1-1-2008

Daily Forecasts of Columbia River Plume Circulation: A Tale of Spring/Summer Cruises

Yinglong J. Zhang
Portland State University

Antonio M. Baptista

Barbara M. Hickey
Portland State University

Byron C. Crump

David A. Jay
Portland State University

See next page for additional authors

Citation Details

Zhang, Yinglong J.; Baptista, Antonio M.; Hickey, Barbara M.; Crump, Byron C.; Jay, David A.; Wilkin, Michael; and Seaton, Charles, "Daily Forecasts of Columbia River Plume Circulation: A Tale of Spring/Summer Cruises" (2008). *Civil and Environmental Engineering Faculty Publications and Presentations*. 23.
https://pdxscholar.library.pdx.edu/cengin_fac/23

Daily forecasts of Columbia River plume circulation: a tale of spring/summer cruises

Yinglong J. Zhang[a1], António M. Baptista[a], Barbara Hickey[b], Byron C. Crump[c], David Jay[d], Michael Wilkin[a] and Charles Seaton[a]

a. Science and Technology Center for Coastal Margin Observation and Prediction,
Oregon Health & Science University,
20000 NW Walker Road, Beaverton, OR 97006, USA

b. School of Oceanography,
University of Washington,
Box 355351, Seattle WA 98195, USA

c. University of Maryland Center for Environmental Studies,
P.O. Box 775, Cambridge, MD 21613, USA

d. Portland State University,
PO Box 751, Portland, OR 97207, USA

Funding sources:
National Science Foundation (OCE-0424602; OCE-0622278), Bonneville Power Administration, and National Oceanic and Atmospheric Administration (AB133F04CN0033)

Submitted to *JGR* on July 11, 2008

[1] Corresponding author; e-mail: yinglong@stccmop.org; phone: (503) 748-1960; fax: (503) 748-1273

Abstract

Semi-operational daily forecasts of circulation from an observatory for the Columbia River estuary-plume-shelf system routinely support oceanographic cruises, by providing 24h-ahead estimates of plume location and structure for planning purposes and for near real-time interpretation of observations. This paper analyzes forecast skill during spring/summer cruises in 2004-2007. Assessment addresses both qualitative descriptions of major plume trends and features and quantitative representation of data from (primarily) vessel-based flow-through, cast and towed systems. Forecasts emerge as robust predictors of plume location and variability, with skill that has grown over time, at least in part due to improvements in model algorithms. When the same version of SELFE is used as the common computational engine, forecast skill is very comparable to the skill of multi-year simulation databases, which are computed retrospectively. As a measure of forecast skill, 55% of the predictions of surface salinities came within 2 km of observations, and directional shifts in response to coastal winds were well predicted. Quantitative skills for other aspects of the plume structure vary. Skill is highest for flow-through (and near-surface TRIAXUS) salinities. Forecasts also capture aspects of the vertical structure of the plume as represented by CTD casts (undulating TRIAXUS). Sub-tidal velocities, as compared against fixed station data, are least well described among examined variables. Overall, circulation forecasts help level the playing field among chief scientists with diverse disciplinary expertise. The chief scientist's expertise on plume physics determines whether forecasts should be used for training and land-assisted planning, or as a sophisticated *en route* planning and interpretation tool. Effective interpretation of vessel observations in context of forecasts requires understanding of physical processes and modeling limitations.

Key words: Forecasting, cross-scale circulation modeling, river plumes, Columbia River plume

1 Introduction

As a major hydrographic feature in the US west coast, the Columbia River (CR) plume transports freshwater as well as nutrients hundreds of miles into the Oregon (OR) and Washington (WA) shelves and beyond. Often reaching northern California to the south and Vancouver Island to the north (Barnes et al. 1972; Hickey et al. 1998), the plume significantly impacts an extensive and complex regional marine and estuarine ecosystem (Hickey et al. 2005; Hickey and Banas 2003; Casillas 1999; Thomas and Weatherbee 2006). Variability at multiple scales, including tidal, seasonal and inter-annual, complicate the understanding of the plume and associated ecosystems. Other than tides, river discharge and shelf winds are dominant drivers of variability.

Ranked second in annual discharge (7000 m^3/s) in the continental United States, the CR is regulated through an extensive system of dams. While dams greatly reduce flow variability relative to pre-development times, the CR still experiences strong seasonal variations. Spring snow melt typically leads to high-discharge freshets, which contrast with late summer minimum flows and with rain-responsive winter discharges. Discharges can exceed 14000 m^3/s during spring freshets or be as low as 2000 m^3/s in late summer. Variation across years is also notable: in recent years, maximum annual discharges have oscillated from 9577 m^3/s in 1997 (with peak freshet of ~15000 m^3/s) to 3965 m^3/s in 2001 (with no noticeable freshet). The volume and surface area of the plume expand in response to increasing river discharge, in ways that also depend on shelf winds (Burla et. al., this issue)

The influence of shelf winds on the CR plume dynamics has long been recognized, with Barnes et al. 1972 describing a bi-seasonal system where the summer plume develops under upwelling-favorable wind and is oriented towards southwest off the Oregon coast, while the winter plume develops northward and attached to the Washington coast in response to Coriolis and downwelling-favorable winds. While this classical bi-seasonal view applies in average, a more complex picture of a "bi-directional" plume has emerged in recent years, in which the summer (winter) plume water can temporarily develop to the north (south) in response to day-scale changes in wind direction (Hickey et al. 2005; Burla et al., this issue). Satellite images suggest that the real picture of the CR plume may be even more complex with multiple patches of the residual plume scattered around Washington and Oregon shelf (Hickey et al. 2005; Thomas and Weatherbee 2006), a view supported by modeling studies (Garvine 1999; Garcia-Berdeal et al. 2002; Baptista et al. 2005; and, in this issue, Burla et al. and Liu et al.) which also identify variability in response to tides and river discharge.

The multi-scale variability of the Columbia River plume poses, in particular, substantial logistical challenges for chief scientists of oceanographic cruises. Numerical forecasts of 3D baroclinic circulation have been used effectively to mitigate these challenges. The forecasts are conducted daily as semi-operational products of the SATURN/CORIE observatory (Baptista 2006.) Other SATURN/CORIE products include real-time observations from multiple stations in the estuary and plume, and near-decade long hindcast simulation databases of 3D circulation.

The SATURN/CORIE forecasts provide both (a) 24h-ahead estimates of plume location and structure for planning purposes, and (b) near real-time retrospective context for interpretation of observations. While a range of chief scientists have empirically endorsed the usefulness of the forecasts, this paper provides a more rigorous skill assessment for a sub-set of RISE and CMOP cruises. Selected were five spring and summer cruises in the 2004-2007 period. The cruises were conducted in different vessels and under chief scientists with different disciplinary backgrounds and with varying degrees of expertise in CR plume dynamics.

After this Introduction, we describe the SATURN/CORIE forecasting system in Section 2. In Section 3 we use a skilled multi-year hindcast simulation database (DB14) to summarize the seasonal and inter-annual of the CR plume (see a more extensive analysis in Burla et al., this issue) and to place cruises in the context of the long-term variation of the plume. Forecast skill assessment, focused on plume salinity, is carried out in Section 4 by quantitative comparison against data collected during four spring and summer cruises in 2005-2007 (flow-through CT, CTD casts, a profiler CTD, and buoys). Because forecast strategies (and versions of the underlying model) have changed over time, we also assess DB14 skill as needed for baseline reference. The Conclusions, Section 5, focus on lessons learned on the usefulness of the forecasts as cruise-supporting tools.

