Article

Evaluation of Structured and Unstructured Models for Application in Operational Ocean Forecasting in Nearshore Waters

Shannon Nudds [1,*], Youyu Lu [1], Simon Higginson [1], Susan P. Haigh [2], Jean-Philippe Paquin [3], Mitchell O'Flaherty-Sproul [2], Stephanne Taylor [1], Hauke Blanken [4], Guillaume Marcotte [5], Gregory C. Smith [3], Natacha B. Bernier [3], Phillip MacAulay [6], Yongsheng Wu [1], Li Zhai [1], Xianmin Hu [1], Jérôme Chanut [7], Michael Dunphy [4], Frédéric Dupont [8], David Greenberg [1], Fraser J. M. Davidson [9] and Fred Page [2]

[1] Ocean and Ecosystem Science Division, Fisheries and Oceans Canada, Bedford Institute of Oceanography, 1 Challenger Drive, Dartmouth, NS B2Y 4A2, Canada; Youyu.Lu@dfo-mpo.gc.ca (Y.L.); Simon.Higginson@dfo-mpo.gc.ca (S.H.); Stephanne.Taylor@dfo-mpo.gc.ca (S.T.); Yongsheng.Wu@dfo-mpo.gc.ca (Y.W.); Li.Zhai@dfo-mpo.gc.ca (L.Z.); Xianmin@ualberta.ca (X.H.); David.Greenberg@dfo-mpo.gc.ca (D.G.)

[2] Coastal Ecosystem Sciences Division, Fisheries and Oceans Canada, St. Andrews Biological Station, 125 Marine Science Drive, St. Andrews, NB E5B 2L9, Canada; Susan.Haigh@dfo-mpo.gc.ca (S.P.H.); Mitchell.O'Flaherty-Sproul@dfo-mpo.gc.ca (M.O.-S.); Fred.Page@dfo-mpo.gc.ca (F.P.)

[3] Meteorological Research Division, Environment and Climate Change Canada, 2121 Transcanada Highway, Dorval, QC H9P 1J3, Canada; Jean-Philippe.Paquin@canada.ca (J.-P.P.); Gregory.Smith2@canada.ca (G.C.S.); Natacha.Bernier@canada.ca (N.B.B.)

[4] Ocean Sciences Division, Fisheries and Ocean Canada, Institute of Ocean Science, 9860 West Saanich Rd., Sidney, BC V8L 4B2, Canada; Hauke.Blanken@dfo-mpo.gc.ca (H.B.); Michael.Dunphy@dfo-mpo.gc.ca (M.D.)

[5] Canadian Centre for Meteorological and Environmental Prediction, National Operation Division, Environmental Emergencies Section, Environment and Climate Change Canada, 2121 Transcanada Highway, Dorval, QC H9P 1J3, Canada; Guillaume.Marcotte@canada.ca

[6] Canadian Hydrographic Service, Fisheries and Ocean Canada, Bedford Institute of Oceanography, 1 Challenger Drive, Dartmouth, NS B2Y 4A2, Canada; Phillip.MacAulay@dfo-mpo.gc.ca

[7] Mercator-Océan, International, 10 rue Hermès, 31520 Ramonville-Saint-Agne, France; jchanut@mercator-ocean.fr

[8] Meteorological Service of Canada, Environment and Climate Change Canada, 2121 Transcanada Highway, Dorval, QC H9P 1J3, Canada; Frederic.Dupont@canada.ca

[9] Ocean and Ecosystem Science Division, Fisheries and Oceans Canada, North Atlantic Fisheries Centre, 80 E. White Hills Rd, St. John's, NL A1A 5J7, Canada; Fraser.Davidson@dfo-mpo.gc.ca

