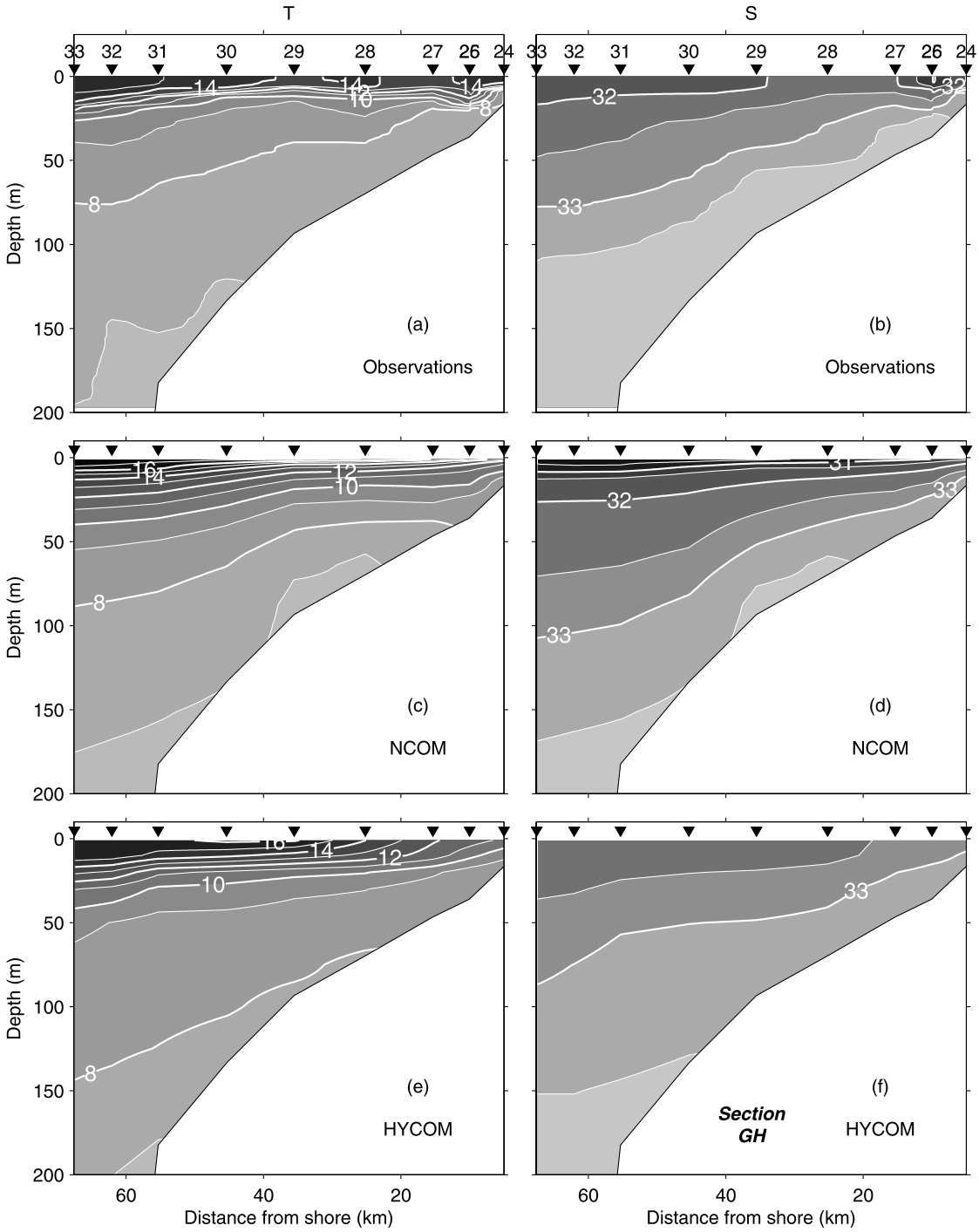

Figure 3. Comparison of (left) *T* and (right) *S* at Section GH between the (a and b) observations and the (c and d) NCOM and (e and f) HYCOM model outputs. The small triangles designate the CTD cast locations, and the cast numbers are shown above the triangles. This is the first of the two GH CTD lines. The line was sampled during a period of light upwelling wind following a relaxation event (∼5 days) as shown by the CR plume water in the top right of Figure 3b.

physical variables. Because it is difficult to assign objectively the weights of the various skills, without losing generality, they will be equally weighted. A simple average of *SS* is obtained for all the available observations in each of the above defined five regions, and an overall *SS* is obtained for each experiment by averaging the *SS* across the five regions (Table 1). Only the 36-h low-pass-filtered time series are used for the sensitivity experiments in assessing the *SS*. Hereafter subscripts "lp" and "hp" designate 36-h low- and high-pass-filtered time series, respectively.

Table 1. Summary of Sensitivity Experiments[a]

Experiment	Jerlov Water Type	Vertical Mixing Scheme	Nudging in the OB Regions	Initial Conditions and OB Forcing	Average Skill Score					
					Far Field		Near Field			
					<20 m	>20 m	<20 m	>20 m	Estuary	Average
g	5	LMD	Tr, M2, M3, Z	NCOM	0.00	0.00	0.00	0.00	0.00	0.00
h	5	LMD	Tr, M2, M3, Z	HYCOM	0.25	−0.14	−0.17	−2.16	0.06	−0.43
i	5	MY2.5	Tr, M2, M3, Z	NCOM	0.03	−0.03	0.08	0.09	0.08	0.05
o	5	GLS	Tr, M2, M3, Z	NCOM	0.14	−0.05	0.20	0.21	−0.01	0.10
s	1	GLS	Tr, M2, M3, Z	NCOM	−0.02	−0.07	−0.03	−0.05	0.01	−0.03
t	3	GLS	Tr, M2, M3, Z	NCOM	0.12	−0.11	−0.06	0.08	0.01	0.01
u	5	GLS	Tr, M2, M3	NCOM	0.14	−0.05	0.20	0.21	−0.01	0.10
v	5	GLS	Tr, M3	NCOM	0.11	−0.03	0.15	0.19	−0.01	0.08
w	5	GLS	Tr	NCOM	0.22	−0.07	0.04	−0.28	−0.03	−0.02
x	5	GLS	Tr, M2	NCOM	0.23	−0.04	0.03	−0.29	−0.04	−0.02
y	5	GLS	Tr, Z	NCOM	0.22	−0.07	0.04	−0.28	−0.03	−0.02
z	5	GLS	No nudging	NCOM	−0.06	−1.36	−0.68	−4.85	−0.09	−1.41

[a]In Table 1 only, Tr, M2, M3 and Z stand for tracers, two-dimensional depth-averaged momentum, three-dimensional momentum, and sea level (zeta), respectively. OB designates open boundary. Note that "<20 m" refers to depth, so this is the surface layer. *SS*, skill score.

4.3.1. Sensitivity Experiment on Jerlov Water Type

[20] The depth of penetration of shortwave radiation is a function of water clarity, represented by the Jerlov water type [*Jerlov*, 1976], which is assigned integer values from 1 through 5 in ROMS, with 1 representing the clearest water and 5 corresponding to the shortest penetration. For a highly productive coast influenced by the CR plume, water clarity is relatively poor in summer, and the vertical penetration of shortwave radiation is short. Experiments with different Jerlov water types indicate that water types 1 and 3 result in warmer water in the subsurface layer (20 m level) and cooler water at the surface (3 m level), and water type 5 gives a T distribution that has the best agreement with the Seacat data at the 3 and 20 m depths from the moored array (figure not shown). Using water type 5, the model skill is significantly improved in the near-field plume (both the near surface and the lower layers) and in the upper layer of the far-field plume (Table 1, cases *o* versus *s* and *t*).

4.3.2. Sensitivity Experiment on Turbulence Closures

[21] Sensitivity was tested for three turbulent closures: the Large-McWilliams-Doney (LMD [*Large et al.*, 1994]), the Mellor-Yamada 2.5 level turbulence closure (MY2.5 [*Mellor and Yamada*, 1982]) and the Generic Length Scale formulation (GLS [*Umlauf and Burchard*, 2003]). For the LMD scheme, both diffusivity due to shear instability and convective mixing due to shear instability are added, surface and bottom boundary layer KPP mixing are turned on, and nonlocal transport is included. Key parameters for LMD include the critical gradient Richardson number $R_{i0} = 0.7$, the double-diffusive density ratio $R_{\rho 0} = 1.9\ m^2\ s^{-1}$, the Brunt-Vaisala frequency limit for convection $-2.0 \times 10^{-5}\ s^{-2}$, the scaling factor for double diffusion of temperature in salt fingering case 0.7, and the molecular viscosity parameters $(\mu, \mu_c^0, \mu_m^0, \mu_s^0, \mu_f) = [1.5 \times 10^{-6}, 10^{-2}, 10^{-3}, 10^{-3}, 10^{-3}]\ m^2\ s^{-1}$. In MY2.5, the *Kantha and Clayson* [1994] stability formulation is used, and key parameters include turbulence closure $(A_1, A_2, B_1, B_2, C_1, C_2, C_3, E_1, E_2) = [0.92, 0.74, 16.6, 10.1, 0.08, 0.7, 0.2, 1.8, 1.33]$, the lower bound on *Galperin et al.* [1988] stability function $G_{h0} = 0.0233$, the time filter coefficient is 0.05, the scale for vertical mixing of turbulent energy $S_q = 0.2$, the upper bound on the turbulent length scale $l_{max} = 0.53$, and the lower bound on turbulent energy $q_{min} = 1.0 \times 10^{-8}$. Horizontal smoothing of buoyancy/shear is activated in both MY2.5 and GLS schemes. Following *Warner et al.* [2005a], the parameters of the k-ε version of the GLS formulation include the Canuto A stability functions [*Canuto et al.*, 2001] and minimum values of vertical eddy viscosity and diffusivity of $5 \times 10^{-6}\ m^2\ s^{-1}$. Other key parameters for the GLS scheme are $(G_{h0}, G_{hcri}, L_1, L_2, L_3, L_4, L_5, L_6, L_7, L_8, G_{hmin}, E_2) = [0.0329, 0.03, 0.107, 0.0032, 0.0864, 0.12, 11.9, 0.4, 0.0, 0.48, -0.28, 1.33]$. The meanings of these parameters can be found in ROMS package program mod_scalars.F. It is found that the GLS scheme results in significant improvements over the other two schemes in the near-field plume, and has the best overall model skill for the whole model domain (Table 1, cases *o* versus *g* and *i*).

4.3.3. Sensitivity Experiment on Nudging Terms and Open Boundary Forcing

[22] In early tests for open boundary conditions, we found that clamped conditions for tracers and 3-D momentum returned worse results than radiation conditions [e.g., *Marchesiello et al.*, 2001]. Thus, the latter are used in conjunction with the *Chapman* [1985] and *Flather* [1976] conditions for the free surface and the 2-D momentum. The model fields are relaxed to the NCOM data on timescales of 3 days at the open boundaries, ramping to 60 days 6 grid points into the interior. Sensitivity experiments show that nudging is necessary for this small model domain (Table 1, comparing cases *o, u, v, w, x, y* and *z*). Nudging different variables results in different improvements in model skills, and the best nudging result is found to be complete nudging (nudging of all the available variables: tracers, 2-D and 3-D momentum). Note that nudging of the free surface does not make any difference when nudging of 2-D momentum is turned on (see the identical results in cases *o* and *u*). Among the five geographic groups, the estuary is least sensitive to nudging choices. Experiments on the choice of open boundary forcing data, NCOM versus HYCOM, indicate

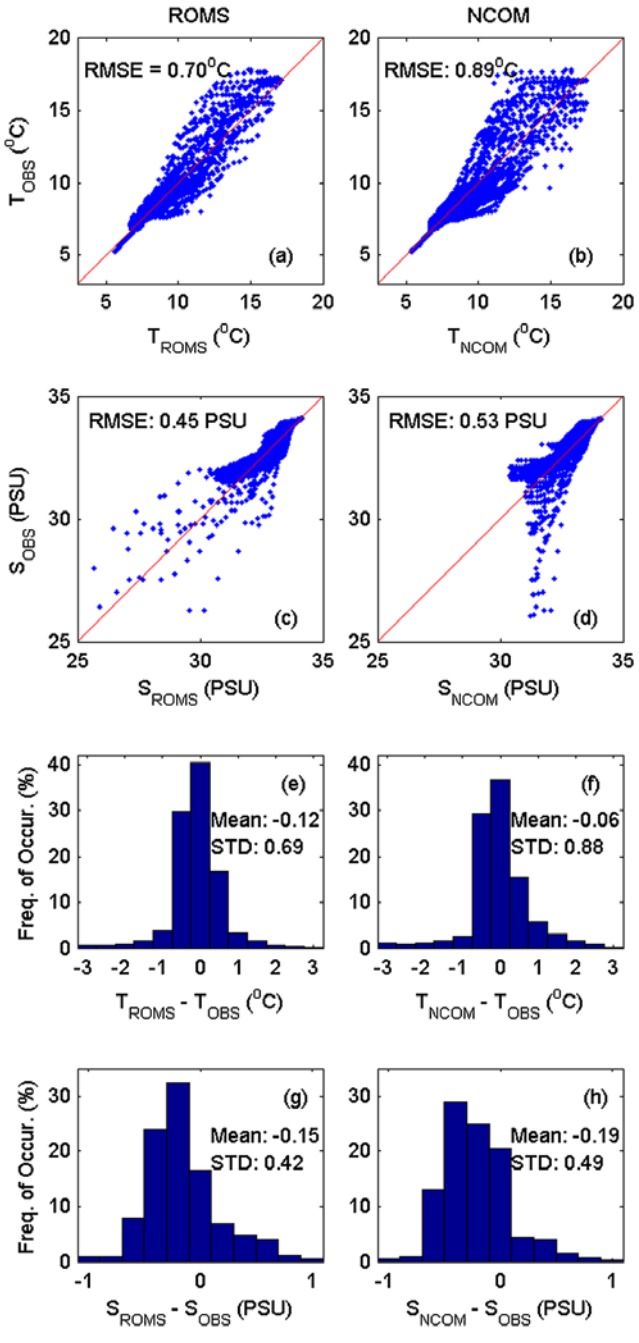

Figure 4. Scatter diagrams of observed versus (left) ROMS and (right) NCOM modeled (a and b) T and (c and d) S for stations on the CTD sections. Histograms of model (e and f) T and (g and h) S errors are also shown.

that the former is better suited for our applications (Table 1, cases *g* versus *h*).