2 Forecasting system in context

2.1 The SATURN/CORIE observatory

One of the earliest end-to-end (data-to-stakeholder) coastal margin observatories, CORIE (Baptista et al. 1998; Baptista et al. 1999; Baptista 2006) was launched in June 1996 with the deployment of a single telemetered conductivity-temperature-depth sensor (CTD). The first sustained forecasts were launched in 1998 using a 2D barotropic model for tidal circulation. While observations and simulations stayed until recently focused on the physical estuary, CORIE has progressively evolved towards the vision (Baptista 2002, Baptista et al. 2008) of observatory-enabled scientific exploration as a potentially transformative tool for oceanography, with implications for regional issues ranging from salmon habitat/passage and hydropower management to navigation improvements and habitat restoration. In 2008, interdisciplinary stations with vertical mobility were deployed, signaling the beginning of the transition of CORIE towards an ambitious end-to-end interdisciplinary observatory, SATURN. As did the original CORIE, the fast-evolving SATURN/CORIE observatory consists of 3 integrated components: a real-time observational network, a modeling system, and a cyber-infrastructure.

The SATURN/CORIE observation network currently consists of ~20 stations in the estuary and plume, which data are regularly used to quality control forecasts and simulation databases, either in the context of forward simulations (Baptista 2006; in this issue: Burla et al.) or of data assimilation (Frolov 2007). Data from the CORIE stations is displayed in real time, albeit with limited quality control, and is then subjected to extensive off-line quality control on a monthly basis. All quality-controlled data and common statistics are available through the web, for either download or visualization, providing the basis for unique insights into the estuarine and plume dynamics (Fain 2001; Chawla et al. 2008; Lohan and Bruland 2006; Bruland et al., this issue). Fig. 1 shows the two SATURN/CORIE stations (OGI01 and OGI02) that are in the plume/shelf region, together with the three RISE stations independently deployed in 2004-2007 (Ed Dever, private communication).

The SATURN/CORIE modeling system currently consists of three sub-systems, all focused on 3D circulation: *simulation databases*, *scenario simulations* and *daily forecasts*. Simulation databases extend continuously from January 1, 1999 through the present (typically lagging calendar time by one year), enabling the analysis of physical variability in the estuary and plume under evolving climate and water regulation. Scenario simulations offer an opportunity to contrast modern Columbia River conditions with reconstructed historical scenarios (e.g., pre-development circa 1880) and future scenarios of change (e.g., associate with climate change or with conditions post a major Cascadia Subduction Zone event). Daily forecasts offer short-term predictions of contemporary conditions, one or a few days ahead, with a level of detail designed to be supportive of *en route* planning and interpretation of oceanographic cruises.

To integrate across extensive observations and simulations, thousands of multi-purpose data and simulation products, and a growing and increasingly diverse user community, the SATURN/CORIE cyber-infrastructure has evolved since inception. It currently consists (Baptista et al. 2008) of a central *data grid* with four external interfaces: the *observation pipeline* for acquiring data from remote sensors, the *forecast factory* for managing the daily operational forecasts and generating the core data products, the *long-term repository* for archiving simulation databases and scenario simulations, and an *integrative web site* for open dissemination of the data. These interfaces, and their earlier renditions, facilitate the continuous evolution of modeling capabilities in general, and forecasting capabilities in particular.

2.2 Circulation forecasts in support of oceanographic cruises

A variety of scientific cruises have been conducted in the Columbia River plume and vicinity since 1998, often in the context of large inter-disciplinary programs sponsored by the National Science Foundation (NSF) and the National Ocean and Atmospheric Administration (NOAA). Of direct relevance for this paper are NSF-supported RISE and CMOP cruises. RISE (River Influences on Shelf Ecosystems) seeks to explore how the river plume modifies biological productivity along the Washington and Oregon continental shelves. CMOP (Science and Technology Center for Coastal Margin Observation and Prediction) explores the use of ocean observatories to advance understanding of the effects of climate and human activity on coastal margins and to educate a diverse workforce.

SATURN/CORIE forecast products (Fig. 1) have been regularly transmitted to on-board computers to aid chief scientist's planning and assess sampling strategies. Conversely, measurements from the ship are transmitted to land computers for near real-time modeling skill assessment. An "ocean appliance" (Howe et al. 2007) brokers the customized management of information flow from and to the vessel. The primary protocol for two-way data transmission is the Ship-to-Shore Wireless Access Protocol (SWAP). The signal is relayed via land-based towers with a transmission range of around 20 nautical miles, depending on the height of the antenna. When this range is exceeded, we also used other means (e.g., satellite phones) to transmit data.

When CORIE daily forecasts were first launched, circa 1998, they required very intensive manual intervention and monitoring and their prediction skill was modest. Today, CORIE include 6 or 7 near-automatic daily forecasts, The supporting cyber-infrastructure (Baptista et al. 2008) ensures automation of most processes, including in the case of forecasts the capture of external forcing from various sources, the launch of simulations based on one of two models (SELFE, Zhang and Baptista 2008; ELCIRC, Zhang et al. 2004), and the generation (by default or on demand) of a suite of post-simulation products such as visual representations and skill assessment metrics.

Forecasts maintained at the time of this writing represent only a snapshot of the forecasting system, as specifics of each forecast and their combination are frequently adjusted to improve skill and to adjust to needs from cruises and other uses. Forecasts (Table 2) may differ in the underlying numerical model, domain and grid resolution, or forcing choices. Multiplicity of forecasts is useful in assessing modeling uncertainty, in differentially refining the representation of specific regions or processes, and in exploring different algorithmic or forcing strategies.

For simplicity, we concentrate in this paper on a single forecast type (‘Development forecast’, created in 2005). We note that the characteristics of simulations within the Development forecast have changed over time, in a continuous search for increased skill. In particular, SELFE versions used in the Development forecast were different for various cruises addressed in this paper: version 1.3g2 was used in June 2005, version 1.3k in August 2005, and version 1.4a in 2006 and 2007 (Table 2). An essential difference is that version 1.4a of SELFE uses an upwind transport algorithm for the salt balance equation, in contrast to (with superior accuracy relative to) the Eulerian-Lagrangian method used in the various variations of version 1.3.

The domain of the Development forecast includes the estuary and the plume, as well as the continental shelf from central CA to Vancouver Island. Fig. 2a shows the extent of the domain and the grid, with additional information on horizontal discretization provided in Fig. 2b and vertical discretization characterized in Table 2. The numerical engine is SELFE (Zhang and Baptista 2008), an unstructured-grid model based on the solution of the shallow-water equation with a finite element/finite volume hybrid method. No data assimilation is used. Each day, two days is simulated, with the full simulation including a one-day hindcast-nowcast and a one-day forecast. External forcing is retrieved and processed automatically. Forcing includes river discharges (Columbia River only); atmospheric conditions from the National Centers for Environmental Prediction (NCEP); and ocean conditions from NCOM (Navy Coastal Ocean Model; Baron et al. 2006). Simulations are launched daily around 11am, to obtain next-day conditions about 12 hours later, i.e., around 11pm of the day of the launch. Results are automatically processed and displayed online together with observational data.