* Correspondence: Shannon.Nudds@dfo-mpo.gc.ca

check for updates

Abstract: The oceanography sub-initiative of Canada's Oceans Protection Plan was tasked to develop high-resolution nearshore ocean models for enhanced marine safety and emergency response, fitting into the multi-scale, multi-level nested operational ocean forecasting systems. For decision making on eventual 24/7 operational support, two ocean models (a structured grid model, NEMO (Nucleus for European Modelling of the Ocean); and an unstructured grid model, FVCOM (Finite Volume Coastal Ocean Model), were evaluated. The evaluation process includes the selection of the study area, the requirements for model setup, and the evaluation metrics. The chosen study area, Saint John Harbour in the Bay of Fundy, features strong tides, significant river runoff and a narrow tidal-river channel. Both models were configured with the same sources of bathymetry and forcing data. FVCOM achieved 50–100 m horizontal resolution in the inner harbour and included wetting/drying. NEMO achieved 100 m resolution in the harbour with a three-level one-way nesting configuration. Statistical metrics showed that one-year simulations with both models achieved comparable accuracies against the observed tidal and non-tidal

water levels and currents, temperature and salinity, and the trajectories of surface drifters, but the computational cost of FVCOM was significantly less than that of NEMO.

1. Introduction

In 2016, Canada launched the CAD 1.5 billion Ocean Protection Plan (OPP) [1] to protect the world's longest coastline and support cleaner, healthier and safer waters. Under the oceanography sub-initiative of the OPP, Fisheries and Oceans Canada (DFO) was tasked to develop high-resolution operational nearshore ocean models for enhanced marine safety and emergency response, specifically electronic navigation and the prediction of oil spill drift trajectory. The nearshore models will eventually fit into the multi-scale, multi-level nested operational ocean-forecasting systems of the Government of Canada, through the collaborative development by Environment and Climate Change Canada (ECCC) and DFO under the Canadian Operational Network of Coupled Environmental Prediction Systems (CONCEPTS) [2] Memorandum of Understanding. The current phase of this OPP sub-initiative focuses on six pilot ports/waterways: Kitimat, Port Metro-Vancouver, Fraser River Port, the St. Lawrence River, Port Hawkesbury, and the Port of Saint John, with plans to extend modelling to other ports in the future. In 2017, during the first year of OPP, a significant effort was made to develop configurations for the Port of Saint John using two widely used, open source, ocean models and evaluate their suitability for OPP applications. The two models, NEMO (Nucleus for European Modelling of the Ocean) [3,4] and FVCOM (Finite Volume Coastal Ocean Model), [5,6] were selected due to their existing applications in Canada.

NEMO is a finite difference model that runs on structured horizontal grids. It was first developed for global and basin-scale applications, and subsequently for coastal applications. Prior to this study, NEMO had not been used for near-shore, port-scale applications. Under CONCEPTS, DFO and ECCC developed a series of operational ocean and sea-ice prediction systems using NEMO, covering the global ocean (with a horizontal resolution of 1/4-degree in longitude/latitude, [7]), regional ocean basins (North Atlantic, Arctic and North Pacific, with 1/12-degree resolution, [8]), and the Great Lakes (with 2 km resolution; [9]). The development of prediction systems for the shelf and coastal oceans off the western and eastern coasts of Canada (with 1/36-degree resolution) is ongoing. These systems run operationally with 24/7 support at ECCC's Canadian Centre for Meteorological and Environmental Prediction (CCMEP).

FVCOM uses finite volume numerics on unstructured horizontal grids. In Canada, and particularly within DFO, there are extensive applications of FVCOM for coastal, near-shore and lake waters. FVCOM is used for simulating both barotropic tides (without including density variations) (e.g., [10,11]), and full baroclinic dynamics (e.g., [12–15]). The use of unstructured grids enables very high horizontal resolution, reaching a couple of metres in nearshore waters in some cases (e.g., [16,17]) while maintaining coarser resolution for offshore areas with larger scale dynamics. Many applications make use of FVCOM's wetting/drying scheme to simulate processes in the intertidal zone (e.g., [17,18]). Under a previous DFO project, FVCOM was used to develop port scale models for five of the six OPP pilot ports, including a baroclinic configuration for the Port of Saint John (without atmospheric forcing).