5. Model-Data Comparisons

[23] In the preliminary model validation, *MacCready et al.* [2009] compared vertical profiles of the temporal mean, the standard deviation and the correlation coefficient of the modeled and observed velocity, T and S at the three moorings on the shelf. They also compared high-pass and low-pass-filtered velocity, T and S time series at selected depths for the central mooring, and made qualitative comparisons on the five CTD sections. In this paper, more extensive model validations are conducted, not only on the shelf, but also in the estuary. For CTD data, error statistics are added, and for moored velocity data, barotropic and baroclinic currents are examined separately. Model-data comparisons are performed in a multiscale approach: in the estuary, near- and far-field plume, respectively, and for time series data, on both tidal and synoptic timescales. Comparison with the HF radar data is on synoptic timescales only. Model output from the best model run (case *o* in Table 1) is used for comparison unless otherwise noted.

5.1. Comparisons With Hydrographic Data

[24] Observed and modeled T and S along Sections GH, RP and CM are shown in the T-S diagrams (Figures 2a and 2b). Our model has better overlay with observations than either the NCOM (Figure 2c) or HYCOM (Figure 2d) models, for low salinity (<30) water of the plume and the main type of water on the shelf: density range 24 kg m^{-3} < σ_θ < 26 kg m^{-3}. The same T and S are shown in scatter diagrams (Figures 4a–4d). Scattered points generally cluster around the slope = 1 diagonal line, indicating good agreement between the model and the observations. Low S points tend to spread farther away from the line. *RMSE* shows that ROMS provides a better fit to the data than NCOM (*RMSE* of T is 0.70 versus 0.89°C, and *RMSE* of S 0.45 versus 0.53, for ROMS and NCOM, respectively). This is expected because ROMS includes the Columbia River, while NCOM does not.

[25] The errors of the T and S hindcasts are slightly biased (Figures 4e–4h), with a mean bias of −0.12°C for T and −0.15 for S and a standard deviation of 0.69°C for T and 0.42 S, respectively. These biases may originate from the NCOM data through the initial conditions and the open boundary forcing, although improvements of ROMS over NCOM are seen in the histograms of errors, i.e., higher frequency of occurrence of smaller errors (Figures 4e–4h). Moreover, mixed layer depth predicted using NCOM is deeper than observed on the central CA coast (Monterey Bay area) during the summer [*Shulman et al.*, 2004, 2007].

[26] Qualitative model-data comparison of S along the five sections is discussed by *MacCready et al.* [2009]. Here comparison is made along the river plume axis only, for both T and S (Figure 5). Highly stratified CR plume water in the upper 20 m layer is seen in both model results and data. The upward tilt of the 8°C isotherm and the 33 isohaline is captured by the model. Model T is slightly colder than observed (~1°C) at the surface, and warmer near the bottom (<1°C). Model S is slightly lower throughout the water column, and the model plume extends farther south than the observed plume by about 18 km as revealed by S = 30 contours. Because NCOM does not have rivers or tides, inclusion of the CR in our model improves S modeling (surface S < 31). Tidal advection and mixing are very important in the estuary and within ~30 km of the CR mouth, but should have little effect over much of this CTD section.

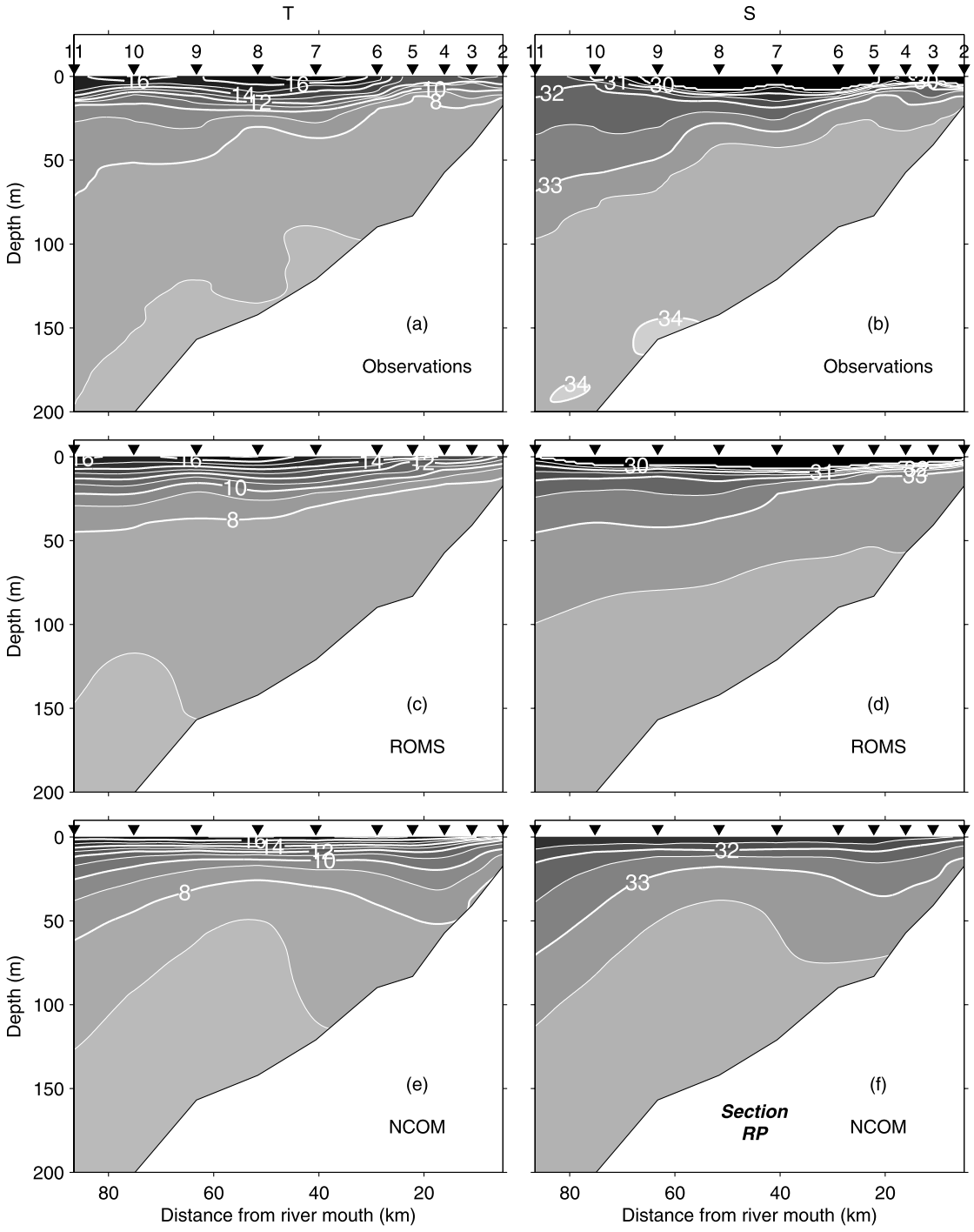

Figure 5. Comparison of (left) temperature and (right) salinity along Section RP between the (a and b) observations and (c and d) ROMS and (e and f) NCOM model outputs. The small triangles designate the CTD cast locations, and the cast numbers are shown above the triangles.

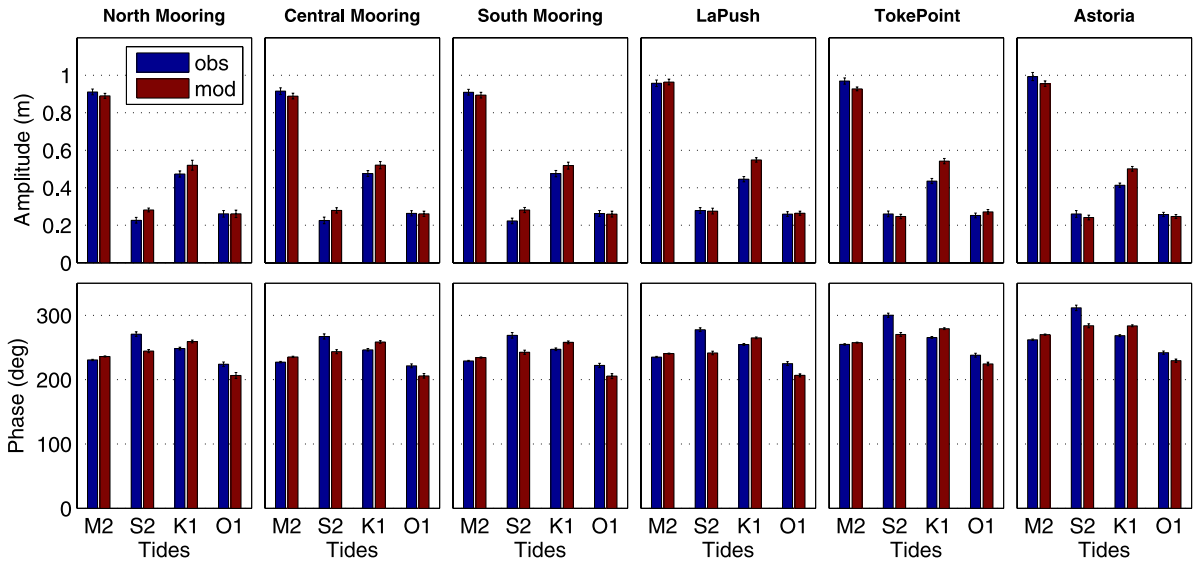

Figure 6. (top) Amplitude and (bottom) phase of sea level for the four major tidal constituents M2, S2, K1 and O1 at the north, central and south moorings, and the three tide gauge stations (La Push, Toke Point, and Astoria). Bottom pressure records are used to approximate sea level at the three mooring sites.

5.2. Tidal Variations on the Shelf and in the Estuary

5.2.1. Tides

[27] Modeled free surface time series are compared to sea level records for major tidal constituents at tide stations La Push, Toke Point, and Astoria, respectively. The bottom pressure records at the three moorings are converted to equivalent sea level assuming a constant density throughout the water column, and then compared with the model free surface. The amplitudes and phases of the four major tidal constituents M2, S2, K1 and O1 are calculated using the T_Tide toolbox [*Pawlowicz et al.*, 2002]. The M2 tide has the largest amplitude (about 0.9∼1 m), and the K1 tide is secondary (∼0.5 m); these are seen in both model and data (Figure 6). The model error of the M2 amplitude is generally less than 3% at the three offshore moorings and at the La Push tide station, and is about 4% in the estuaries (Toke Point and Astoria stations). A slight amplification of the M2 tidal amplitude from shelf to coast is seen in both model and observations, and increased phase lags for all the four major tides are seen in the estuaries (Figure 6). Prediction of tidal height at these stations is only a weak test of model performance, however, because they are all so similar to the tide at the open boundaries.

5.2.2. Barotropic Tidal Currents

[28] Barotropic tidal current ellipses of the four major tidal constituents are also obtained using the T_Tide tool-

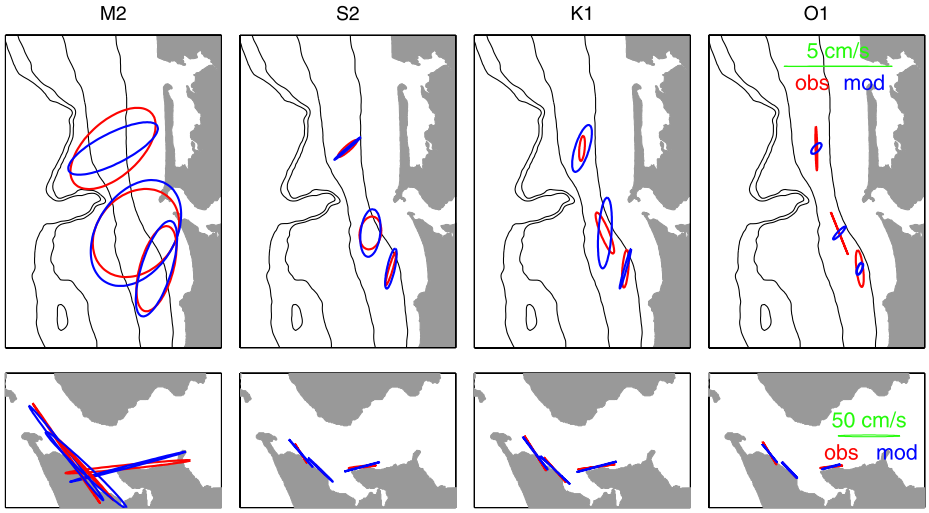

Figure 7. Barotropic tidal ellipses of the four major tidal components (from left to right, M2, S2, K1 and O1) on the (top) shelf and in the (bottom) Columbia River south channel. Note that different velocity scales are used for the shelf and the estuary, respectively.