2.3 Simulation databases as reference baseline

Because the specifics (including underlying SELFE version) of the Development forecast change over time, we resort in this paper to results of a SATURN/CORIE simulation database as a consistent baseline against which to interpret changes in forecast skill. There are currently three major simulation databases, self-redundant by design, and differing on the underlying model (SELFE or ELCIRC, Zhang et al. 2004), the domain (estuary-plume-shelf or estuary), and on choices in domain discretization and model parameterization. As baseline for this paper we use DB14, which is most similar to the Development forecast in that it (a) utilizes SELFE, (b) covers the estuary, plume and shelf (from northern California to southern British Columbia), and (c) uses the same horizontal and vertical discretization. A dominant difference between DB14 and the Development forecast is that forcing for DB14 is known retrospectively rather than being itself forecasted. Another important difference is that DB14 consistently uses version 1.4a of SELFE, while the Development forecast uses a range of versions over time (see Table 2).

3 Characteristics of the virtual CR plume

3.1 Contextual variability

To place the spring and summer cruise periods analyzed in this paper in a longer-term context of variability, we summarize here characteristics of the CR plume for a two-year period (2005-2006) as represented by DB14 simulations. A more extensive treatment of the topic of CR plume variability using multiple CORIE simulation databases is provided by Burla et al. (this issue). Other recent studies of seasonal and inter-annual variability of the plume include Thomas and Weatherbee (2006), who analyzed the six years of satellite data 1998-2003 and concluded that inter-annual differences in winter plume are mainly due to wind forcing but summer inter-annual differences are dominated by differences in discharge volume.

Due to its large river discharge, the CR plume is "supercritical" (Fong and Geyer 2002) in the sense that the Coriolis-induced coastal jet cannot carry away the entire river discharge without help from coastal wind and ambient currents. Characteristic of super-critical behavior, the plume tends to develop a circular bulge (Horner-Devine et al. 2006) with discharge-dependent diameter that is then altered by the coastal wind and large-scale ocean currents. An Empirical Orthogonal Function (EOF) analysis of model results (Burla et al 2008) suggests that the first EOF represents the natural tendency of the CR plume to veer right under Coriolis, while the 2nd EOF corresponds to the bi-directional plume; the two EOFs account for approximately 44% and 21% of the total variance respectively.

To characterize the plume variability relative to external forcing, we will use here the plume thickness, area, and centroid location as well as volume (not shown, but used to compute thickness). Plume volume and area are calculated as:

$$\{1\} \qquad P_a = \int \hat{S}_0 dA,$$

where the volume and area integrations are done in a region that encompasses the entire plume[2], and $\hat{S}$ (and the surface value $\hat{S}_0$) is a measurement of local "freshness":

$$\{2\} \qquad \hat{S} = \max(0, 1 - S / S_{ref}),$$

in which S is the salinity, S_{ref} is a cut-off ocean salinity. The plume thickness is then:

$$\{3\} \qquad P_t = P_v / P_a .$$

The location of the plume centroid is computed as:

$$\{4\} \qquad \mathbf{x}_p = \frac{\int \mathbf{x} \hat{S}_0 dA}{\int \hat{S}_0 dA} .$$

Similar but slightly different definitions of these plume metrics are used in Burla et al. (this issue).

Fig. 3 re-enforces the notion that freshwater discharge and coastal wind (especially alongshore wind) are major drivers of the variability of plume characteristics[3]. Discharge varies seasonally in relatively predictable ways. Maximum discharges (often > 10,000m^3/s) occur during spring freshets, typically in late May and early June. During winter months, the discharge tends to peak in correspondence to heavy rainfall. Moderate to low discharges occur between winter and the spring freshet, and minimum discharges occur between summer and fall (~3000m^3/s). Inter-annual variability is also significant, as illustrated by comparing 2005 and 2006, with 2006 having a larger and longer spring freshet, as well as larger discharges in winter.

Except for some periods in summer and early fall, coastal winds are strong (>10m/s) and storms (up to 20m/s) are particularly frequent during winter and spring months. Studies conducted during the summer "quiet" months have found wind reversal to be frequent which leads to the formation of bi-modal plumes (e.g., Hickey et al. 2005). Fig. 3 shows that significant reversals also occur during winter

---

[2] The bounding box we used is 43^{o}N to the south (Newport, OR), 48^{o}N to the north (Juan de Fuca Strait), and 150 km offshore from the mouth of the estuary). Due to the relatively low S_{ref}=28 psu used in Figs. 3-5, extending the bounding box have no effects on the results.

[3] The influence from large-scale coastal currents, and in particular the California Current, is also noteworthy. While we have not analyzed this influence in detail, sensitivity tests suggest that this influence is secondary relative to the influence of river discharge and coastal winds, at least in the near-plume region.

months, when a combination of large discharge and strong and highly variable wind makes the plume dynamics more variable than is often recognized.

The plume area and thickness are in general inversely correlated (Fig. 3; cf. Eq. (3)). The plume area is largest under upwelling favorable wind and high discharge (e.g., spring freshet in May 2005 and May 2006), and smallest during low discharge periods (e.g., August 2005). Notably, the very high discharge during Nov-Dec. 2006 did not result in a large plume because strong downwelling wind pushed the plume tightly against the coast, and the plume was almost non-existent on several days within this period (also similarly in early January 2006). The plume thickened considerably during this period (>30m), and the large variation in thickness around the days when the plume vanishes resulted from the near singularity in Eq. (3). Thickness ranges from <2m during summer months to over 40m during winter storms, and can change by as much as 20m in a 2-5 day period. Peaks in plume thickness generally align with those in downwelling wind, with larger peaks during higher discharges, although there appears to be a phase lag of ~0.2 days, likely due to nonlinear effects (e.g., inertia and turbulent mixing).

Variation of the plume centroid shows complex patterns (Fig. 3). While the plume is always constrained to the west of the river mouth, the north-south excursion of the plume can be very large (e.g., ~110 km from May 20 to June 3 2005, with more common excursions of ~40 km within ~5 days.) The north-south movement is much faster than the east-west movement, most likely due to geographic constraint and the generally stronger along-shore wind component. In general, the plume centroid moves with the coastal wind largely as expected. It moves southwestward during upwelling, and north to northeastward during relaxation and downwelling, with some phase lag. The on- and offshore movement of the centroid is mainly determined by the east-west component of the wind. Thus, depending on the combination of the two components of the wind, northwest or southeast movement of the plume is also common (e.g., around Feb. 15 2006). The movement is also controlled, to a lesser extent, by the discharge; in both 2005 and 2006 we observe quickest movements during the spring freshets.

In addition to the primary driving forces of wind and discharges, the plume also responds to tidal modulation and mixing. This is apparent during summer months, when the other two forces are relatively weak. The spring-neap variation can be clearly seen in the plume thickness, and to a lesser extent, in the centroid location (Fig. 5). Tidal mixing plays an important role in the near-plume processes.

3.2 Cruise conditions

The five cruises discussed in this paper were conducted from May through August, in 2004-2007, and offer a perspective into the very significant variability that occurs within the 'summer plume' regime of Barnes et al. (1972). Fig. 6 allows for a quick overview of the differences in forcing (river discharge and shelf winds) and plume surface signature across cruises, while Fig. 4a-e provides details into the variation in forcing and aggregate plume metrics (area, thickness and location) within each cruise. All plume characteristics are computed from DB14, except for Aug. 2007 where development forecast is used (because DB14 is not yet available for 2007.)