Several studies have compared different configurations of the same root model, but few have compared fully baroclinic, structured and unstructured models with similar resolution over the same domain. Huang et al. [19] compared FVCOM with a structured grid model, ROMS (the Regional Ocean Modelling System), [20] but the study focused on idealized test cases and used barotropic configurations only. Trotta et al. [21] examined the use of NEMO and the Shallow Water Hydrodynamic Finite Element Model (SHYFEM) [22], in a downscaling context for a relocatable ocean platform for

forecasting, but they did not directly compare the two models. More recently, and most relevant to our study, Biastoch et al. [23] performed a comprehensive comparison of a nested configuration of NEMO and the unstructured Finite Element Sea Ice-Ocean Model (FESOM) [24], but this was for global configurations and focused on large scale ocean circulation.

The models developed for OPP applications will eventually be run with 24/7 operational support, but amongst the Canadian government agencies, the capacity to do so only exists within the ECCC. ECCC currently uses NEMO for operational ocean forecasting. To add the operation of FVCOM to ECCC would entail additional resources (and cost) compared to utilizing the existing NEMO infrastructure. Hence, ECCC required the assessment of the performance of both NEMO and FVCOM for consideration in decision making. To achieve this, DFO and ECCC jointly developed an evaluation process, summarized here, to objectively compare both the predictive accuracy of the key parameters required for OPP applications, and the efficiency in terms of computational cost, between NEMO and FVCOM.

This paper describes the principles and factors that were considered in designing the evaluation process (Section 2), and application to the Port of Saint John including the metrics and sample results of the evaluation (Section 3). The selected examples do not cover the full results of the NEMO/FVCOM evaluation, and do not demonstrate the full strength of either model. The evaluation guided on-going research in the development and improvement of both models. More comprehensive descriptions of the configurations and results of both models are or will be documented elsewhere, e.g., the NEMO configuration by Paquin et al. [25]. Finally, the proposed process and metrics can be generalized and modified to evaluate the configurations developed for other regions, and with models other than NEMO and FVCOM, for research and operations.

2. Factors Considered for Evaluation

The evaluation process includes the selection of the study area, the requirements for the model setup, and the metrics for evaluating the models. The evaluation was designed to objectively (quantitatively) assess the accuracy and efficiency of NEMO and FVCOM for operational forecasting at port scales for parameters of interest to the OPP applications. Here, the term *accuracy* pertains to the models' ability to reproduce the observations, and *efficiency* refers to the computer resources that are required to run the models in an operational context. The objectives of OPP are to improve electronic navigation and to predict oil spill drift, thus the parameters of interest for the evaluation were water level, currents, water temperature and salinity (density), and surface drifter trajectories. As with any experimental design, it is important to consider the ability of the evaluation to detect contrasts in the models and to force the strongest contrast possible. Thus, the evaluation was designed to challenge the models with respect to accuracy and efficiency. Consequently, the evaluation helped gain a better understanding of the strengths and weaknesses of the models, and although the evaluation focused on one port, the results can be reasonably expected to extend to other ports.

Of the six OPP pilot ports, one was selected for the evaluation process. The selection of the study area was based on (1) the regional oceanography being sufficiently complex to include key dynamic processes; (2) the availability of forcing data (from atmosphere, rivers, and open ocean) to drive the port models; and (3) the availability of sufficient observational data to assess the predictive accuracy of the models with respect to the parameters of interest. The parameters of interest at Canadian ports are typically influenced by complicated coastlines, bathymetry, tides, river runoff, the open ocean, and, in some cases, the presence of sea-ice. Due to the urgent timeline of the OPP (i.e., the evaluation was to be completed within the first year of OPP), the explicit inclusion of sea-ice and the simulation of the intertidal zone (with a wetting/drying scheme) were not required, but to ensure that FVCOM was used to its full potential, the use of the wetting/drying scheme was encouraged. At the time of the evaluation, a wetting/drying scheme was not available in NEMO. Forecasting the surface waves at port scales was not considered for the current phase of the OPP.