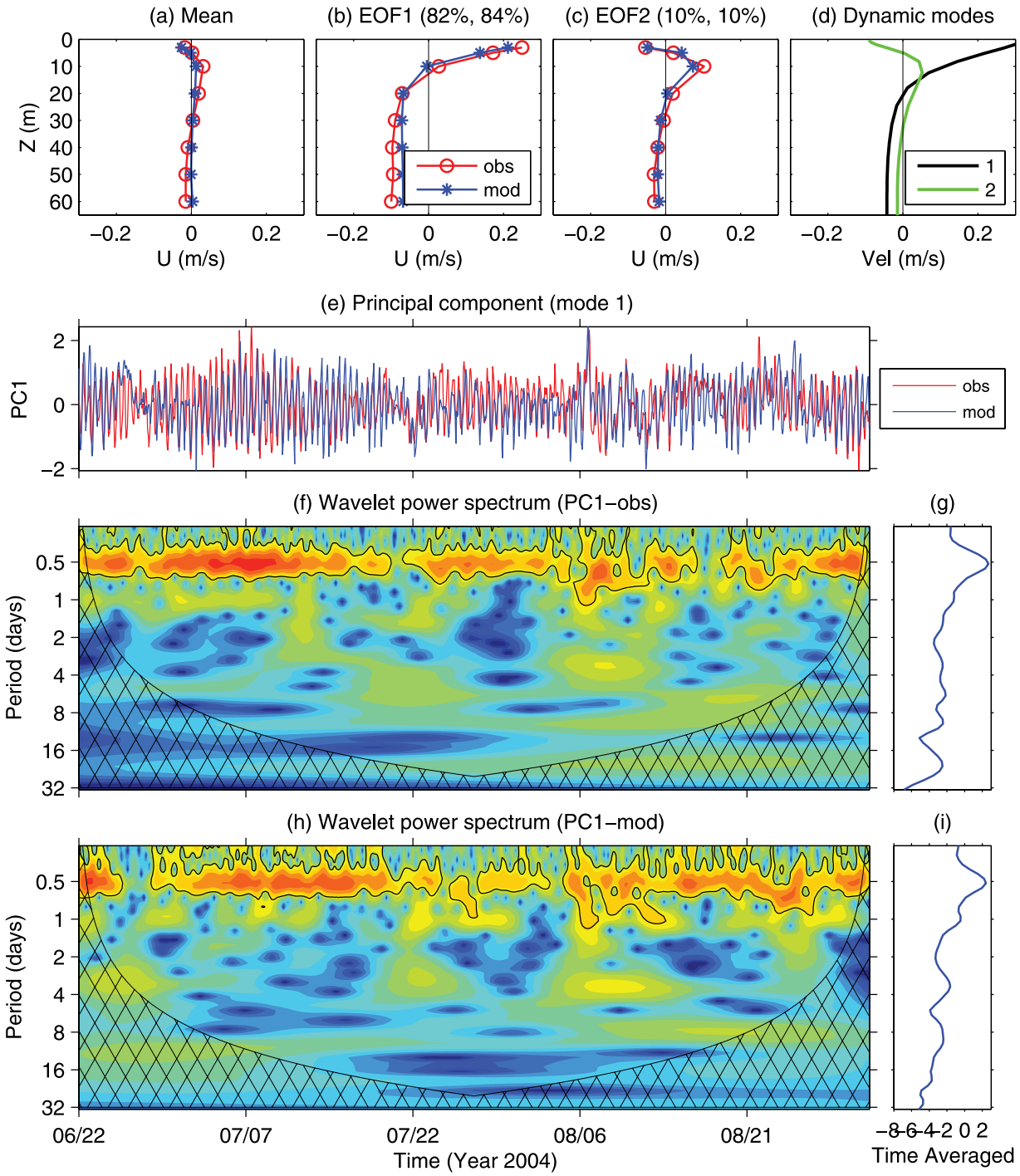

Figure 8. EOF analysis of the hourly baroclinic east-west velocity component vertical profiles at the central mooring: (a) mean and (b and c) first two EOF modes of the *u* profiles, and (e) the first mode principal component (PC) time series. The variances accounted for by the EOF modes are indicated as two percentages in the parentheses on top of the EOF plots; the first and the second are for the observed and the modeled time series, respectively. (d) The first two dynamic modes of the horizontal velocity estimated from modeled stratification. Wavelet analysis of the first mode principal component time series (Figures 8f–8i): (f and h) wavelet power spectrum and (g and i) time-averaged power spectrum for the observed (Figures 8f and 8g) and modeled (Figures 8h and 8i) time series. Red and blue indicate high and low wavelet power spectrum values, respectively; both the time-varying and the time-averaged wavelet power spectra are shown in the base 2 logarithm. The regions of greater than 95% confidence are shown with black contours. Cross-hatched regions on either end indicate the "cone of influence," where edge effects become important.

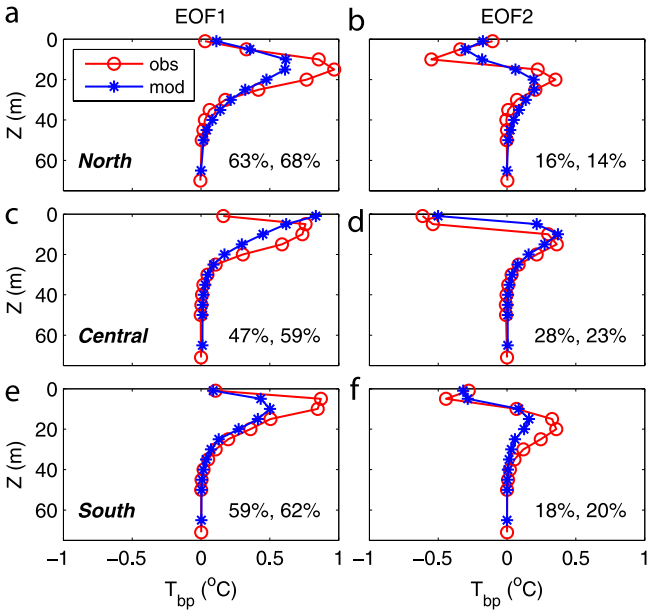

Figure 9. (left) Mode 1 and (right) mode 2 EOFs of the T_{bp} profiles at the (a and b) north, (c and d) central and (e and f) south moorings on tidal timescales (6–36 h band-pass-filtered). The variances accounted for by the EOF modes are indicated as two percentages in each plot; the first and second are for observed and modeled time series, respectively.

box. Barotropic tidal currents are dominated by the M2 constituent with a major axis of 3∼4 cm s^{-1} on the shelf and 80∼100 cm s^{-1} in the south channel of the CR estuary. Modeled major axis error of M2 tidal ellipse is 0.1∼0.4 cm s^{-1} on the shelf (Figure 7). Those for K1 and O1 have significant model-observation differences, likely because those tidal components are weaker and the signal-to-noise ratios are generally lower. In contrast, the comparisons are much better in the estuary (Figure 7). The tidal ellipses are highly polarized along the south channel of the CR estuary for all the four major tidal constituents. Modeled major axis error of M2 tidal ellipse is 2.0 ∼ 6.3 cm s^{-1}. A detailed analysis of the tidal currents derived from the recent observations in this area is reported by Dever and Kosro (submitted manuscript, 2009).

5.2.3. Baroclinic Tidal Currents

[29] Baroclinic tidal currents on the shelf are also dominated by the M2 semidiurnal components; the major axes of the M2 ellipses near the surface are about 2∼3 times larger than those at middle and near bottom depths (figure not shown). Note that the CR plume is a source of strong nonlinear internal wave generation [*Nash and Moum*, 2005; *Orton and Jay*, 2005; *Pan et al.*, 2007]; conventional harmonic analysis, assuming the time series to be stationary, may not work well for real ocean currents in this region where the stratification changes drastically depending on the plume location. For nonstationary time series, wavelet analysis is more appropriate. Previous wavelet analysis of the tidal records in this area can be seen in the works of *Jay and Flinchem* [1997, 1999].

[30] The time domain empirical orthogonal functions (EOF) method is widely used in oceanography to extract dominant patterns/structures from time series of spatial maps [e.g., *Davis*, 1976]. It separates the data set into data-dependent, empirical orthogonal modes. Generally speaking, each mode has an associated variance, a spatial pattern EOF, and a principal component (PC) time series. Leading EOFs show spatial patterns of dominant variability in the data, and their PCs give their weightings, which are functions of time. The EOF is used to extract dominant vertical structure of the baroclinic tidal currents from hourly time series of velocity profiles at one mooring site; the associated PC is further analyzed using the wavelet power spectrum to illustrate its variation in both time and frequency domains.

[31] The EOF analysis is performed on the hourly velocity vertical profiles (u component, with the depth-averaged component removed) at the central mooring for modeled and observed time series (Figures 8a–8e). Analysis with the v component (not shown) gives almost the same result. Although the mean currents are relatively weak (<5 cm s^{-1}) throughout the water column (Figure 8a), the first mode currents are stronger and vertically sheared with reversed signs in surface and subsurface layers (Figure 8b). The model successfully captures this pattern. In particular, the first two EOF modes of the vertical profiles of velocity from both the model output and the observations resemble the first two internal wave modes (horizontal velocity vertical profiles) derived from the average stratification of the model water (Figures 8b–8d). The principal component time series show the temporal variation of the baroclinic velocity (Figure 8e). To examine the time series in the frequency domain, wavelet power spectra of the first mode principal component time series are calculated using the wavelet toolbox of *Torrence and Compo* [1998] including the wavelet power spectrum rectification [*Liu et al.*, 2007]. The model successfully reproduces the dominant variability in the semidiurnal frequency band (centered at a period of 0.5 day), as well as the modulation of the semidiurnal baroclinic tidal currents by synoptic and spring-neap variations (Figures 8f–8i). The modeled wavelet spectrum of the semidiurnal baroclinic tidal currents is frequently weaker than the observed (Figures 8f and 8h), indicating that the model underestimates baroclinic currents; this is also seen in the time averaged wavelet power spectra (Figures 8g and 8f). The speed of the first mode wave is 0.41 m s^{-1}, and the wavelength for the M2 tidal period is about 18 km; our model domain is large enough to accommodate multiple wavelengths over the shelf. However, the effects of remotely forced internal waves are not represented well through the open boundary forcing, because NCOM does not include tides.

[32] A strong baroclinic signal is to be expected here, due to the intense tidal currents and stratification of the CR plume. Whether or not this signal propagates away from the CR mouth as a first mode internal wave is a question for future research. *Nash and Moum* [2005] and *Orton and Jay* [2005] have already observed that in part the plume front generates (O 100 m wavelength) nonlinear internal waves; however the model grid does not resolve these features.

5.2.4. *T* and *S* Variations on Tidal Timescales

[33] In contrast to the velocity field, the T in the observed near-field plume has a smaller variance on tidal timescales than on synoptic timescales (Figure 9). The first mode EOFs

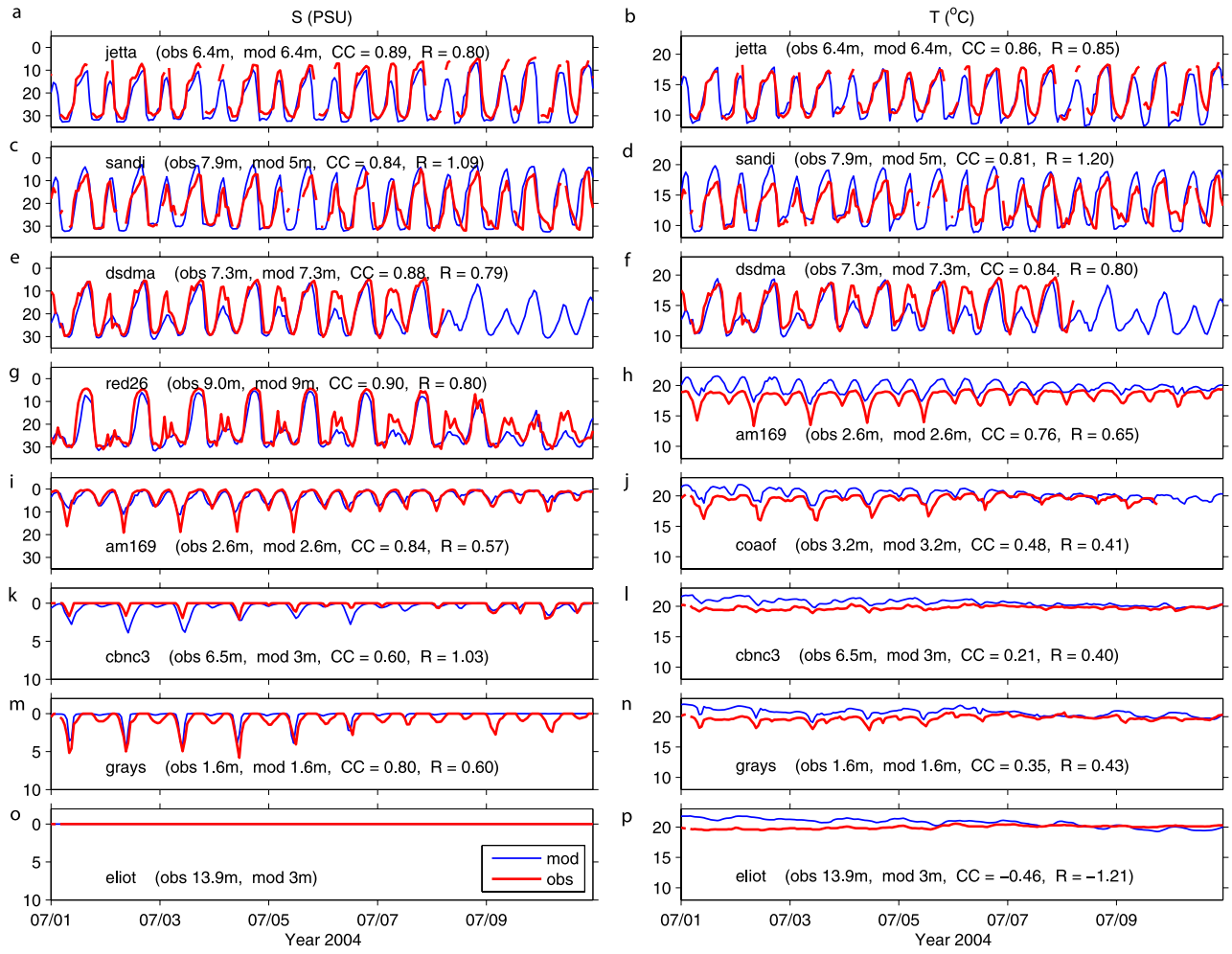

Figure 10. Observed and modeled hourly (left) S and (right) T time series in the Columbia River estuary. Note the y axis scales are different for Figures 10k, 10m, and 10o. When an actual data depth is not available in the model due to the limited accuracy in resolving the bathymetry, a closest data level is used instead, e.g., at Station sandi the T and S sensors were at 7.5 m, but the deepest available model output at that grid point is 5 m. The station name, sensor depth, model data depth, correlation coefficient (CC), and regression coefficient (R) are listed in each plot. R is calculated from a linear regression model: $T_{mod} = R \times T_{obs}$ + residual. Both observed and modeled salinity are identically 0 at Station eliot (Figures 10o) during this period. The correlation coefficients are significant at 95% confidence except for the temperature at Stations cbnc3, grays and eliot (Figures 10l, 10n, and 10p). From top to bottom, the stations are located approximately from west (river mouth) to east in the estuary.