The three cruises in 2005 and 2006 offer a contrast in river discharge regimes. The June 1-18, 2005 cruise took place shortly after the freshet and the discharge was moderate (~7000 m^3/s) and declining (Fig. 4b). By the time the next cruise (Aug. 4-25, 2005) started, the discharge had reached its yearly minimum (~3000 m^3/s; Fig. 4c). The wind was highly variable, especially in June, but was generally weak during the August cruise. The decline of the plume size during this period can be seen clearly in Fig. 4c. The mean surface plume and the standard deviation during the two cruises are shown in Fig. 6, which demonstrates the strong variability of the plume. Due to several strong downwelling events during the first cruise, the residual plume shows a circular bulge outside the mouth and a northward propagating jet that hugs the coastline, although there is a considerable plume south of the mouth as well, reminiscent of the upwelling periods. Note that downwelling favorable wind re-enforces the natural mode of the plume while upwelling favorable wind counters it. The mean plume in the August cruise, however, is a typical southwestward tending plume located offshore. Large anomalies exist in the near-field plume, with larger values in June (Fig. 6). The large anomalies to the north of the mouth even during the primarily upwelling period of August 2005 are a testament to the bi-directional nature of the plume (Hickey et al. 2005; Fig. 6).

The May 10-June 6 cruise in 2006 took place during the freshet and thus exhibited distinctive characteristics from the two 2005 cruises. The wind was again variable with six strong downwelling events alternated by relaxation and upwelling periods. The large river discharge is the most dominating feature of this period. For example, plume thicknesses are considerably different during two downwelling events of comparable magnitude (the line and the left boundary of the box in Fig. 4d), with the large freshwater discharge during the second downwelling event (the box in Fig. 4d) more than doubling the plume thickness. Although the forcing conditions and the plume centroid location at the beginning and end of the cruise were very close to each other, the plume area first dwindled to less than half due to the downwelling events and then increased to more than twice the initial value towards the end of the period. The mean plume resembles that of June 2005, albeit with larger extent, and the standard deviation is all much higher than in June 2005 (Fig. 6). The maximum anomaly reaches as high as 17 psu.

The first half of the Aug. 13-30 2007 CMOP cruise took place under primarily downwelling wind which relaxes to upwelling-favorable wind during the second half of the cruise. Such a wind condition plus the persistent downwelling wind that occurred one week prior to the cruise leads to a maximum thickness around August 21 (Fig. 4e). As compared to the August 2005 cruise, the plume extent during the '07 cruise is larger and with a visible northward component (Fig. 6). The size difference is due to the larger discharge (~4000 versus ~6000 m^3/s). The northward extent is a reflection of 'old plume' water, associated with a strong downwelling event that occurred the week prior to the cruise, and is thus a witness to the temporal memory of the system.

Conditions for the July 2004 RISE cruise were similar to those for August 2005, although with a significant event of wind relaxation around July 20-21 (Fig. 4a), increasing the footprint of the northward extent of the plume relative to August 2005. The plume is also more attached to the coast, likely because of the smaller period of sustained upwelling.

4 Forecast skill assessment

As discussed in Section 2.2, of the several CORIE forecasts available we will concentrate on the skill assessment on the full-domain, SELFE-based *development* forecast, which has run continuously since 2005[4]. Note again that different versions of SELFE were in use at the time of various 2005-2007 cruises (Table 2). Most data presented in this section were collected by vessel-based instruments, and there is some uncertainty regarding the exact depth associated with the data collection. The corresponding model results were extracted at the estimated instrument depth. Each daily forecast

[4] In 2004, only ELCIRC-based forecasts were being run. Results shown in this paper for 2004 are, unless otherwise noted, from DB14.

consists of one-day hindcast simulation and one-day forecast simulation; for consistency, the results shown here were concatenated for the full cruise period from the one-day forecast simulations.

4.1 Tracking exercise

One of the most stringent tests we have conducted for the forecast skill was the use of near real-time model outputs to help find plume fronts in Aug. 21-23 2007, a period of establishing upwelling conditions after a downwelling period with lingering plume memory (Figs. 4e&7). Each day, a pre-defined cruise path, based on the development forecast results, was made available to the chief scientist on the R/V Wecoma. The paths consisted of hourly "waypoints" that were intended to sample across plume fronts multiple times (Fig. 7). A similar exercise was conducted in Aug 29-30.

We measure the model errors as the distance between the ship location and corresponding surface salinity isoline predicted by the forecast (Fig. 7). The average error is 3.6 km. The error is within 2 km more than 55% of the time, and within 5 km more than 70% of the time (Fig. 8b); only less than 10% of the time is the error over 8 km. This level of skill was considered useful by the chief scientist, who reported that the ship was often very near the plume front judging by the different colors of the two types of water. Remaining errors in locating the plume are attributed in part to grid resolution (the grid is typically coarser than 0.5 km in the plume region) and in part due to errors in wind forcing (partially illustrated, at a single offshore station, in Fig. 8a) and in over-estimated thermocline depth in the NCOM ocean forcing (not shown). While SATURN/CORIE simulation databases are not yet available for 2007, our experience with other years indicates that hindcast simulations, which are forced with improved winds and discharges, tend to lead to overall improvement over the forecasts. All uncertainties considered, the forecasts capture plume location in a useful manner, despite the plume’s large variability during the period.

4.2 Flow-through salinity data

The comparison of flow-through data from any of the 2004-2007 cruises against simulations (forecasts or DB14) is illustrative of both the complexity of the CR plume dynamics and the challenges of assessing modeling skill. For instance, for the cruise in July 2004, DB14 results shows overall impressive skill in tracking regime changes, but grossly over-predict (at the vessel location) the plume extent July 17-18 in response to a fast transition from upwelling to downwelling conditions. Note also that errors appear to be substantially different in different regions of the plume. For example, if we divide the plume region into three sub-regions: bulge (defined, for this purpose, as the region within 10 km of the mouth of the CR), north plume (north of the mouth but outside the bulge) and south plume, visual inspection (Fig. 9b; note band identifying the location of the vessel in the bulge or in the south/north plume) suggests that errors are often largest during this particular cruise in the highly varying bulge.

Fig. 14 suggests that forecasts and DB14 skills are comparable “in bulk”. However, using flow-though data from the three cruises in 2005-2006 to assess forecast skill and to compare it against DB14 skill, we will show that absolute and relative performances are highly nuanced. The contextual information for the plume during these cruises has been summarized in Section 3. Due to the range of discharge regimes during these periods (from low flow in Aug. 2005, to moderate flow in June 2005 to high flow freshet in June 2006), the comparisons shown below are representative of many other cruises. In all three cruises, forecast results were sent near real-time to the R/V Wecoma and used in support of *en route* cruise planning by the same chief scientist (Hickey).

Figs. 10a, 11a and 12a show the tracks of the vessel relative to the plume position at the start of each cruise. The modeled plume was very responsive to the wind. Significant variability of plume size within and across cruises was mainly controlled by a combination of the magnitudes of river discharge and wind regime. Plume orientation exhibited different "modes", sometimes even with multiple patches (e.g., June 16, 2005; Fig. 10a). With a more discriminating scale than that used in Figs. 10a, 11a and 12a, we are able to see more patches along the WA and OR coast that were likely remnants of old CR plume (not shown). The bi-directional mode described by Hickey et al. (2005) can be observed on June 7 and 14, 2005 (Fig. 10a), for example.

A broad range of salinities were covered in the flow-through data (Figs. 10b, 11b and 12b). Spatial and temporal variation of salinity was reasonably represented in the forecast. In certain periods, forecast errors appear more responsive to errors in external forcing than in others; for instance, RMS errors for both salinities and wind were largest during the June 2005 cruise, and visual inspection suggests the period June 05-09, 2005 as illustrative of the strong influence in forcing error. However, we caution that it would be over-simplistic to attribute errors only to forcing in general and to wind forcing in particular; for example, the memory of the system significantly complicates the interpretation.