Various aspects of the model setup were built into the process of evaluation, including spatial resolution, the inclusion of dynamic processes, forcing fields, the duration of the simulation, computational cost, and the variables, frequencies and format of the model output. The required horizontal grid spacing of the models was 100 m or less to resolve the horizontal gradients of currents and the presence of eddies due to nonlinear processes. The required vertical grid spacing was 1 m or less near the surface, to resolve the currents in the upper layer that are important for navigation and oil spill drift. The models included the full baroclinic dynamics to simulate the variations due to surface momentum and buoyancy fluxes, river runoff, and open ocean forcing into the port. The models were subjected to the same forcing, which included atmospheric forcing from the operational weather forecasting system, large-scale oceanic forcing from the operational regional ice-ocean forecasting system, and available tidal and river runoff. At the current stage of the OPP, the port models did not include any data assimilation, partly due to the lack of sufficient real time observational data. A common time frame for the simulation was determined by the available observations. The duration of the simulation was 15 months to allow for a spin-up of the models and the evaluation of a full annual cycle. For the proper evaluation of model efficiency, both models were run on the same high-performance computer facility of the Government of Canada. For operational applications, the required run-time of the models was 0.5 h (or less) for a 48 h simulation. The frequency of the model output was minimally 0.5 h for the proper evaluation of tidal variations.

The metrics for the model data comparison were defined based on the existing expertise of the team, and through expert consulting and literature research. The metrics were defined for the tidal and non-tidal components of sea level and currents, vertical profiles of water temperature and salinity, time series of sea surface temperature (SST) at fixed stations, and the trajectories of surface drift. The quantitative metrics are statistically robust to measure the discrepancy between the model solution and observational data. Prior to performing the statistical comparison, the model results were extracted from the grid node nearest to the locations of the observations. Time series analysis also included a comparison of the energy spectrum. Because the domains of both models cover the area beyond the port, the evaluation was carried out for the "inner harbour" (port) and "outer harbour" separately, with an emphasis put on the inner harbour. In addition to the quantitative evaluation, the models were evaluated qualitatively based on known features of the regional oceanography, e.g., the presence of the river plume, tidal fronts, eddies, etc.

Finally, when possible, the models were compared to existing operational products that are used in Canada. This includes the Scotia-Fundy-Maine WebTide (hereafter referred to as WebTide) solutions for tidal elevation and currents [26], and the regional ice-ocean prediction system (RIOPS) that covers the North Atlantic, Arctic and North Pacific with 1/12-degree resolution [8]. Note that neither WebTide or RIOPS were developed for near-shore operational applications and are not high-resolution port solutions, but were the existing operational products for the area at the time of this evaluation.

3. Results

This section describes how the evaluation process, outlined in Section 2, was applied to the NEMO and FVCOM models for the Port of Saint John. Samples of the evaluation results are based on the model solutions near the end of the evaluation process (December 2017). The evaluation identified some deficiencies or errors in both models, and the subsequent work beyond this evaluation process led to improved model solutions, e.g., documented in Paquin et al. for NEMO [25].

3.1. Regional Oceanography

The Port of Saint John was selected as the study site, in part because of its complicated regional oceanography. The port is in Saint John Harbour (hereafter SJH) located in the Bay of Fundy (Figure 1). The area presents tides among the largest in the world, with a maximum range reaching 16 m in the upper Bay of Fundy [27]. The tidal range in SJH can exceed 8 m at spring tide [28,29]. SJH receives significant freshwater influx from the Saint John River (hereafter SJR) which ranges from about