of the 6∼36 h band-pass-filtered T profile time series (hereafter T_{bp}) from the T loggers show that the maximum T_{bp} variation in the water column is located in the subsurface layer (5∼20 m depth). This feature is partly present in the model. This vertical structure of the T_{bp} variability is associated with the plume front/pycnocline and the strong baroclinic tidal currents in the near-field plume. The modeled maximum T_{bp} variation in the subsurface layer is about half of that observed.

[34] The dominant tidal features in the estuary are reproduced by the model. In particular, T and S in the CR estuary exhibit strong tidal variations near the river mouth and in the south river channel; and T and S tidal fluctuations decrease toward the upper end of the estuary (Figure 10). These T and S features are seen in both model and data. The low correlation coefficients of the T at some stations (e.g., coaof, cbnc3, grays and eliot) may be misleading, because different linear trends (i.e., low-frequency offsets) are seen between the modeled and observed time series, which significantly reduce the values of the linear correlation.

5.3. Synoptic Variations on the Shelf and in the Estuary

5.3.1. Comparisons of Velocity, T and S on the Shelf

5.3.1.1. Comparison With Moored ADCP Velocity Data

[35] As with tidal timescales, the velocity time series at the moorings are decomposed into barotropic and baroclinic

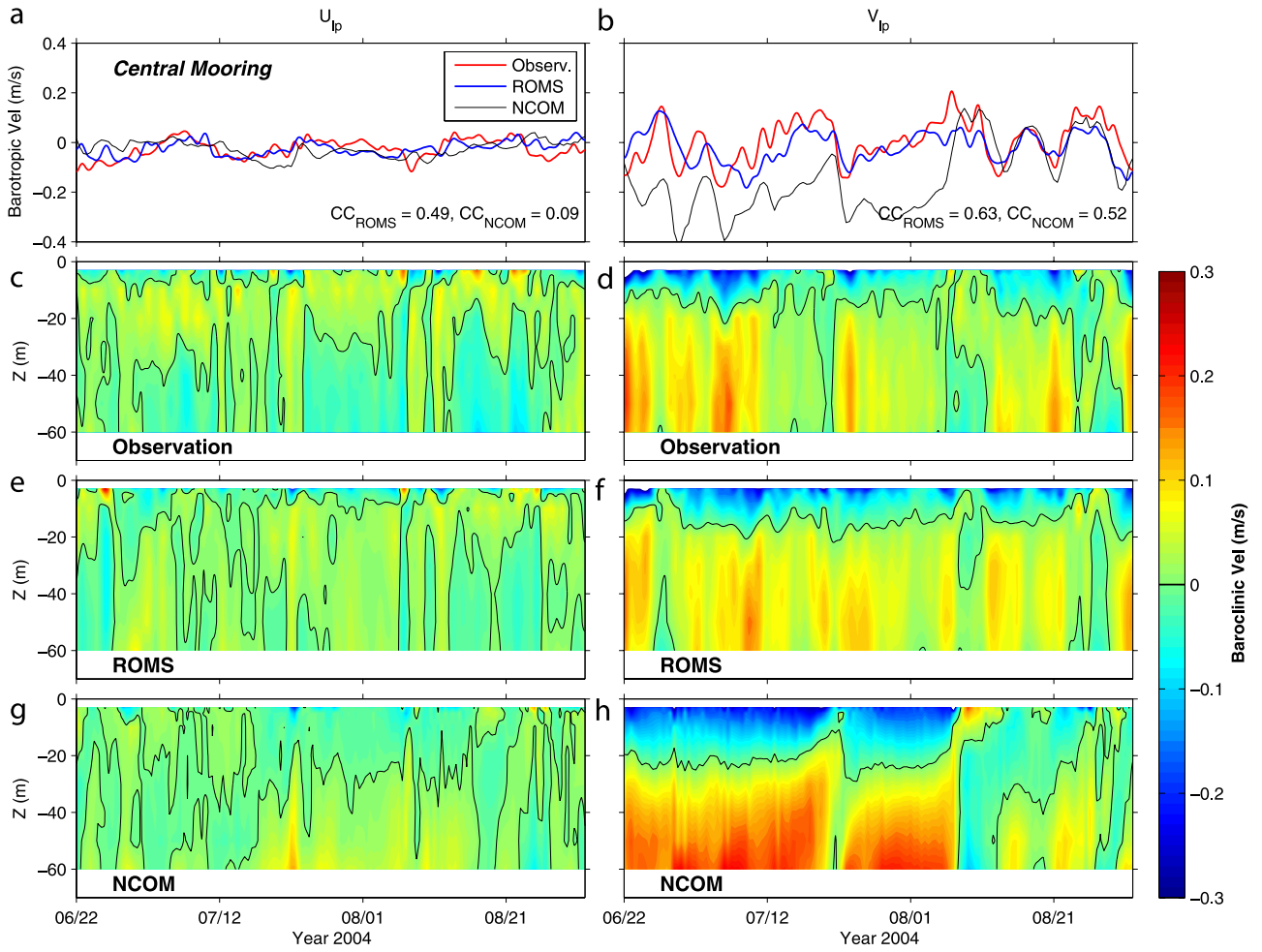

Figure 11. Comparison of 36-h low-pass-filtered east-west (U_{lp}) and north-south (V_{lp}) velocity components at the central mooring between observations and ROMS and NCOM model outputs. The velocity time series are decomposed into (a and b) barotropic and ((c and d) ADCP observations, (e and f) ROMS, and (g and h) NCOM) baroclinic components. Correlation coefficients between the observed and modeled barotropic velocity components (CC_{ROMS} for ROMS, and CC_{NCOM} for NCOM) are also shown in the bottom right corner of Figures 11a and 11b. The correlation coefficients for V_{lp} are above the 95% level of confidence.

components on synoptic timescales. Modeled and observed velocity components at the central mooring are shown in Figure 11. The east-west component (U_{lp}), roughly in the across-shelf direction, is weaker and thus hard to simulate due to the low signal-to-noise ratio, while the north-south component (V_{lp}) is stronger and better agreement (CC = 0.63 versus 0.49) is found between model and observations for both the barotropic and baroclinic components. ROMS is significantly better than NCOM in velocity simulation, e.g., the NCOM simulated southward current is generally stronger by 30 cm s^{-1} in July 2004, which is not seen in ROMS. Among the three moorings, the V_{lp} model-data comparison at the central mooring is the best; V_{lp} is overestimated by 10 cm s^{-1} at the north mooring, and the currents are generally weaker and thus noisier at the south mooring (figures not shown). As noted by *Shulman et al.* [2007], relatively small offsets in the modeled features can lead to an unrealistically poor assessment of model results when comparisons are done with point mooring observations. This is particularly true for the north mooring, which is located just north of the northern edge of the CR plume; a slight difference in the plume position (with the mooring in or out of the plume edge) may result in large differences in the velocity field.

5.3.1.2. Comparison With HF Radar Surface Current Data

[36] Modeled surface currents are interpolated onto the HF radar data points which have an overall geometric dilution of precision (GDOP [*Wells and Beck*, 1987; *Chapman et al.*, 1997]) less than or equal to 4. The velocity time series are low-pass-filtered and EOF analysis is performed (Figure 12). The model-data correspondence between the mode 1 EOF structure and its amplitude time series is very strong (CC = 0.86, R = 0.95). The mode 2 EOF still has a strong correspondence in the spatial structure, though the time series correlation is lower. About 80% of the

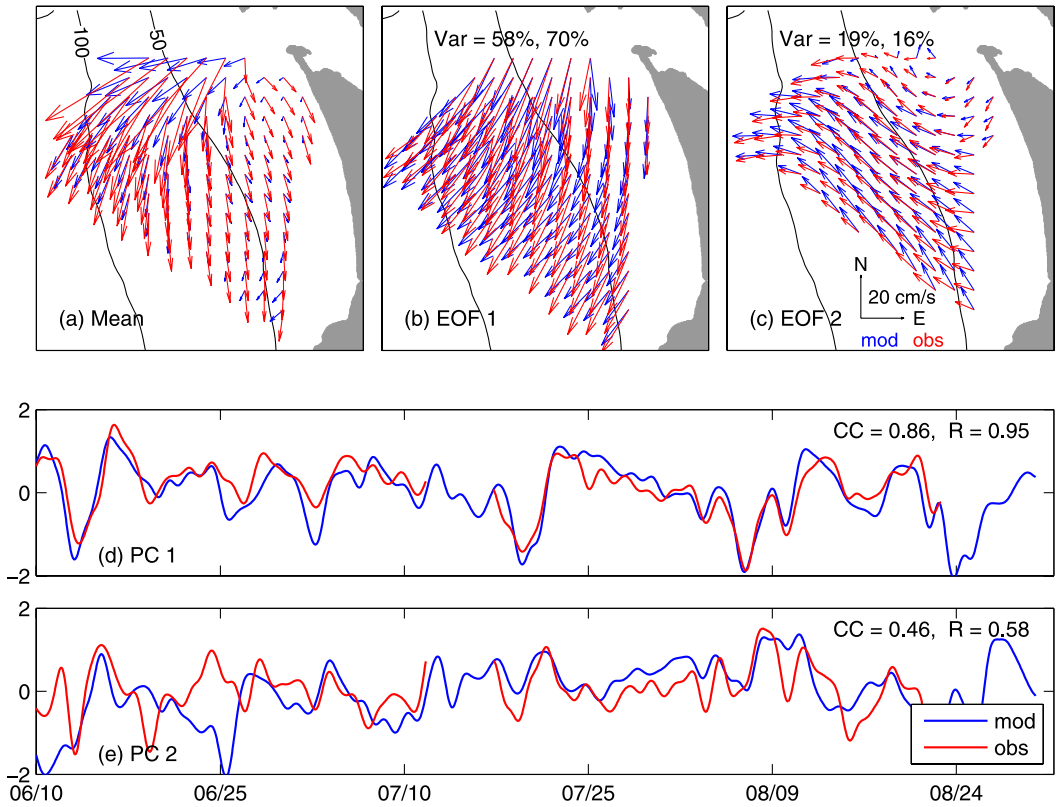

Figure 12. EOF analyses of the 36-h low-pass-filtered surface velocity from the model output and from the HF radar. (a) The mean, (b) the first, and (c) the second mode EOF velocity vectors, and the associated principal component (PC) time series of (d) the first and (e) the second modes. The two percentages in Figures 12b and 12c give the variances accounted for by the specific modes derived from the model and the observations, respectively. Correlation (*CC*) and regression (*R*) coefficient between the principal components of the observation and the model output are also shown in the top right corner of Figures 12d and 12e. The correlation coefficients are above the 95% level of confidence.

variance can be accounted for by the first two modes. The model-observation agreement for mean currents is not as good as those for the EOFs; some angle offsets are seen in weaker current areas and the mean current strength is significantly underpredicted near the CR mouth. On average, the HF radar surface currents were more strongly southward, by 0.109 m s^{-1}, and more strongly westward, by 0.051 m s^{-1}, than the model currents; the average north/south differences grew with distance from the coast, and the average westward differences were strongest offshore from the mouth of the Columbia River. These strong southward currents sweeping offshore are common in the HF radar measurements during southward winds, and often indicate a larger-scale alongshore flow steered offshore at the Columbia River mouth. *RMSE* between model and HF measured EOFs were 0.085 m s^{-1} and 0.087 m s^{-1} for the first and second EOFs, respectively.