The comparison of forecast skill during the 2005-2006 cruises against DB14 skill is particularly interesting (Figs. 10b, 11b, and 12b; Tables 3, 4; and particularly Table 5). Overall, DB14 far outperforms forecasts during June 2005, but slightly underperforms (especially based on correlation coefficients) forecasts in May 2006. In August 2005, determination of whether DB14 or the forecast has better skill is complex: forecasts slightly outperform DB14 based on RMS errors, yet DB14 is far superior (in all regions of the plume but the south plume) based on correlation factors. Note that correlation factors can be interpreted as indicative of model skill in representing plume variability/trends around the mean state, while RMS errors capture direct mismatches including those between the mean states. In general, RMS forecast errors are substantially smaller in the low-discharge August 2005 cruise than in the large discharge June 2005 and May 2006 cruises – but this is not true for correlations factors, which improve regardless of river discharge with each new cruise. For DB14, RMS errors are lower for lower river discharges, but correlation factors stay at similar levels across all three cruises.

The explanation for the above behavior is not straightforward, and it is confounded by the fact that all metrics are being computed from a relatively scarce dataset with each of the three regions. The case can however be made that a large part of the explanation lies on the evolving versions of SELFE used in the forecasts over time. Indeed, the SELFE versions used in 2005 differ within Development forecasts, and between forecasts and DB14 (Table 2.) Both forecasts and DB14 use SELFE version 1.4a in 2006 (Table 2). Version 1.4a has been independently verified to be the most accurate among the SELFE versions used in this paper (Zhang and Baptista 2008). We thus interpret the results as suggesting, as the first order message, that version 1.4a brought substantially enhanced skill to the forecasts. We also interpret the results as suggesting that although salinity values are better estimated in low discharge conditions, plume variability and trends are captured throughout all discharge regimes at an approximately consistent level. The marginally better performance of the forecast versus DB14 in May 2006 is difficult to interpret, given the confounding uncertainties in the forecasted river discharge, ocean conditions (in particular, thermocline depths from NCOM forecasts) and actual instrument depth (due to ship motion.)

Across all 2004-2007 cruises, forecasts have different skills in different regions (Table 3). Using the previously introduced three sub-regions, RMS errors show least skill for the bulge than for the north and south plume (due to the larger variation of the mean state therein), although correlations show a more consistent cross-region skill. Different factors may affect differentially the skill in each region. For instance, introducing the freshawater discharges from the Grays Harbor and Willapa Bay watersheds in the forecasts for August 2007 (“Experimental forecast”) substantially improves correlation coefficients for the North plume (from 0.82 to 0.91) while leaving correlations for all other regions unaffected (Table 3; Fig. 13b).

4.3 Salinity data from casts

The vertical structure of the plume is determined by a combination of winds, river discharge and tides, as well as by processes that hydrostatic models (like SELFE and ELCIRC) and current grid resolutions are not designed to address (e.g., salt fingering and internal waves, Nash and Moum 2005). CTD casts taken from the R/V Forerunner and the R/V Wecoma provide insights both into the response of plume structure to forcing and into the ability of models to represent observed structure. The R/V Forerunner casts were obtained in day-long cruises, which partially overlap with Wecoma and/or Pt. Sur cruises in 2005-2007.

The casts will be analyzed in three groupings. The first grouping consists of six R/V Forerunner cruises with observations along similar pathways (Figs. 15 & 16). Five of these cruises were in June, with river discharges ranging from 6,100 to 6500 m^3/s in 2005 and 9000 to 10000 m^3/s in 2006; the August cruise has substantially smaller river discharge ~4,500 m^3/s. Taken as an aggregate, and non-withstanding different tide and wind regimes, salinity patterns for these six cruises suggest that increased freshwater discharge leads to increased plume thickness near the mouth of the estuary (e.g., compare Fig. 15a with Figs. 15b-c and Figs. 16a-c), as anticipated from the theory of buoyant jets. In general, the forecasts captured the observed response near surface to river discharge meaningfully, although (likely because the different SELFE versions used in the forecasts for different cruises) not as clearly as DB14 does. In the deeper depths, however, the change of SELFE version from forecast to DB14 seems to have often led to under-prediction of the salinity.

The second grouping (Fig. 17) consists of two 2005 R/V Forerunner cruises, geographically concentrated at the entrance of the estuary or in the near-field coastal jet to the North. Results suggest that DB14 has better skill than the Development forecast, that both types of simulation capture usefully (but not in detail) elements of the variability of the plume. The third grouping (Fig. 18) consists of 4 days of casts from the R/V Wecoma: Aug. 21-23 and 25, 2007. In this case, only forecasts are available. We observe a general degradation of skill with increasing distance from the coast or increasing depth, possibly a reflection of uncertainties in forcing ocean conditions.

4.4 Salinity data from a TRIAXUS fish

Salinity data was also collected by a TRIAXUS, towed from the Pt. Sur during the 2004-2006 cruises (Figs. 19-21). The data were gathered near the surface in 2005 but at 10-40 m below surface in 2004 and 2006.

For the surface data (June 1-4, 2005), both the Development forecast and DB14 have a very strong skill to describe the TRIAXUS data (visually, Fig. 19b; and based on RMS errors and correlation coefficients, see Table 3). While that skill appears at first sight disproportionately high by contrast with the skill reported for the flow-through data (Section 4.2; Table 3), the explanation lies in the much shorter time span of the available TRIAXUS data (after gaps are removed for hourly averaging), which avoids the periods of much larger model errors during the transition from upwelling to downwelling, starting on June 5 (see Fig. 10b for context). Between the Development forecast and DB14, the latter has a higher degree of skill.

Comparisons against undulating data are less satisfactory, consistent with the lesser ability of the model to describe vertical structure versus surface “fingerprint” of the plume. However, performance is better for July 11-27, 2004 (where the observed variability is smaller and is better captured by the model) than it is for June 5-12, 2006 (Fig. 20). Note that TRIAXUS data set has many gaps in July 2004, and, in particular, misses the July 17 event where flow-through data for the surface plume was poorly represented (Section 4.2; Fig. 9b). We do not have significant overlapping flow-through data for June 2006 (Fig. 21). DB14 is more skilled than Development forecast for June 2006, but mostly because of background salinity values (forecast does well in variability); we do not have development forecast for July 2004.

4.5 Sub-tidal velocities from buoys

The previous types of data were collected from vessels, and offer useful insights into plume gradients and geometry. Complementary to these data are in-situ measurements from moorings and buoys, which have the advantage of decoupling space from time, and thus serve as witness to specific events. Long-term time series were gathered at 3 RISE and 2 CORIE offshore buoys (Fig. 1), and are being used in separate papers for a systematic analysis of the modeling skill of SATURN/CORIE simulations (e.g., Burla et al, this issue.) Here we use the velocity information at OGI01 to illustrate the challenges of forecasting sub-tidal velocities.

Velocity variation is particularly difficult to model in this region because the transient plume can significantly alter ambient currents. While tides are an important part of the process, the mean transport is primarily governed by the sub-tidal velocity. Fig. 22a shows comparisons of low-passed velocity, at multiple depths, at OGI01, for June 2006. The forecasts capture near surface velocity reasonably well, but miss important events in the deeper layers. The improvement from the development forecast to

DB14 is marginal at best (Fig. 22b), which suggests that if forcing errors are responsible, these errors are shared between the two types of simulation. A possibility is that errors result from the initial and boundary conditions obtained from NCOM for temperature and salinity.