500 m^3 s^{-1} during summer low-water conditions [29] to a maximum of about 10,000 m^3 s^{-1} during the spring freshet [30]. The combination of tidal flow and river runoff generates sharp density fronts which propagate through the vicinity of the port. Due to the large tidal amplitudes in the SJH and throughout the Bay of Fundy, there are significant inter-tidal zones (tidal flats) in the area. The local geography also leads to a complex flow system in the estuary. As SJR discharges into the harbour, water flows over a shallow sill into the Reversing Falls (see inset in Figure 1), a gorge in which rapids usually reverse in direction. Above the sill, the river expands into a river system that receives fresh water input from a large watershed that includes many lakes, hereafter called the river estuary. The Reversing Falls sill, between the harbour and the river estuary, constrains the upstream propagation of tides into the river estuary [28]. The flow conditions in the Reversing Falls change dramatically over the course of a tidal cycle, with the flow alternating between upstream and downstream. Typically, during the flood tide, the sea level in the harbour rises above the water level in the river estuary and there is a strong flow of saline water into the river system [31]. During the ebb tide, the sea level in the harbour drops to a lower level than in the river estuary and a mix of river and sea water flows back out into the harbour [31]. During freshet conditions, if river levels are sufficiently high, the river discharge dominates the flow through the Reversing Falls, such that there is no reversal of the flow. In the river estuary, a two layer system of fresher surface water and brackish bottom water, with a salinity of up to 20 ppt [28], is established in Long Reach (shown in Figure 1) and is present for most of the year, except during the spring and fall freshets when it may be flushed away [28,32]). In winter, the ice is present on the river estuary, but only thin ice appears near the shore in the harbour [28].
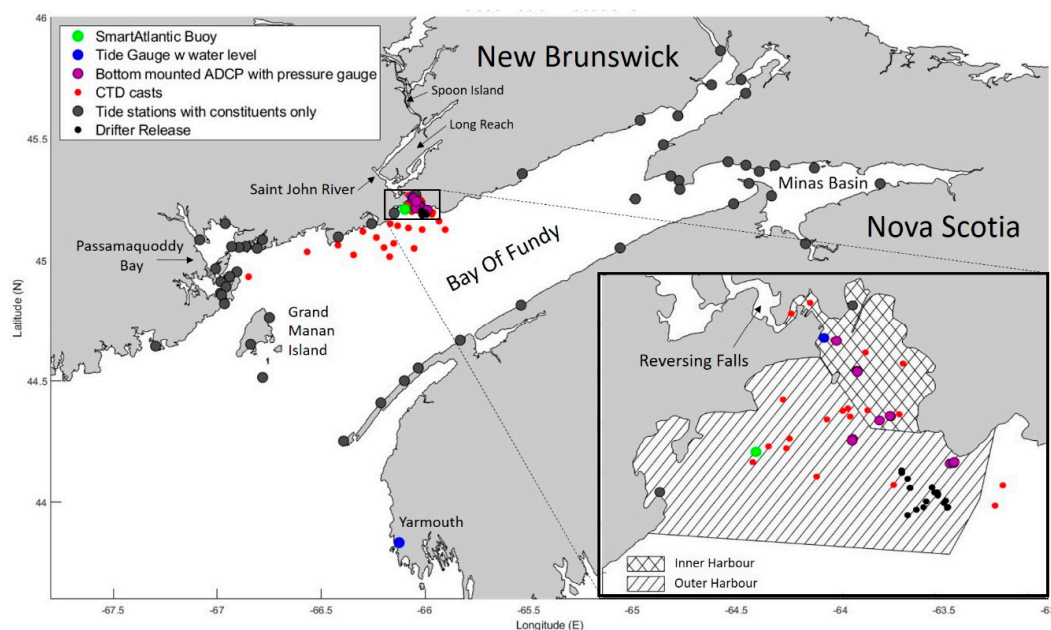
Figure 1. Map of the study area including the locations of the ocean observations, and the boundaries of the inner harbour and outer harbour.

3.2. Available Ocean Observations

A reason for choosing the Port of Saint John as the study area was the quantity and quality of available ocean observations during a one-year window, from 1 May 2015 to 30 April 2016, which was subsequently set as the time period for the model simulation and evaluation. The types and locations of the available observations are shown in Figure 1. The map also divides the study area into three sub-areas: the inner harbour, the outer harbour, and the greater Bay of Fundy. Table 1 provides a summary of the observations, including the variables that were measured by each instrument and the number of observations in each of the three sub-areas. Note that extensive effort was put into collecting and carrying out the quality control on the observational data by various DFO projects.