[37] Note that the data periods for the EOF analysis are slightly different, i.e., the observations have two 1-week gaps, while the model results are continuous. EOF analysis of the model output with the observation gaps removed results in only a 1% increase in the variance of the first mode, and no visual change to the EOF patterns. Also note that the HF radar current depth may be uniform across the shelf, while the model surface layer (between the first two sigma surfaces) becomes thicker offshore; the top sigma layer thickness changes from 0.3 to 0.7 m between the 50 and 110 m isobaths. The maximum vertical current shear between the first two sigma layers is 7∼12 cm s^{-1} across the 50∼100 m isobaths on synoptic timescales. The accuracy of the forcing winds (from the MM5 regional model) also strongly affects the comparison. The time mean wind stress at the central mooring was 79% higher than observed, with a 15° counterclockwise rotation relative to the observed [*MacCready et al.*, 2009]. However, the mean wind stress is small compared to the significant wind stress events. Other factors that could affect the agreement include, but are not limited to, accuracy of HF radar, model errors due to numerical schemes, initial conditions etc., all of which are beyond the scope of this paper.

[38] Agreement between the model and the HF radar data is much better than with the moored ADCP data, because the radar surface currents are dominated by wind-driven frictional flow, which apparently is well represented by the model. Flows in deep layers are more dominated by ambient flows, which are a function of the open boundary forcing

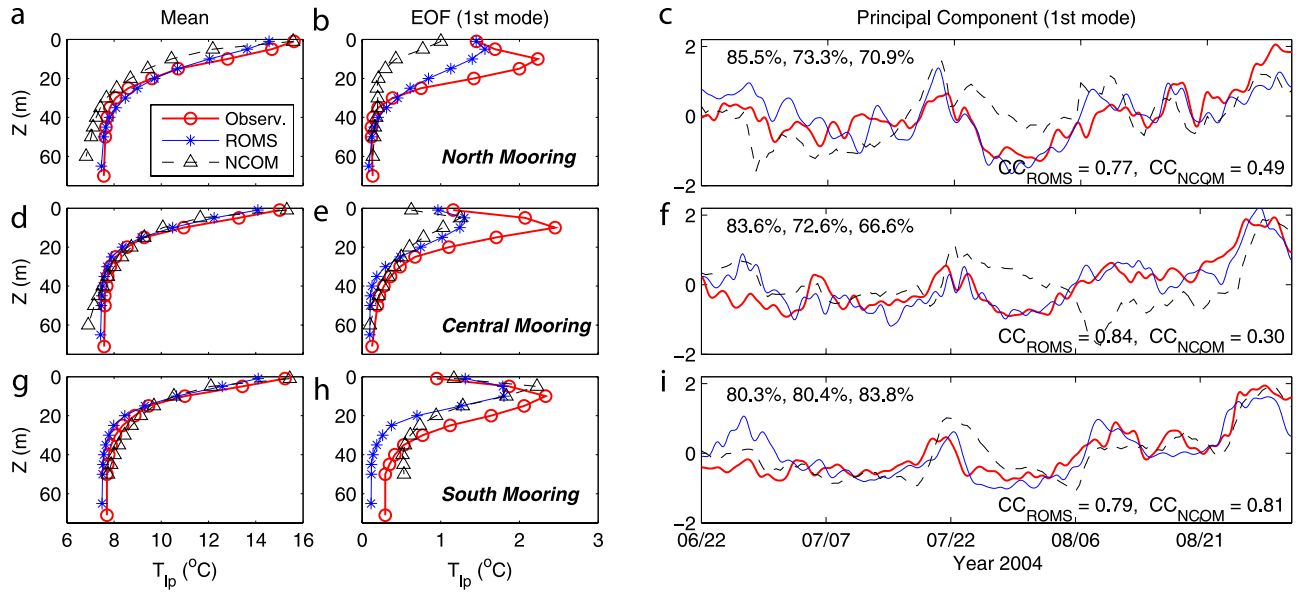

Figure 13. EOF analyses of the 36-h low-pass-filtered temperature (T_{lp}) profiles from the T logger observations, ROMS and NCOM model output at the (a, b, and c) north, (d, e, and f) central and (g, h, and i) south moorings. (left) The mean temperature profiles, (middle) the first mode EOF temperature profiles and (right) the associated principal component time series. The three percentages in the top left corner of Figures 13c, 13f, and 13i give the variances accounted for by the first modes, derived from the observations, ROMS and NCOM output, respectively. Correlation coefficients between the principal components of the observation and the model output (CC_{ROMS} for ROMS, and CC_{NCOM} for NCOM) are also shown in the bottom right corner of Figures 13c, 13f, and 13i. The correlation coefficients are above the 95% confidence level except for CC_{NCOM} at the north and central moorings.

(i.e., quality of the NCOM data) as well as local barotropic responses to wind-forcing. The ambient flow includes the seasonal coastal jet as well as coastal trapped waves.

5.3.1.3. Comparison With Moored *T*

[39] EOF analysis is also performed on the 36-h low-pass-filtered temperature (T_{lp}) profiles from the model and T logger observations for each mooring (Figure 13). The modeled water is somewhat colder than observed with a maximum error of about 1°C near the surface. Similar to tidal timescales (T_{bp}, Figure 9), both modeled and observed mode 1 EOFs show that T_{lp} variability decreases with depth and that the maximum T_{lp} variability is located in a

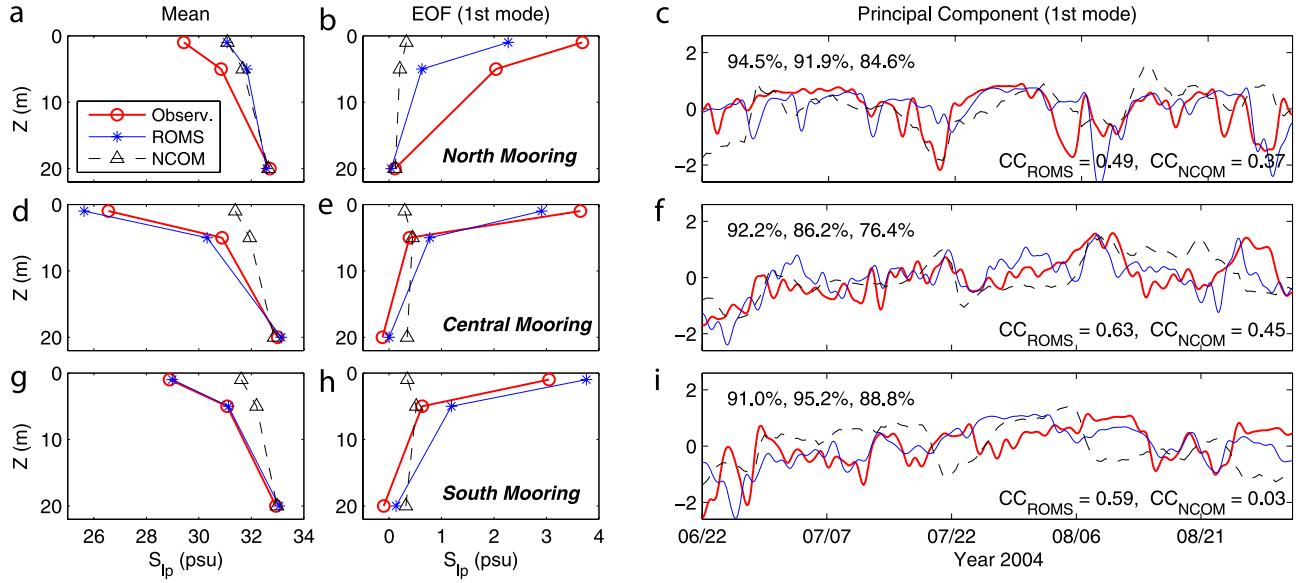

Figure 14. Same as Figure 13 but for the salinity data from Seacat measurements. Only the correlation coefficients CC_{ROMS} at the central and south moorings are above the 95% level of confidence.

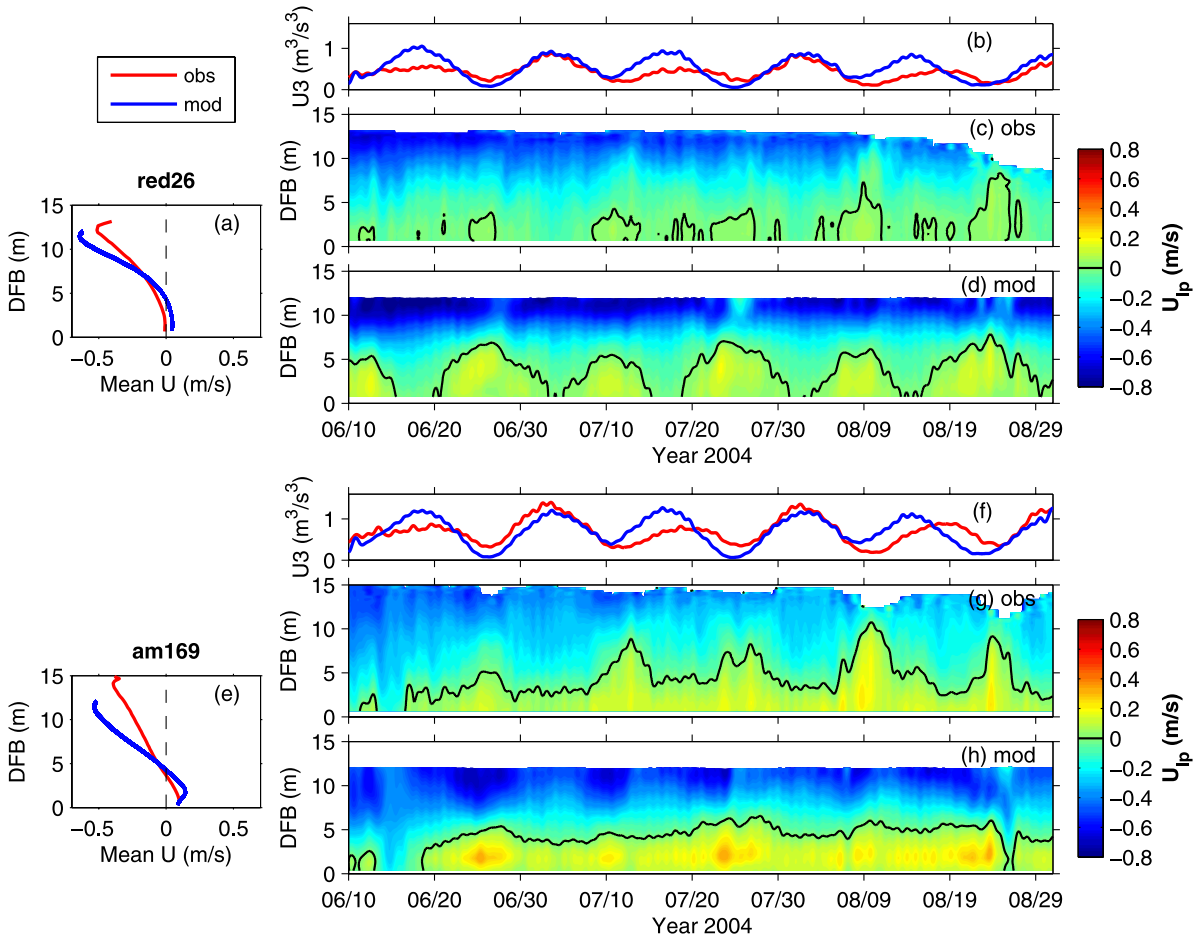

Figure 15. Observed and modeled vertical profiles of (left) the mean and (right) the 36-h low-pass-filtered principal axis current velocity (U) at the two ADP stations: (a–d) red26 and (e–h) am169 in the Columbia River estuary. Both Figures 15b and 15f show the 36-h low-pass-filtered $|\hat{u}|^3$ time series, where $\hat{u}$ is the depth-averaged velocity in the principal axis direction. DFB designates "depth from bottom."

subsurface layer (Figure 13, middle). The magnitude of the maximum T_{lp} variability is twice as large as that of T_{bp}, demonstrating that the T variation is mostly on synoptic timescales in the near field plume. T_{lp} variability is generally weaker than observed, consistent with the fact that the observed thermocline is sharper than the modeled thermocline. Generally, temporal evolution of the modeled and observed T_{lp} profiles are significantly coherent (Figure 13, right).

5.3.1.4. Comparison With Moored *S*

[40] A similar EOF analysis performed on the moored salinity data at 1, 5 and 20 m levels (Figure 14). Both the modeled and the observed temporal mean S vertical profiles show fresher water near the surface, and the best comparison between modeled and observed S occurs at the south mooring. Both the modeled and the observed mode 1 EOFs show the largest S_{lp} variability at the surface (1 m level) associated with the surface trapped CR plume. The near-surface S_{lp} variability is somewhat overestimated at the south mooring by <1, but underestimated at the north and central moorings by less than 2 and 1, respectively. As with the velocity field, the weakest S_{lp} comparison is found at the north mooring, as observed in the mean and first mode EOF vertical profiles and the principal component time series. Note that the north mooring is frequently located near the northern edge of the CR plume; a slight change in the plume position could result in large differences in salinity. Also, the north mooring is located "upstream" of the CR plume and is thus affected by the NCOM open boundary (the mean ROMS and NCOM S profiles overlap at the north mooring) more than at the other moorings. However, ROMS performs better than NCOM in S_{lp} at all three sites.