In all cases, the errors are much larger for the east-west component than the north-south component, especially at deeper depths (Figs. 22b), where the influence of large-scale currents (e.g., California Current system) may be more important. Note that we have not imposed any velocity boundary conditions from NCOM at the ocean boundary and therefore large-scale currents may not be accurately represented in the models. Comparison of velocity variances indicates that the model tends to under-estimate the velocity variability, especially for the east-west component (Fig. 22a), possibly due to low signal-to-noise ratio for this component. The under-estimation of variability may be related to numerical dissipation inherent in the models (Zhang and Baptista 2008) as well as errors in representing turbulent mixing. We have since obtained improved results with newer versions of SELFE.

5. Conclusions and implications

With the foundation of an efficient and robust numerical model (SELFE, Zhang and Baptista 2008), we have been able to forecast the circulation in the CR estuary and plume on a routine basis, with meaningful predictive skill, as a part of the SATURN/CORIE observatory. Given the regional importance of the CR, these forecasts have the potential to be an important supporting tool towards the scientific understanding of the Pacific Northwest coastal margin ecosystem. The forecasts are an open resource, available to the scientific community as outreach of the Science and Technology Center for Coastal Margin Observation and Prediction. By placing those forecasts aboard vessels, we have provided chief scientists in oceanographic cruises anticipatory (24h-ahead) insights into plume location, size, and gradients. The forecasts have already been beneficially applied in a variety of NSF- and NOAA-funded cruises, with chief scientists of very different disciplinary backgrounds, supporting scientific objectives as diverse as characterizing microbial communities, nutrient fluxes, salmon habitats, bird distributions, and harmful algal blooms.

Using a small sub-set of those cruises (specifically, one CMOP and four RISE cruises in 2004-2007) as example, we documented in this paper the level of predictive skill that can be currently achieved for the CR plume. This level of skill has increased over time, as first ELCIRC (not discussed here) and then SELFE matured as cross-scale baroclinic circulation models. For further progress, improvements in external forcing (both ocean and atmospheric circulation at sub-basin scale) are considered essential. Not addressed in this paper are the predictive capabilities for the estuary, which (benefiting from the more constrained and more tidally dominated nature of the estuary) tend to be superior to those for the plume.

Characterizing the level of skill of a forecast in general terms is challenging. Of the skill assessments presented here, the metric (a distance) used to document the plume tracking capabilities for August 2007 is perhaps the most useful and intuitive, albeit least conventional. Getting a vessel located next to the “feature” of interest is clearly a first-level concern for any chief scientist aboard an oceanographic cruise. SATURN/CORIE forecasts have often been able to minimize “guessing”, an accomplishment of practical significance in as broad and complex environment as the CR plume. This is particularly useful during periods of fast transition between wind regimes. High-gradients in the plume are particularly well forecasted: over 55% of the time plume concentrations near gradients were predicted within 2 km during the tracking exercise of August 2007. This effectively opens the door to a wide range of scientific applications, in particular in relation to adaptive sampling schemes using vessels and other mobile platforms such as gliders and autonomous underwater vehicles. Novel fast and model-independent data assimilation schemes (Frolov 2007) will be invaluable in exploring such opportunities.

Integrative metrics (such as RMS error and correlation factors) are also important, although mostly retrospectively. While it might be difficult to relate in near real-time an integrative metric to a small scale feature of interest, these metrics offer the trained investigator fundamental insights into relative skills of different forecasts. They also offer the ability to place forecast skill in perspective relative to algorithmic changes in models, to changes in model parameters, and to the nature and predictive skill of external forcing. This is best done retrospectively, and with the benefit of a reference long-term simulation database. For instance, Table 5 strongly suggests (but does not demonstrate) that SELFE version 1.4a has superior skill relative to version 1.3g2, and (arguably only based on correlation factors, with RMS errors offering a confounding factor) to version 1.3k. Consolidating and translating those insights into practical near real-time guidance to a chief scientist, regardless of disciplinary background, is a worthy future challenge.

The intuition of a chief scientist remains an invaluable tool. For instance, Hickey pointed to limitations in the forecast skill of the plume north of the CR, suggesting that the absence of freshwater from WA rivers could be a factor, before correlation factors (e.g., in Table 3, compare correlation factors for Dev-N and Exp-N) backed numerically that intuition. One of the most important contributions of the forecasts is thus arguably to develop and assist the intuition of the chief scientists.

Graphical representations of the plume available aboard in near real time, in formats comfortable to the particular chief scientist, are a major benefit. As an example of customization, a window into the multiple self-redundant forecasts (differences among which offer insights into uncertainty) can either be useful or distracting/overwhelming to different chief scientists.

Unsurprisingly, graphical representations offer qualitative insights that are highly complementary to what metrics provide. For instance, Fig. 15 offers compelling evidence of the superior ability of DB14 (versus Dev forecasts) to represent near-surface plume structure, and in particular, the effect of river discharge on plume thickness. In this particular case, this visual representation may outweigh quantitative ways to characterize model skill, given general data scarcity and the inability of metrics such as RMS to discriminate errors in the representation of a specific target feature versus errors elsewhere in the water column.

While forecasts are strictly hydrodynamic at this point, there is an opportunity to use empirical relationships to bring forecast products closer to ecological variables of interest. For instance, as suggested by Bruland et al. (this issue), nitrate concentration in the CR plume has a strong empirical relationship with salinity. While plume salinity and nitrate observations are relatively scarce, forecasts and hindcasts of salinities may (within the constraints of model uncertainty) be able to fill in gaps by helping predict or re-construct nitrate fields through empirically derived correlations, for periods and/or regions when/where data is not available. Also pointed out by Bruland et al. (this issue), temperatures and/or salinities in the CR estuary may offer an effective index for shelf upwelling. Research to develop such an index is underway based on CORIE observational data (Charles Seaton, personal communication). Once the index is robustly tested based on observational data, its computation will become an integral by-product of SATURN/CORIE forecasts and hindcasts, again eyeing the potential to fill past gaps and/or to anticipate prevailing conditions.

While this aspect was purposefully de-emphasized in this paper, we note in closing that careful treatment of many aspects influencing numerical accuracy, efficiency and robustness (taken in combination) matters a great deal in representing the cross-scale complexity of the CR plume. The use of unstructured grids enables great flexibility in spatial resolution, but it alone will not guarantee the necessary efficiency, robustness or accuracy. For example, the slow convergence rate of ELCIRC (not shown) means that aggressive refinement is not always possible, and therefore the higher-order schemes used in SELFE are instrumental in modeling important aspects of plume circulation efficiently.

The choice of numerical algorithm must be guided by a judicious balance of considerations, because formal numerical accuracy does not always translate in optimal practical skill for complex systems like the CR. For instance (comparisons not shown), while the performance of ELCIRC in conserving volume and mass is superior to SELFE for the CR system, the overall accuracy and the ability to represent plume (and estuary) features is far superior in SELFE. Similarly, while high-order conservative schemes exist (e.g., discontinuous Galerkin method), efficiency considerations limit their value to forecasting systems such as SATURN/CORIE <u>at the moment</u>. Our experiences with evolving from ELCIRC to SELFE as the default model in the SATURN/CORIE system suggest that forecasting systems should be designed from inception to allow for interchanging computational engines, as new models evolve and computational barriers are removed.