Table 1. A summary of the observations that were used for the evaluation.

Instrument	Variable	Number of Observations		
		Inner Harbour	Outer Harbour	Greater BoF
SmartAtlantic Buoy	Sea Surface Temperature	0	1	0
Tide Gauge	Water Level	1	0	1
Historical Tide Gauges	Tidal Constituents	1	1	46
Moored ADCP	Currents	6	5	0
	Water Level			
CTD Profiles	Temperature	23	25	32
	Salinity			
Drifters	Surface Drifter Trajectories	0	134	0

All tide gauge and tide station data were obtained from the Canadian Hydrographic Service (CHS). The "constituent only" stations provided the harmonic constants for sets of tidal constituents based on the analysis of past water level time series observations, but did not provide observational time series during the evaluation period. The time series of water levels were available from two real-time tide gauges in the study area: one in the inner SJH and one at Yarmouth. The SJH gauge was critical for evaluating the tidal and non-tidal components of the modelled water level in the inner harbour. The Yarmouth gauge, located near the mouth of the Bay of Fundy, was used to assess the lateral open boundary condition and large-scale performance of the models.

DFO provided observations from 11 moored Acoustic Doppler Current Profilers (ADCPs) that were deployed between August 2015 and February 2016: six in the inner harbour and five in the outer harbour. The ADCPs were either moored a few metres off the bottom or within tripod bottom mounts and recorded the velocity in 1 m bins. To ensure confidence in the observations, a minimum of 75% ping return cut-off was used to define the valid observations. The ADCPs were also set to record the bottom pressure. The bottom pressure data were converted into variations of water level. The "inverse barometric" effect, accounting for the water level variations due to surface atmospheric pressure, was computed from the forcing data and added to the water level derived from the bottom pressure measurement.

Conductivity-temperature-depth (CTD) profiles were also provided by the DFO and were used to evaluate the modelled temperature and salinity fields. There were 80 CTD casts collected at 35 stations in and around SJH between June 2015 and April 2016. Although it is difficult to fully evaluate the large-scale models with point measurements, the locations and timings of the CTD casts made it possible to evaluate the seasonal changes in both the water properties and the depth of the mixed layer at representative locations in the three sub-areas. A SmartAtlantic [33] buoy located just outside the SJH provided the only time series measurement of the sea surface temperature. The buoy data, provided by the Fisheries and Marine Institute of the Memorial University of Newfoundland, were essential for evaluating the seasonal cycle of SST. Monthly mean SST and SSS (sea surface salinity) over the Bay of Fundy were evaluated in a qualitative manner to identify any large-scale discrepancies between the two models.

Lastly, the DFO deployed 134 surface drifters near SJH (in the outer harbour) between July 2015 and January 2016. Four different types of surface drifters were used: Seimac Accurate Surface Tracker barrel-shaped drifters (hereafter referred to as Barrel), MetOcean iSphere spherical drifters (hereafter referred to as Sphere), MetOcean CODE/Davis drifters (hereafter referred to as Davis), and Oceanetic Surface Circulation Tracker drifters (hereafter referred to as Sponge). The drifters provided records of drift trajectories over short periods of time (typically only one day). They were used to evaluate the drifter simulations based on the modelled surface current and the influence of surface winds (described in Section 3.7), and more generally, the models' ability to predict drift, which is one of the primary applications of the models under OPP.

3.3. Setup of NEMO and FVCOM Models

The models are based on NEMO version 3.6 and FVCOM version 3.2.1. Both models solve the 3D equations controlling the variations of ocean currents, water levels, temperature and salinity. The hydrostatic and Boussinesq approximations are applied. The minimum water temperature was set to freezing temperature. This is a simplified approach to account for sea-ice formation in winter, without turning on the sea-ice modules in either model.