5.3.2. Comparisons of Velocity, *T* and *S* in the Estuary

[41] Comparisons of currents on synoptic timescales are performed at stations red26 and am169 only, because the velocity record at the other ADP station (tansy) is short, about one month in length. These two stations are located in the south river channel (Figure 1), and the low-frequency currents are directed primarily in the principal axis current direction. The latter is determined as the direction of the maximum variance [*Emery and Thompson*, 1998] using the 36-h low-pass-filtered depth averaged velocity. The princi-

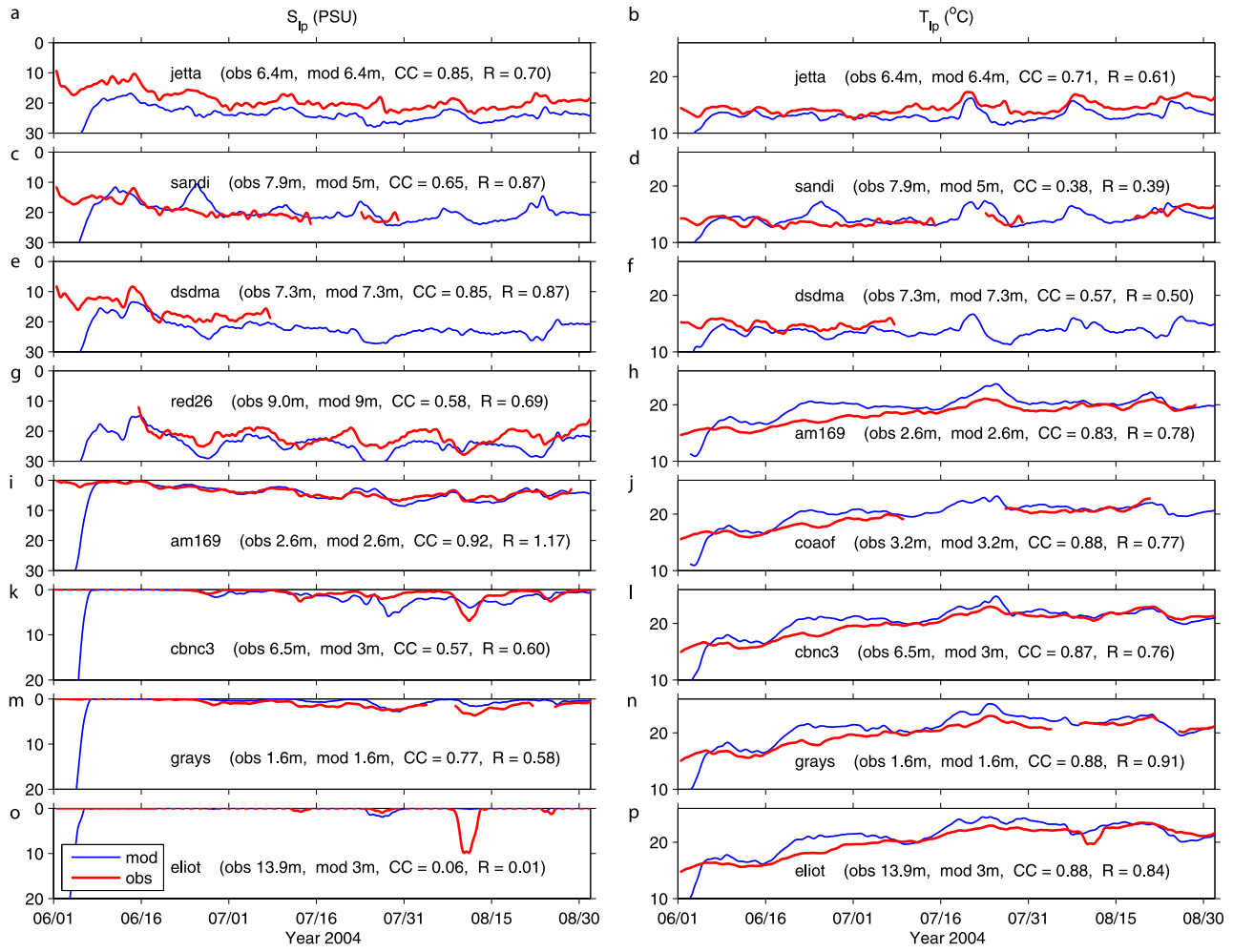

Figure 16. Observed and modeled 36-h low-pass-filtered salinity (S_{lp}) and temperature (T_{lp}) time series in the Columbia River estuary. The station name, sensor depth and data depth extracted from the model, correlation coefficient (*CC*) and regression coefficient (*R*) are shown in each plot. The correlation coefficients are above the 95% confidence level at stations red26, am169, and cbnc3 for S_{lp}, at stations jetta and eliot for T_{lp}.

pal axis current direction is aligned with the river channel, i.e., −46.4° and −42.0° from the east (positive anticlockwise) at station red26, and 5.8° and 6.8° at station am169, respectively, for observed and modeled currents. The principal axis currents (U_{lp}) are obtained by projecting the 36-h low-pass-filtered currents on their principal axis directions. In general, modeled U_{lp} has a two-layer vertical structure [e.g., *Chawla et al.*, 2008] similar to observed, with weaker inflow near the bottom and stronger outflow near the surface (Figure 15). However, modeled vertical shear in U_{lp} is stronger than observed, likely due to the shallower channel depth in the model. Even though less vertical shear is expected in the shallower estuary due to enhanced mixing, the surface outflow is stronger for a given freshwater flux. The temporal evolution of the U_{lp} profile is related to the spring-neap cycle of the flow [e.g., *Jay and Smith*, 1990] at station red26, as indicated by the 36-h low-pass-filtered $|\hat{u}|^3$ time series ($\hat{u}$ is the hourly depth-averaged velocity in the principal axis direction). This is less obvious at station am169, which is located farther up the estuary.

[42] The modeled 36-h low-pass-filtered salinity (S_{lp}) is higher than the observed (by 3∼5) at station jetta, which is located near the river mouth; however, this salinity offset is not seen at the other stations along the CR estuary (Figure 16). The decreasing trend of S_{lp} along the estuary is seen in both model and observations. Modeled T_{lp} is slightly lower than observed at station jetta, but slightly higher toward the inner part of the estuary; in general, modeled T_{lp} compares well with the observed, especially for the seasonal warming in the estuary.

6. Evaluation of Model Skill

6.1. Root-Mean Squared Error

[43] Both *MSE* and *RMSE* are often used to quantify model performance [e.g., *Wilkin*, 2006; *Barth et al.*, 2008]. The latter has the same unit as the variable. As in calculating the *SS* for the sensitivity experiments, the *RMSE* is

Table 2. Root-Mean-Squared Errors of the Model Variables From the Best Model Run[a]

	CTD Sections				Moorings on the Shelf						
	<20 m		>20 m		<20 m			>20 m			
Variables	RP	GH, CM	RP	GH, CM	North	Central	South	North	Central	South	Estuary Stations
U_{lp} (m s^{-1})	–	–	–	–	0.10	0.07	0.07	0.03	0.03	0.02	0.04[b]
V_{lp} (m s^{-1})	–	–	–	–	0.19	0.11	0.11	0.10	0.06	0.05	
T_{lp} (°C)	1.04	1.55	0.26	0.38	1.21	1.11	1.29	0.29	0.23	0.42	1.52
S_{lp}	0.88	0.60	0.22	0.35	1.61	1.17	1.10	–	–	–	2.97
U_{hp} (m s^{-1})			–		0.12	0.11	0.12	0.06	0.05	0.05	0.18[b]
V_{hp} (m s^{-1})			–		0.13	0.11	0.11	0.06	0.05	0.05	
T_{hp} (°C)			–		0.72	0.77	0.65	0.19	0.12	0.14	1.14
S_{hp}			–		0.68	1.08	0.57	–	–	–	3.34

[a]Experiment *o*. *RMSE,* root-mean-squared errors.
[b]The velocity is evaluated for the along channel component.

estimated in surface and deep layers separated by the 20 m level (Table 2). For *U*, *V*, *T* and *S*, larger *RMSE*s are found in the surface than the deep layer as expected since the natural variance is greater there. In the near-field plume, the *RMSE* of U_{lp} is smaller than that of V_{lp} (0.07∼0.10 versus 0.11∼0.19 m s^{-1} in surface layer), with the largest *RMSE* at north mooring. The *RMSE* of T_{lp} is 1.11∼1.29°C and 0.29∼0.42°C in the surface and the deep layers, respectively, with the smallest *RMSE* at central mooring. The *RMSE* of S_{lp} is 1.10∼1.61 in the surface layer (Seacat *S* is not measured in the deep layer), with the largest *RMSE* at north mooring. The *RMSE*s of U_{hp} and V_{hp} are about the same: 0.11∼0.13 and 0.05∼0.06 m s^{-1} in surface and deep layers, respectively. The *RMSE* of T_{hp} is 0.65∼0.72 and 0.12∼0.19°C in the surface and the deep layers, respectively, which is smaller than that of $T_{lp.}$ The *RMSE* of S_{hp} (0.57∼1.08) is also smaller than that of S_{lp}. Along the CTD sections, the *RMSE* of *T* in the surface layer is smaller at Section RP (1.04°C) than at Sections GH and CM (1.55°C) in the far-field plume; while the *RMSE* of *S* is larger at Section RP than at Sections GH and CM (0.88 versus 0.60). The *RMSE* of *T* in the deep layer are about the same as those in the near-field plume. In the CR estuary, the *RMSE* of the along-channel velocity is smaller than those on the shelf on synoptic timescales (0.04 versus 0.07∼0.19 m s^{-1}), but larger for tidal flow (0.18 versus 0.11∼0.13 m s^{-1}). The *RMSE*s of *T* and *S* are generally larger in the estuary than in the near-field plume, again because of the greater natural variance there.

6.2. Willmott Skill

6.2.1. Assessment of ROMS

[44] Since *RMSE* has real units, it is difficult to compare model skill across variables that have different units. For this kind of comparison, the nondimensional *WS* has an advantage. The *WS* is used to quantitatively evaluate the individual physical parameter hindcasts and to assess overall model skill of the best model run (case *o* in Table 1). As in calculating the *SS* for the sensitivity experiments, the model domain is divided into five regions defined by the estuary, near- and far-field plume, and the near surface and deep layers separated at 20 m. The choice of the five regions is motivated by scientific interest rather than by data availability. The main focus of this estuary-plume-shelf cross-scale modeling effort is to simulate the CR plume, a feature trapped in the top few meters of the water column. Thus, in averaging the *WS*, more weight is put on the near surface layer where the plume is trapped. This is equivalent to dividing the shelf water column into a thinner near surface layer and a thicker deep layer, and giving their average *WS* values equal weights in assessing the overall model skill. Similarly, more weights per volume are put on the estuary and the near-field plume than the far-field plume. For each moored time series, the *WS* scores are calculated at every observed depth on synoptic and tidal timescales, respectively. The skills for different physical variables are further averaged in each region and on each

Table 3. Average Willmott Skill Scores for the CTD Sections[a]

	Section RP		Sections GH and CM	
	T	*S*	*T*	*S*
ROMS (<20 m)	0.90	0.80	0.85	0.75
ROMS (>20 m)	0.53	0.48	0.50	0.49
NCOM (<20 m)	0.50	0.44	0.81	0.64
NCOM (>20 m)	0.47	0.53	0.56	0.50

[a]*WS,* Willmott skill.

Table 4. Average Willmott Skill Scores for Synoptic and Tidal Timescales at the North, Central and South Moorings in the Near Field of the Plume

		<20 m			>20 m		
Timescales	Variables	North	Central	South	North	Central	South
Synoptic (ROMS)	U_{lp}	0.48	0.57	0.43	0.46	0.75	0.48
	V_{lp}	0.60	0.66	0.39	0.62	0.68	0.53
	T_{lp} (Tlogger)	0.75	0.78	0.76	0.69	0.70	0.55
	T_{lp} (Seacat)	0.62	0.71	0.75	–	–	–
	S_{lp} (Seacat)	0.75	0.74	0.69	–	–	–
Synoptic (NCOM)	U_{lp}	0.29	0.39	0.37	0.40	0.50	0.54
	V_{lp}	0.56	0.40	0.56	0.66	0.48	0.59
	T_{lp} (Tlogger)	0.49	0.49	0.78	0.31	0.44	0.75
	T_{lp} (Seacat)	0.51	0.54	0.75	–	–	–
	S_{lp} (Seacat)	0.35	0.36	0.54	–	–	–
Tidal (ROMS)	U_{hp}	0.55	0.60	0.38	0.59	0.78	0.49
	V_{hp}	0.45	0.65	0.51	0.47	0.51	0.55
	T_{hp} (Tlogger)	0.41	0.46	0.38	0.31	0.34	0.22
	T_{hp} (Seacat)	0.30	0.49	0.33	–	–	–
	S_{hp} (Seacat)	0.41	0.45	0.43	–	–	–
	η_{hp}	0.97	0.97	0.97	–	–	–

Table 5. Average Willmott Skill Scores in the Estuary

		Stations			
Timescales	Variables	jetta	red26	am169	Average
Synoptic	U_{lp}[a]	–	0.84	0.88	0.86
	T_{lp}	0.60	0.77	0.77	0.71
	S_{lp}	0.57	0.64	0.94	0.72
	η_{lp}[a]	0.70	0.74	0.70	0.71
Tidal	U_{hp}[a]	–	0.95	0.39	0.67
	T_{hp}	0.94	0.90	0.85	0.90
	S_{hp}	0.94	0.90	0.85	0.90
	η_{hp}[a]	0.97	0.97	0.97	0.97

[a]The velocity is evaluated for the along channel component (U), and the sea level (η) is evaluated at tide gauge stations La Push, Toke Point and Astoria, respectively.

timescale. Finally, an average skill is obtained among the five regions and on the two timescales.