The use of multiple forecasts to study the same feature is an effective way to get insights into errors and uncertainties in complex systems. Inter-comparisons between multiple self-redundant forecasts (and also the use of simulation databases as baseline reference) shed light on different manifestation of errors across different models and modeling options. For example, although SELFE results are generally better than ELCIRC's, the improvement is not uniform (e.g., Burla et al. this issue); one of the confounding factors is error compensation, which is common in modeling this type of systems. Our long-term plan for improving forecast skill calls for categorizing errors and uncertainties from various sources (code, parameterizations, external forcing) with the help of the data assimilation techniques emerging from the work of Frolov (2007).

Acknowledgement

The SATURN/CORIE forecasting system is a team effort, involving an eclectic set of people with complementary expertise. The team is represented in the authorship list of this paper by Joseph Zhang, a lead numerical modeler; Antonio Baptista, who provides overall modeling and scientific oversight for the team; Michael Wilkin, responsible for the field operations; and Charles Seaton, who has been chief scientist aboard multiple M/V Forerunner cruises. Other members of the team who contributed to the results presented include Bill Howe and Paul Turner (cyber-infrastructure) and Michela Burla, Sergey Frolov and Nate Hyde (analyses of simulations). The National Science Foundation (OCE-0622278; OCE-0424602) provided financial support for their contribution to this paper. The Northwest Fisheries Science Center of the National Oceanic and Atmospheric Administration and Bonneville Power Administration provided funding and early incentive for development of forecasting capabilities in support of cruises; thanks are due, in particular, to Ed Casillas and Bill Peterson for their support and scientific collaboration.

Barbara Hickey, a physical oceanographer, has been an early adopter of the CORIE forecasting technology, and her extensive multi-year feedback has substantially improved the quality of our simulations; she was the Wecoma chief scientist in the 2005 and 2006 cruises. Byron Crump, a microbiologist, has since 2007 used CORIE forecasts in support of adaptive sampling of microbial-relevant plume gradients; he was the Wecoma chief scientist in the 2007 cruise. David Jay, a physical oceanographer, was the Pt. Sur chief scientist in the 2005 and 2006 cruises; he provided TRIAXUS data retrospectively for a 'blind test' of the forecast skill. The National Science Foundation (OCE-0424602 for Crump; OCE-0239089 for Hickey; OCE-0622278 for Jay) provided financial support for their contribution to this paper. Any statements, opinions, findings, conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views or policies of the federal sponsors, and no official endorsement should be inferred.

REFERENCES

Baptista, A. M., M. Wilkin, P. Pearson, P. Turner, C. McCandlish, P. Barrett, S. Das, W. Sommerfield, M. Qi, N. Nangia, D. Jay, D. Long, C. Pu, J. Hunt, Z. Yang, E. Myers, J. Darland, and A. Farrenkopf (1998), Towards a Multi-Purpose Forecast System for the Columbia River Estuary, Oceans'98, MTS, Baltimore, MD.

Baptista, A. M., M. Wilkin, P. Pearson, P. J. Turner, C. McCandlish, and P. Barrett (1999), Coastal and Estuarine Forecast Systems: A multi-purpose infrastructure for the Columbia River, Earth System Monitor, vol. 9, pp. 1-2, 4-5, 16.

Baptista, A.M. (2002), Environmental Observation and Forecasting Systems, in Encyclopedia of Physical Science and Technology, vol. 5, R. A. Meyers, Ed.: Academic Press, pp. 565-581.

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.

Baptista, A. M. (2006), *CORIE: the first decade of a coastal-margin collaborative observatory*, Oceans'06, MTS/ IEEE, Boston, MA.

Baptista, A.M., W. Howe, J. Freire, D. Maier and C. Silva (2008), Scientific Exploration in the Era of Ocean Observatories, Computing in Science and Engineering, 10 (3), 53-58

Barnes, C.A., A.C. Duxbury and B.-A. Morse, (1972), Circulation and selected properties of the Columbia River effluent at sea, The Columbia River Estuary and Adjacent Ocean Waters, Pruter, A.T., Alverson, D.L. (Eds.),. University of Washington Press, Seattle, WA, pp. 41–80.

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 Modelling*, 11, 347-375.

Burla, M., A.M. Baptista, Y.L. Zhang, S. Frolov (2008), Seasonal and inter-annual variability of the Columbia River plume: a perspective enabled by multi-year simulation databases, *J. Geophy. Res*. (submitted).

Casillas, E. (1999), *Role of the Columbia River estuary and plume in salmon Productivity*, Ocean conditions and the management of Colombia River salmon, proceedings of a symposium, G. A. Bisbal (Ed.), Portland, OR, Northwest Power Planning Council.

Fain, A.M. V., D.A. Jay, D.J. Wilson, P.M. Orton, and A.M. Baptista (2001), Seasonal and Tidal Monthly Patterns of Particulate Matter Dynamics in the Columbia River Estuary, *Estuaries*, 24(5),770-786.

Fong, D.A. and R.W. Geyer (2002), The alongshore transport of freshwater in a surface-trapped river plume, *J. Phys. Oceano.*, 32, 957-972.

Frolov, S. (2007), Enabling Technologies for Data Assimilation in a Coastal-Margin Observatory, Ph.D. Dissertation, OHSU, Portland, OR.

Garcia-Berdeal, I., B.M. Hickey and M. Kawase (2002), Influence of wind stress and ambient flow on a high discharge river plume, *J. Geophys. Res.*, 107, 3130.

Garvine, R.W. (1999) Penetration of buoyant coastal discharge onto the continental shelf: A numerical model experiment. *J. Phys. Oceano.*, 29, 1892-1909.

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. 10339-13368.

Hickey, B.M. and N.S. Banas (2003), Oceanography of the U.S. Pacific Northwest Coastal Ocean and Estuaries with Application to Coastal Ecology, Estuaries, 26, 1010-1031.

Hickey, B. M., S. Geier, N. Kachel and A. MacFadyen (2005), A bi-directional river plume: the Columbia in summer, *Cont. Shelf Res.*, 25, 1631-1656.

Horner-Devine, A.R., D.A. Fong, S.G. Monismith and T. Maxworthy (2006), Laboratory experiments simulating a coastal river inflow, *J. Fluid Mech*., 555. 203-232.

Howe, B., N. Hagerty, E. Van Matre, D. Maier, A. M. Baptista, C. Seaton, P. Turner (2007), The Ocean Appliance: Complete Platform Provisioning for Low-Cost Data Sharing, IEEE/MTS Oceans 2007, Vancouver, BC.

Lohan, M. C., and K. W. Bruland (2006), Importance of vertical mixing for additional sources of nitrate and iron to surface waters of the Columbia River plume: Implications for biology, *Marine Chemistry*, 98, 260-273.

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.

Thomas, A. C. and R. A. Weatherbee (2006), Satellite-measured temporal variability of the Columbia River plume, *Remote Sensing of Environment*, 100, 167-178.

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.

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

TABLE CAPTIONS

Table 1: Cruise periods

Table 2: Configurations of select forecast and hindcast simulations

Table 3: RMS errors and correlation coefficients for near-surface salinities

Table 4. Aggregate error statistics

Table 5: RMS errors and correlation coefficients for 2005-2006 forecasts, normalized by DB14

FIGURE CAPTIONS

Fig. 1: SATURN/CORIE forecasts are delivered to on board web servers to assist *en route* cruise planning and data interpretation. Data from cruises (here, RISE and CMOP) and fixed observation networks (here, RISE and a sub-set of SATURN/CORIE) are used for near real-time and *a posteriori* forecast skill assessment. Information flow, from sensor data to realistic forcing and simulation products, is enabled by the SATURN/CORIE cyber-infrastructure.