The setup of the model domains considered the large-scale ocean forecasting system for providing the lateral open boundary condition (OBC). At the time as the evaluation, the available large-scale model was RIOPS which has a nominal horizontal resolution of 1/12-degree longitude/latitude, about 7.5 km in the study area. The finer coastal ocean prediction system for the east coast of Canada, at the nominal resolution of 1/36-degree, was still in the planning stages.

To take the OBC from RIOPS, FVCOM had the advantage of using a single unstructured-grid with variable horizontal resolution. Based on previous experience, the FVCOM domain was set as shown in Figure 2a, which encompasses the Bay of Fundy and Gulf of Maine, and extends offshore to include the Scotian Shelf and the shelf break. The horizontal cell size ranges from 14 km offshore to 48 m in SJH (Figure 2b). There are 21 geometrically spaced vertical sigma-levels resulting in layer thicknesses ranging from centimeters at the surface, to hundreds of meters at depth in the shelf break area.
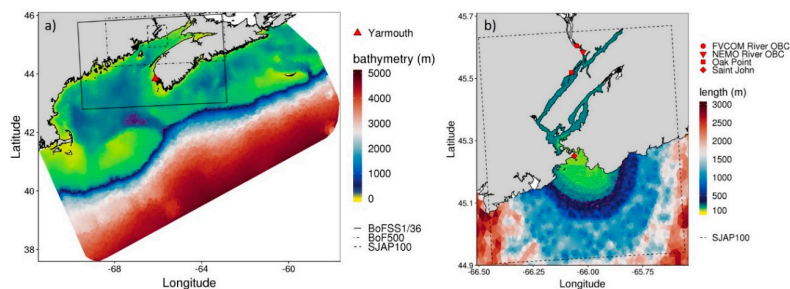
Figure 2. (**a**) Bathymetry over the FVCOM (Finite Volume Coastal Ocean Model) grid and the outlines of the Bay of Fundy and the western Scotian Shelf model component with a nominal horizontal resolution of 1/36-degree (BoFSS1/36) (solid line), the Bay of Fundy model component with an approximate 500 m resolution (BoF500) (dash-dot line), and the inner and outer Saint John Harbour model component with an approximate 100 m resolution (SJAP100) (dashed line) model domains. (**b**) Colormap showing the variable resolution of the FVCOM (as element side length in m) with the outline of the SJAP100 model domain (dashed line). Location of the river boundary, where the open boundary condition (OBC) is applied, is shown for the FVCOM (circle) and NEMO (Nucleus for European Modelling of the Ocean) (inverted triangle). Yarmouth (triangle), Saint John (diamond) and Oak Point (square) are marked for reference.

The structured-grid of NEMO does not have the flexibility to vary the horizontal grid size greatly with a single design. Instead, the approach taken was to develop three configurations (shown in Figure 2a) that are connected with one-way nesting. The grids of the three components are all aligned with the tri-polar ORCA grids created by the DRAKKAR Group [25,34]. The outer component covers the Bay of Fundy and the western Scotian Shelf with a nominal horizontal resolution of 1/36-degree, hence referred to as BoFSS1/36. The intermediate component covers the Bay of Fundy with an approximate 500 m resolution (hereafter BoF500). The inner component covers the inner and outer harbour with an approximate 100 m resolution (hereafter SJAP100). Note that the Saint John river estuary system, including the Reversing Falls, is included in SJAP100, roughly represented in BoF500, but totally excluded in BoFSS1/36. The one-way nesting is achieved by the larger domain component providing the OBC forcing to the next level smaller domain component successively, i.e., from RIOPS to BoFSS1/36, from BoFSS1/36 to BoF500, and from BoF500 to SJAP100. NEMO uses z-levels in the vertical space, with "bottom partial cells" for the accurate representation of the varying bathymetry, and the "variable volume level" scheme [35] to allow for the stretching and compression of the level thickness

according to the changing water levels. Regarding the set-up of vertical levels, BoFSS1/36 uses the same 50-level setup as RIOPS but only has 28 active levels because its domain does not reach the deep ocean; BoF500 and SJAP100 have the same setup with 41 and 35 active levels, respectively, due to the difference in the maximum water depth between the two domains. The finest vertical resolution is near the surface, 1 m for all three components.