[45] Because time series are not available for CTD data, spatial series are used instead. For each CTD transect, the *WS* scores are calculated for T (or S) at each depth taking the spatial points at one depth as a series, and the *WS* values are depth-averaged for surface and deep layers (Table 3). Note that this is a more conservative approach than, e.g., calculating the *WS* for all CTD data below 20 m on a section. For hindcasting T and S, the model has better skill in the surface ($WS \approx 0.8$) than in the deep layer ($WS \approx 0.5$). Note that the water mass in the deep layer is mainly determined by NCOM through open boundary exchange, while that in the surface layer is locally modified by atmospheric forcing and river outflow. Better model skill in near surface T indicates the effectiveness of the bulk formulations [*Fairall et al.*, 1996a, 1996b; *Liu et al.*, 1979] in ROMS and the reliability of the MM5 radiation data. The improvement in S hindcast is mainly because of the presence of the river plume, which is not included in NCOM. For near surface T and S, the model has slightly better skill for the plume axis section (RP) than those sections in the far-field plume regions, and the skill for T is greater than that for S (WS = 0.85∼0.90 versus 0.75∼0.80).

[46] Among the three moorings, the model has greatest skill at the central mooring for the observed U, V, T, and S on both tidal and synoptic timescales (Table 4), because it is most reliably affected by the plume, which is well simulated by our model. The model skill for velocity is slightly better in the deep layer, while that for T is slightly better in the surface layer; this occurs on both synoptic and tidal timescales. On synoptic timescales (Table 4), the model has the most skill for T_{lp} ($WS = 0.69 \sim 0.78$), the second for S_{lp} ($WS = 0.62 \sim 0.75$), and the least for U_{lp} and V_{lp} ($WS = 0.39 \sim 0.66$) in the surface layer. On tidal timescales, the model has the most skill for sea level (η_{hp}) (WS = 0.97); the next best for U_{hp} and V_{hp} (WS = 0.38 ∼ 0.51); the third for T_{hp} (WS = 0.38 ∼ 0.46); and the least skill for S_{hp} (WS = 0.30 ∼ 0.49) in the near-field plume. Our model skills for the moored T are comparable to skills for ROMS on the Southeast New England Shelf [*Wilkin*, 2006], for which WS = 0.46∼0.77.

[47] In the estuary, the *WS* scores are calculated at three stations (jetta, red26 and am169) for along-channel velocity (U), T and S, and at the three tide gauge stations La Push, Toke Point, and Astoria for sea level (Table 5). The model generally has better skill on tidal timescales than on synoptic timescales, reflecting the very strong tidal forcing in the estuary. On synoptic timescales, the model has the most skill for U_{lp} ($WS = 0.86$), the second for T_{lp}, S_{lp} and η_{lp} (WS = 0.71 ∼ 0.72). On tidal timescales, the model has the most skill for η_{hp} ($WS = 0.97$). Note that model skills for T_{lp} and S_{lp} are about the same ($WS \approx 0.7$) in the surface layer of the near-field plume and in the estuary, as is the model skill for η_{hp} ($WS \approx 0.97$). However, the skill for T_{hp} and S_{hp} is much higher in the estuary than on the shelf ($WS \approx 0.9$ versus 0.4). Our model skills for η (0.71∼0.97), S (0.72∼0.90) and U (0.86) are comparable to those of the ROMS applications in the Hudson River estuary [*Warner et al.*, 2005b], in which the skills for η, S and U are 0.85∼0.95, 0.85, 0.68∼0.89, respectively.

[48] For each physical variable, an average skill is taken across the five regions. The average model skills for velocity, T and S are 0.6∼0.7 and 0.5∼0.6, respectively, on synoptic and tidal timescales (Table 6). For each region/layer, an average skill is also obtained by averaging the skills for all the available physical variables. The model has better average skill in the surface layer ($WS \approx 0.6 \sim 0.8$) than in the deep layer ($WS \approx 0.5$) on both timescales. In the surface layer, the average model skill in the near-field plume is lower than those in the far-field plume and in the estuary; this may be due to the high complexity of the plume in the

Table 6. Average Willmott Skill Scores in the Five Regions

		Far Field		Near Field			
Timescales	Variables	<20 m	>20 m	<20 m	>20 m	Estuary	Average
Synoptic	U_{lp}[a]	–	–	0.52 (0.43)	0.59 (0.53)	0.86	0.66 (0.48)
	T_{lp}	0.85 (0.81)	0.50 (0.56)	0.80 (0.54)	0.59 (0.49)	0.71	0.69 (0.60)
	S_{lp}	0.75 (0.64)	0.49 (0.50)	0.75 (0.43)	0.48 (0.53)	0.72	0.64 (0.53)
	η_{lp}	–	–	–	–	0.71	
	Average	0.80	0.50	0.69	0.55	0.75	0.66
Tidal	U_{hp}[a]	–	–	0.53	0.57	0.67	0.59
	T_{hp}	–	–	0.42	0.29	0.90	0.54
	S_{hp}	–	–	0.37	–	0.90	0.64
	η_{hp}	–	–	0.97	–	0.97	0.97
	Average	–	–	0.57	0.43	0.86	0.62
Average		0.80	0.50	0.63	0.49	0.81	0.65

[a]The velocity is evaluated for both the U and V components on the shelf, but for the along channel component (U) only in the estuary. The scores are calculated from ROMS output except those in parentheses calculated from NCOM output.

near-field. Of course, the skills for the far-field plume are assessed only with CTD data (no velocity observations are available). The average model skills for the synoptic and tidal timescales are 0.66 and 0.62, respectively. The overall model skill is 0.65.

6.2.2. Assessment of NCOM

[49] As for ROMS output, the *WS* scores are used to evaluate the NCOM model so that relative skills of the two models can be compared. Since the NCOM does not have the CR or tides, the assessment is based on the same observations on the shelf and for synoptic timescales only. The average *WS* scores for simulating the CTD sections and the moored time series are listed in Tables 3 and 4, respectively.

[50] The NCOM model has about the same skill ($WS \sim 0.5$) as the ROMS model for T and S in the layer deeper than 20 m along the CTD sections (Table 3). Improvement of model skills by ROMS over NCOM (*WS* from $\sim$0.5 to $\sim$0.8) is mainly seen in the upper layer (<20 m) and along the river plume (Section RP), where the CR influence and tidal mixing are most important. This indicates the importance of including river and tides in our modeling system. On the other hand, major improvement is within a limited small area, emphasizing the importance of the outer domain model in providing basic water mass information. Any bias in NCOM water T and S would also be received by the ROMS model.

[51] NCOM's skills for moored T_{lp}, S_{lp} and U_{lp} generally decrease from south to north among the three moorings in the near field-plume area, and the skill for V_{lp} is weakest at the central mooring (Table 4). This may indicate an important influence of the CR plume on coastal trapped waves. When the northward propagating coastal trapped waves reach the latitude of the CR mouth, they can be modified by the strong CR plume. The NCOM model domain is large enough to convey information on coastal trapped waves generated off northern California [*Hickey and Banas*, 2003], but it does not have the river; such modification is of course not captured by the NCOM model, and this may decrease its model skills at the central mooring where the plume influence is most important. Interaction of the plume and coastal trapped waves deserves further study, but is beyond the scope of this paper. The low skills of NCOM at the north mooring may also be due to NCOM bias (e.g., too strong southward currents, colder water) because this site is closer to its northern open boundary at 49°N. Thus, it is easy to understand why major improvement of model skills by ROMS over NCOM is mainly seen at the north and central mooring and in the upper 20 m layer (Tables 4 and 6).

7. Summary and Discussion

[52] Numerical simulations of the CR estuarine-plume-shelf circulation in summer 2004 have been conducted using a ROMS model nested within the NCOM model and driven by realistic atmospheric, tidal and river forcing. Model sensitivity to a number of parameter choices is tested using a relative model skill (*SS*) according to a criterion proposed by *Oke et al.* [2002]. The best model run is validated with a variety of simultaneous observations in summer 2004.

[53] Overall, the best model run has about equal skill at tidal and subtidal properties. Tidal circulation and water properties are best simulated in the estuary, which is strongly forced and damped. In contrast, the worst model performance is for tidal properties on the shelf. Subtidal currents are again best in the estuary. However, subtidal temperature and salinity are best simulated in the surface waters (<20 m) on the shelf – even inside the river plume.

[54] By nesting within the NCOM regional model, some important remote forcings (e.g., coastal trapped waves) are added; on the other hand, some of the NCOM biases are received (e.g., deeper thermocline and stronger southward currents). The influence of the biases may be throughout the entire water column (this calls for more accurate large-scale models). Fortunately, the surface currents are more effectively modified by the surface forcing than the deep currents; thus as long as surface wind stress and heat flux forcing are reliably specified and the NCOM biases are not too large, the surface current simulations are promising. By including tides and river, our model provides improvements over the NCOM model in simulating the T and S distributions in the near-field plume where tidal advection and mixing play an important role. The *RMSE* of T and S are within 1.3°C and 1.6 in the upper 20 m layer of the near-field plume on synoptic timescales, and within 0.8°C and 1.1 on tidal timescales.

[55] Among the three moorings in the near-field plume, the model provides the best simulation at central mooring. This is apparently because it is located within the CR plume most of time, and thus is most reliably affected by the plume which is well simulated by our model. The north mooring is often located near the northern edge of the CR plume. A slight difference in the plume position may result in large difference in S and velocity. Also, the north mooring is located in the ''upstream'' of the CR plume in the California Current which flows southward on the shelf, thus it is most affected by the NCOM bias (too strong southward currents) through the north open boundary. The south mooring is often located on the southern edge of the CR plume; both generally weaker currents and more eddy activities south of the plume decrease model performance there.

[56] Based on the definition of *Willmott* [1981], a comprehensive skill assessment method is proposed to evaluate the estuary-plume-shelf cross-scale modeling system in simulating different physical variables. The model domain is divided into five dynamical regions defined by estuary, near- and far-field plume, near surface and deep layers. Average *WS* is obtained for each region by averaging the *WS* of different physical variables. By defining a thinner near surface layer and a thicker deep layer (separated by the 20 m level on the shelf), a simple average of the *WS* across the five regions results in more skill weight in the near surface layer where the plume is trapped. This weighting scheme is chosen not because of data availability, rather it is motivated by our scientific interests (the plume).

[57] **Acknowledgments.** This work was supported by NSF grants OCE 0239089 (to B. M. Hickey and P. MacCready), OCE 0238021 (to E. P. Dever) and OCE 0237710 (to P. M. Kosro). The NCOM-CCS data were provided by NRL, and the moored observations in the CR estuary were obtained courtesy of Antonio Baptista and the CORIE team of the Oregon Health and Science University. Helpful discussions with S. Springer,

A. Barth and D. Darr on ROMS and its nesting to the larger-domain models are appreciated. This is RISE contribution 31.

References

Banas, N. S., P. MacCready, and B. M. Hickey (2009), The Columbia River plume as cross-shelf exporter and along-shelf barrier, *Cont. Shelf Res.*, *29*, 292–301.

Baptista, A. M. (2002), Environmental observation and forecasting systems, in *Encyclopedia of Physical Science and Technology*, vol. 5, 3rd ed., edited by R. A. Meyer, pp. 565–581, Academic, San Diego, Calif.

Baptista, A. M., et al. (1998), Towards a multi-purpose forecast system for the Columbia River estuary, paper presented at Ocean Community Conference '98, Mar. Technol. Soc., Baltimore, Md.

Baptista, A. M., et al. (1999), Coastal and estuarine forest systems: A multipurpose infrastructure for the Columbia River, *Earth Syst. Monit.*, *9*(3), 1–2, 4–5, 16.

Baptista, A. M., Y. L. Zhang, A. Chawla, M. Zulauf, C. Seaton, E. P. Myers III, J. Kindle, M. Wilkin, M. Burla, and P. J. Turner (2005), A cross-scale model for 3D baroclinic circulation in estuary-plume-shelf systems: II. Application to the Columbia River, *Cont. Shelf Res.*, *25*, 935–972.

Barron, C. N., A. B. Kara, P. J. Martin, R. C. Rhodes, and L. F. Smedstad (2006), Formulation, implementation and examination of vertical coordinate choices in the global Navy Coastal Ocean Model (NCOM), *Ocean Modell.*, *11*(3–4), 347–375.

Barth, A., A. Alvera Azcarate, and R. H. Weisberg (2008), A nested model study of the loop current generated variability and its impact on the West Florida Shelf, *J. Geophys. Res.*, *113*, C05009, doi:10.1029/2007JC004492.

Battisti, D. S., and B. M. Hickey (1984), Applications of remote wind forced coastal trapped wave theory to the Oregon and Washington coasts, *J. Phys. Oceanogr.*, *14*, 887–903.

Bleck, R. (2002), An oceanic general circulation model framed in hybrid isopycnic-cartesian coordinates, *Ocean Modell.*, *4*, 55–88.

Canuto, V. M., A. Howard, Y. Cheng, and M. S. Dubovikov (2001), Ocean turbulence. Part I: One-point closure model—Momentum and heat vertical diffusivities, *J. Phys. Oceanogr.*, *31*, 1413–1426.

Chapman, D. C. (1985), Numerical treatment of cross-shelf open boundaries in a barotropic coastal ocean model, *J. Phys. Oceanogr.*, *15*, 1060–1075.

Chapman, R. D., L. K. Shay, H. C. Graber, J. B. Edson, A. Karachintsev, C. L. Trump, and D. B. Ross (1997), On the accuracy of HF radar surface current measurements: Intercomparisons with ship-based sensors, *J. Geophys. Res.*, *102*(C8), 18,737–18,748.

Chassignet, E. P., et al. (2007), The HYCOM (HYbrid Coordinate Ocean Model) data assimilative system, *J. Mar. Syst.*, *65*, 60–83.