Fig. 2. (a) Model grid used for development forecasts, and (b) grid resolution shown as equivalent radius in km.

Figure 3. Plume characteristics for years (a) 2005 and (b) 2006. The bars indicate the cruise periods. Unless indicated otherwise, all variables are from a SELFE hindcast database (DB14) and 30-hour low-pass filtered using a 4th order Butterworth filter. Tidal range is calculated at the NOAA Astoria tide gauge location. The wind time series is the measured values at NDBC's Columbia River buoy. Discharge is measured at Beaver Army station. The cut-off ocean salinity is S_{ref}=28 psu. Box 1 in (a) corresponds to upwelling periods, during which the plume generally became thinner, and centroid moved offshore and southward. Box 2 indicates that the combination of downwelling wind and freshet lead to dramatic movement of plume in the north-south direction and thickening of the plume. Box 3 shows that downwelling in a low-discharge period only led to a small increase in the thickness. For year 2006 as shown in (b), in contrast to 2005, the river discharge in winter of 2006 was comparable to that during the annual spring freshet. The plume thickness exhibited large oscillation near the beginning and end of the year.

Figure 4. Plume characteristics during (a) July 2004, (b) June 2005, (c) Aug. 2005, (d) May 2006, and (e) Aug. 2007 cruises. See Fig. 3 for explanation of variables used; the only exception is in 2007, where the plume characteristics is calculated using the development forecast (DB14 results not available for this year).

Figure 5. During a relatively constant (low) discharge and weak wind period in the Aug. 2005 cruise (indicated as the bar), the spring-neap cycle can be clearly seen in the plume thickness, and to a lesser extent, the centroid locations (which were also influenced by the shifting wind). See Fig. 3 for explanation of variables used; note that no variables are filtered.

Figure 6. Statistics of plume during five cruises. The wind time series is the measured values at NDBC's Columbia River buoy, and discharges are measured at Beaver Army station (extracted from DB14).

Fig. 7. “Feature tracking” exercise conducted in Aug. 2007. The red dots are the positions of Wecoma, and the black dotted lines are the corresponding salinity isolines for the measured salinity values at the instrument depth (4 m) from the forecast. See Fig. 8 for the error statistics. The forecast skill is generally good; most times the chief scientist was able to locate the plume front (judging by the different water colors) with help from the forecast.

Fig. 8. (a) The salinity values, the model errors in terms of the minimum distance between the Wecoma position and the corresponding salinity isoline, and the wind errors at the 39 way points shown in the previous figure. The correlation coefficient between wind errors and distance errors is only 0.01, again indicating that the plume responds to the time history of the wind. (b) Cumulative percentages as a function of model error distances. Over 55% of the time the error distance is within 2km, and >70% of the time it is within 5km. The model performance is satisfactory, given the errors in wind forcing, and the fact that the model resolution in the region is coarser than 0.5km.

Figure 9. Flow-thru comparisons during July 2004 RISE cruise. (a) Wecoma cruise paths (red lines) and the corresponding surface plume (at 28 and 30 psu; black dotted lines) at the start of each cruise as predicted by the development forecast. (b) Comparison of salinities at instrument depth (approximately 4 m from the surface). Since the dev forecast was not yet deployed for this cruise, DB14 results are used. The salinities have been hourly averaged. The wind is measured at the Columbia River buoy.

Figure 10. Flow-thru comparisons during June 2005 RISE cruise. (a) Wecoma cruise paths (red lines) and the corresponding surface plume (at 28 and 30 psu; black dotted lines) at the start of each cruise as predicted by the development forecast. Partially guided by the forecast, the cruise paths cross the plume many times to study the gradients. (b) Comparison of salinities at instrument depth (approximately 4 m from the surface). The salinities have been hourly averaged. The wind is measured at the Columbia River buoy.

Figure 11. Flow-thru comparisons during Aug. 2005 RISE cruise. (a) Wecoma cruise paths (red lines) and the corresponding surface plume (at 28 and 30 psu; black dotted lines) at the start of each cruise as predicted by the development forecast. (b) Comparison of salinities at instrument depth (approximately 4 m from the surface). The salinities have been hourly averaged. The wind is measured at the Columbia River buoy.

Fig. 12. (a) May 2006 Wecoma cruise paths (red lines) and the corresponding surface plume (at 28 and 30 psu; black dotted lines) at the start of each cruise as predicted by the dev forecast. The wind regime during this period is very variable, resulting in different plume orientations. (b) Comparison of near-surface salinities. The discharges are high in this spring freshet period.

Figure 13. Flow-thru comparisons during Aug. 2007 CMOP cruise. This cruise was concurrent with the feature tracking exercise shown in Fig. 7. (a) Wecoma cruise paths (red lines) and the corresponding surface plume (at 28 and 30 psu; black dotted lines) at the start of each cruise as predicted by the

development forecast. (b) Comparison of salinities at instrument depth (approximately 4 m from the surface). The DB14 results are not available for this cruise; results from two forecasts are shown here. The salinities have been hourly averaged. The wind is measured at the Columbia River buoy.

Fig. 14 Scatter plot of data and model (development forecast and DB14) for near-surface flow-through measurements collected by Wecoma and Pt Sur. Both model and data have been hourly averaged. The comparison for DB14 is done from 2004 to 2006, whereas the comparison for the development forecast is done from 2006 and 2007; same version of SELFE (v1.4a) was used for the comparisons. The RMSE and correlation coefficients can be found in Table 4.

Fig. 15 CTD cast comparison in 2005 cruises by Forerunner, near the CR mouth. The cast locations are shown on top of each column; the first casts are indicated as squares, the last as circles, and the rest as pluses. During these cruises, the M/V Forerunner followed similar patterns. Therefore the influence of increasing discharge on the plume thickness is visible from the casts.

Fig. 16 CTD cast comparison in the 2006 cruise by Forerunner, near the CR mouth. The cast locations are shown on top of each column; the first casts are indicated as squares, the last as circles, and the rest as pluses. During these cruises, the M/V Forerunner followed similar patterns. Therefore the influence of increasing discharge on the plume thickness is visible from the casts.

Fig. 17 Other CTD casts in 2005 by Forerunner. Cast locations are shown on top of each column. The first casts are indicated as squares, the last as circles, and the rest as pluses.

Fig. 18 Other CTD casts in 2007 cruise by Wecoma. Cast locations are shown on top of each column. The first casts are indicated as squares, the last as circles, and the rest as pluses.

Fig. 19 Surface CTD data from Pt Sur in June 2005. (a) Cruise path. (b) Salinity comparison along the path.

Fig. 20 CTD data collected from Pt Sur at *deeper* depths in July 2004. (a) Cruise path. (b) Salinity comparison along the path at the specified depths. The RMS error is 0.84 psu and the correlation coefficient is 0.75.

Fig. 21 CTD data collected from Pt Sur at *deeper* depths in June 2006. (a) Cruise path. (b) Salinity comparison along the path at the specified depths. The RMS error is 1.26 psu (dev) and 0.82 (DB14), and the correlation coefficient is 0.68 (dev) and 0.66 (DB14).

Figure 22. (a) Comparison of residual velocity at ogi01 (Fig. 1), during June 2006 cruise. The depths are relative to free surface. Ellipses for the total velocity (including the tides) are shown next to each depth. (b) Correlation coefficients for u (dashed lines) and v (solid lines).