Chawla, A., D. A. Jay, A. M. Baptista, M. Wilkin, and C. Seaton (2008), Seasonal variability and estuary-shelf interactions in circulation dynamics of a river-dominated estuary, *Estuaries Coasts*, *31*(2), 269–288.

Cummings, J. A. (2005), Operational multivariate ocean data assimilation, *Q. J. R. Meteorol. Soc., Part C*, *131*(613), 3583–3604.

Davis, R. E. (1976), Predictability of sea surface temperature and sea level pressure anomalies over the North Pacific Ocean, *J. Phys. Oceanogr.*, *6*, 249–266.

Divins, D. L., and D. Metzger (2002), NGDC Coastal Relief Model, Natl. Geophys. Data Cent., Boulder, Colo. (Available at http://www.ngdc.noaa.gov/mgg/coastal/coastal.html)

Dong, C., and J. McWilliams (2007), A numerical study of island wakes in Southern California Bight, *Cont. Shelf Res.*, *27*, 1233–1248.

Egbert, G. D., and S. Y. Erofeeva (2002), Efficient inverse modeling of barotropic ocean tides, *J. Atmos. Oceanic Technol.*, *19*(2), 183–204.

Emery, W. J., and R. E. Thompson (1998), *Data Analysis Methods in Oceanography*, 634 pp., Pergamon, New York.

Fairall, C. W., E. F. Bradley, J. S. Godfrey, G. A. Wick, J. B. Edson, and G. S. Young (1996a), Cool-skin and warm-layer effects on sea surface temperature, *J. Geophys. Res.*, *101*, 1295–1308.

Fairall, C. W., E. F. Bradley, D. P. Rogers, J. B. Edson, and G. S. Young (1996b), Bulk parameterization of air-sea fluxes for tropical ocean-global atmosphere Coupled-Ocean Atmosphere Response Experiment, *J. Geophys. Res.*, *101*, 3747–3764.

Flather, R. A. (1976), A tidal model of the northwest European continental shelf, *Mem. Soc. R. Sci. Liege, Ser. 6*, *10*, 141–164.

Fox, D. N., W. J. Teague, C. N. Barron, M. R. Carnes, and C. M. Lee (2002), The Modular Ocean Data Assimilation System, *J. Atmos. Oceanic Technol.*, *19*, 240–252.

Galperin, B., L. Kantha, S. Hassid, and A. Rosati (1988), A quasi-equilibrium turbulent energy model for geophysical flows, *J. Atmos. Sci.*, *45*, 55–62.

Garcia-Berdeal, I., B. M. Hickey, and M. Kawase (2002), Influence of wind stress and ambient flow on high discharge river plume, *J. Geophys. Res.*, *107*(C9), 3130, doi:10.1029/2001JC000932.

Haidvogel, D. B., H. G. Arango, K. Hedstrom, A. Beckmann, P. Malanotte-Rizzoli, and A. F. Shchepetkin (2000), Model evaluation experiments in the North Atlantic Basin: Simulations in nonlinear terrain-following coordinates, *Dyn. Atmos. Oceans*, *32*, 239–281.

Haidvogel, D. B., et al. (2008), Ocean forecasting in terrain-following coordinates: Formulation and skill assessment of the Regional Ocean Modeling System, *J. Comput. Phys.*, *227*, 3595–3624.

Halliwell, G. (2004), Evaluation of vertical coordinate and vertical mixing algorithms in the HYbrid Coordinate Ocean Model (HYCOM), *Ocean Modell.*, *7*, 285–322.

He, R., and J. L. Wilkin (2006), Tides on the Southeast New England Shelf: A view from a hybrid data assimilative modeling approach, *J. Geophys. Res.*, *111*, C08002, doi:10.1029/2005JC003254.

Hickey, B. M., and N. S. Banas (2003), Oceanography of the U.S. Pacific Northwest coast and estuaries with application to coastal ecology, *Estuaries*, *26*, 1010–1031.

Hickey, B. M., L. J. Pietrafesa, D. A. Jay, and W. C. Boicourt (1998), The Columbia River plume study: Subtidal variability in the velocity and salinity fields, *J. Geophys. Res.*, *103*, 10,339–10,368.

Hickey, B. M., S. L. Geier, N. B. Kachel, and A. MacFadyen (2005), A bidirectional river plume: The Columbia in summer, *Cont. Shelf Res.*, *25*, 1631–1656.

Hickey, B. M., A. MacFadyen, W. Cochlan, R. Kudela, K. Bruland, and C. Trick (2006), Evolution of chemical, biological, and physical water properties in the northern California Current in 2005: Remote or local wind forcing?, *Geophys. Res. Lett.*, *33*, L22S02, doi:10.1029/2006GL026782.

Jay, D. A., and E. P. Flinchem (1997), Interaction of fluctuating river flow with a barotropic tide: A test of wavelet tidal analysis methods, *J. Geophys. Res.*, *102*, 5705–5720.

Jay, D. A., and E. P. Flinchem (1999), A comparison of methods for analysis of tidal records containing multi-scale non-tidal background energy, *Cont. Shelf Res.*, *19*, 1695–1732.

Jay, D. A., and J. D. Smith (1990), Circulation, density distribution and neap-spring transitions in the Columbia River Estuary, *Prog. Oceanogr.*, *25*, 81–112.

Jerlov, N. G. (1976), *Marine Optics*, *Elsevier Oceanogr. Ser*, vol. 14, 321 pp., Elsevier, Amsterdam.

Kantha, L. H., and C. A. Clayson (1994), An improved mixed layer model for geophysical applications, *J. Geophys. Res.*, *99*, 25,235–25,266.

Kara, A. B., C. N. Barron, P. J. Martin, L. F. Smedstad, and R. C. Rhodes (2005), Validation of interannual simulations from the 1/8° Global Navy Coastal Ocean Model (NCOM), *Ocean Modell.*, *11*, 376–398.

Kindle, J. C., R. Hodur, S. deRada, J. Paduan, L. K. Rosenfeld, and F. P. Chavez (2002), A COAMPS™ reanalysis for the eastern Pacific: Properties of the diurnal sea breeze along the central California coast, *Geophys. Res. Lett.*, *29*(24), 2203, doi:10.1029/2002GL015566.

Large, W. G., J. C. McWilliams, and S. C. Doney (1994), Oceanic vertical mixing: A review and a model with a nonlocal boundary layer parameterization, *Rev. Geophys.*, *32*, 363–403.

Li, M., L. Zhong, and W. C. Boicourt (2005), Simulations of Chesapeake Bay estuary: Sensitivity to turbulence mixing parameterizations and comparison with observations, *J. Geophys. Res.*, *110*, C12004, doi:10.1029/2004JC002585.

Liu, W. T., K. B. Katsaros, and J. A. Bussinger (1979), Bulk parameterization of air–sea exchanges of heat and water vapor including the molecular constraints at the interface, *J. Atmos. Sci.*, *36*, 1722–1735.

Liu, Y., and R. H. Weisberg (2005), Momentum balance diagnoses for the West Florida Shelf, *Cont. Shelf Res.*, *25*, 2054–2074.

Liu, Y., X. S. Liang, and R. H. Weisberg (2007), Rectification of the bias in the wavelet power spectrum, *J. Atmos. Oceanic Technol.*, *24*(12), 2093–2102.

Liu, Y., P. MacCready, and B. M. Hickey (2009), Columbia River plume patterns in summer 2004 as revealed by a hindcast coastal ocean circulation model, *Geophys. Res. Lett.*, *36*, L02601, doi:10.1029/2008GL036447.

Lutjeharms, J. R. E., P. Penven, and C. Roy (2003), Modelling the shear edge eddies of the southern Agulhas Current, *Cont. Shelf Res.*, *23*, 1099–1115.

MacCready, P., R. D. Hetland, and W. R. Geyer (2002), Long-term isohaline salt balance in an estuary, *Cont. Shelf Res.*, *22*, 1591–1601.

MacCready, P., N. S. Banas, B. H. Hickey, E. P. Dever, and Y. Liu (2009), A model study of tide- and wind-induced mixing in the Columbia River estuary and plume, *Cont. Shelf Res.*, *29*, 278–291.

Marchesiello, P., J. C. McWilliams, and A. Shchepetkin (2001), Open boundary conditions for long-term integration of regional oceanic models, *Ocean Modell.*, *3*, 1–20.

Marchesiello, P., J. C. McWilliams, and A. Shchepetkin (2003), Equilibrium structure and dynamics of the California current system, *J. Phys. Oceanogr.*, *33*, 753–783.

Mass, C. F., et al. (2003), Regional environmental prediction over the Pacific Northwest, *Bull. Am. Meteorol. Soc.*, *84*(10), 1353–1366.

Mellor, G. L., and T. Yamada (1982), Development of a turbulence closure model for geophysical fluid problems, *Rev. Geophys.*, *20*, 851–875.

Murphy, A. H. (1992), Climatology, persistence, and their linear combination as standards of reference in skill scores, *Weather Forecast.*, *7*, 692–698.

Nash, J. D., and J. N. Moum (2005), River plumes as a source of large-amplitude internal waves in the coastal ocean, *Nature*, *437*, 400–403.

Oke, P. R., J. S. Allen, R. N. Miller, G. D. Egbert, J. A. Austin, J. A. Barth, T. J. Boyd, P. M. Kosro, and M. D. Levine (2002), A modeling study of the three-dimensional continental shelf circulation off Oregon. Part I: Model-data comparisons, *J. Phys. Oceanogr.*, *32*, 1360–1382.

Orton, P. M., and D. A. Jay (2005), Observations at the tidal plume front of a high-volume river outflow, *Geophys. Res. Lett.*, *32*, L11605, doi:10.1029/2005GL022372.

Pan, J., D. A. Jay, and P. M. Orton (2007), Analyses of internal solitary waves generated by the Columbia River plume with a SAR image, *J. Geophys. Res.*, *112*, C07014, doi:10.1029/2006JC003688.

Pawlowicz, R., B. Beardsley, and S. Lentz (2002), Classical tidal harmonic analysis including error estimates in MATLAB using T_TIDE, *Comput. Geosci.*, *28*, 929–937.

Peliz, A., J. Dubert, D. B. Haidvogel, and B. Le Cann (2003), Generation and unstable evolution of a density-driven Eastern Poleward Current: The Iberian Poleward Current, *J. Geophys. Res.*, *108*(C8), 3268, doi:10.1029/2002JC001443.

Shulman, I., J. Kindle, S. Derada, S. Anderson, B. Penta, and P. Martin (2004), Development of hierarchy of nested models to study the California Current System, in *Estuarine and Coastal Modeling, Proceedings of 8th International Conference on Estuarine and Coastal Modeling*, edited by M. L. Spaulding, pp. 74–88, Am. Soc. of Civ. Eng., Reston, Va.

Shulman, I., J. Kindle, P. Martin, S. deRada, J. Doyle, B. Penta, S. Anderson, F. Chavez, J. Paduan, and S. Ramp (2007), Modeling of upwelling/relaxation events with the Navy Coastal Ocean Model, *J. Geophys. Res.*, *112*, C06023, doi:10.1029/2006JC003946.

Tinis, S. W., R. E. Thomson, C. F. Mass, and B. M. Hickey (2006), Comparison of MM5 and meteorological buoy winds from British Columbia to northern California, *Atmos. Ocean*, *44*, 65–81.

Torrence, C., and G. P. Compo (1998), A practical guide to wavelet analysis, *Bull. Am. Meteorol. Soc.*, *79*, 61–78.

Umlauf, L., and H. Burchard (2003), A generic length-scale equation for geophysical turbulence models, *J. Mar. Res.*, *61*, 235–265.

Warner, J. C., C. R. Sherwood, B. Butman, H. Arango, and R. P. Signell (2005a), Performance of four turbulence closure models implemented using a generic length scale method, *Ocean Modell.*, *8*, 81–113.

Warner, J. C., W. R. Geyer, and J. A. Lerczak (2005b), Numerical modeling of an estuary: A comprehensive skill assessment, *J. Geophys. Res.*, *110*, C05001, doi:10.1029/2004JC002691.

Wells, D., and N. Beck (1987), Guide to GPS positioning, 2nd ed., Can. GPS Associates, Fredericton, N. B., Canada.

Wilkin, J. L. (2006), The summertime heat budget and circulation of southeast New England shelf waters, *J. Phys. Oceanogr.*, *36*, 1997–2011.

Wilkin, J. L., and W. Zhang (2006), Modes of mesoscale sea surface height and temperature variability in the East Australian Current, *J. Geophys. Res.*, *112*, C01013, doi:10.1029/2006JC003590.

Willmott, C. J. (1981), On the validation of models, *Phys. Geogr.*, *2*, 184–194.

Zhang, Y. L., and A. M. Baptista (2008), SELFE: A semi-implicit Eulerian-Lagrangian finite-element model for cross-scale ocean circulation, *Ocean Modell.*, *21*(3–4), 71–96.

Zhang, Y. L., A. M. Baptista, and E. P. Myers (2004), A cross-scale model for 3D baroclinic circulation in estuary-plume-shelf systems: I. Formulation and skill assessment, *Cont. Shelf Res.*, *24*, 2187–2214.

N. S. Banas, B. M. Hickey, and P. MacCready, School of Oceanography, University of Washington, Seattle, WA 98195-5351, USA.

E. P. Dever and P. M. Kosro, College of Oceanic and Atmospheric Sciences, Oregon State University, Corvallis, OR 97331-5503, USA.

Y. Liu, College of Marine Science, University of South Florida, 830 First Street South, St. Petersburg, FL 33701, USA. (yliu18@gmail.com)