A high-resolution coastline dataset provided by the CHS was used to define the coastline represented in FVCOM and each component of NEMO. The model bathymetry within the Bay of Fundy and part of the Gulf of Maine was interpolated from a high-resolution dataset constructed from CHS and the Ocean Mapping Group, University of New Brunswick (OMG) data sources. Over the area not covered by the CHS/OMG data, NEMO used the global high-resolution SRTM30 bathymetry product [36] and FVCOM used bathymetry from the Scotia-Fundy-Maine grid of WebTide. As a wetting/drying scheme is unavailable in NEMO version 3.6, several modifications were made to the coastline and bathymetry, mainly in the upper Bay of Fundy and in the SJH which has large intertidal areas. Part of the intertidal zone was excluded, and part had the water depth increased to avoid drying and to maintain numerical stability. The modifications ensured the good representation of the M2 tides in and near the SJH [25].

Except for the river estuary system, which was initialized with the same data as the NEMO BoF500 simulation, the FVCOM model was initialized with the 3D temperature and salinity from the RIOPS solution on 1 February 2015, and was ramped up from rest (zero sea level and velocity) over a period of 36 h. The outer component of NEMO (BoFSS1/36) was initialized with the RIOPS solution on 1 February 2015, including 3D temperature and salinity, sea levels, and currents. BoF500 was initialized with the BoFSS1/36 solution on 1 April 2015, but with the temperature and salinity in the river estuary system taken from the solution of a separate 1-year simulation. Finally, SJAP100 is initialized on 10 April 2015, from the BoF500 solution. The simulations prior to 1 May 2015 (the start of the evaluation duration), are treated as the spin-up of the models. For FVCOM, the model spin-up occurs between 1 February and 30 April 2015. This spin-up period is considered sufficient for the winter conditions of the study area [25].

Input data for the lateral OBC included tidal and non-tidal components. For the FVCOM model and the BoFSS1/36 of NEMO, the non-tidal water levels for the OBC, at hourly intervals, were obtained from the de-tided RIOPS solution. The NEMO OBC also included non-tidal depth-averaged velocity from RIOPS. De-tiding is necessary because the RIOPS tidal solution contains significant errors in the Bay of Fundy. The input for OBC also included the daily averaged 3D temperature and salinity (without de-tiding) from RIOPS. The tidal water levels (and depth-averaged velocity for NEMO only) were computed from the harmonics of five major tidal constituents (M2, N2, S2, K1 and O1) from the WebTide solution [26]. Other tidal constituents will be included in the future development and operational application of the models. The BoFSS1/36 obtained a sufficiently accurate tidal solution near the open boundary of BoF500, and hence, BoF500 used both the tidal and non-tidal forcing for OBC from the solution of the BoFSS1/36, at 0.5 h intervals. Consequently, the SJAP100 took the full OBC forcing from the BoF500, also at 0.5 h intervals. Regarding the setup of the OBC, NEMO applied a "radiation" scheme for the barotropic (depth-averaged) current normal to the open boundary [37], and a "flow relaxation" scheme for temperature, salinity and the baroclinic current over a "relaxation zone" [38]. FVCOM enforced the water level, temperature and salinity at the open boundary nodes. A sponge layer was implemented at the open boundary with a radius of 25 km. Within the sponge layer, the longitudinal and latitudinal velocities, u and v, were reduced by a factor of u/(1+ Cspg u2) and v/(1+ Cspg v2). The friction coefficient, Cspg, was specified to be 0.002 at the open boundary nodes and linearly decreases to zero over the radius of the sponge layer.

The Saint John River runoff was included in FVCOM, and the BoF500 and SJAP100 components of NEMO. For both the NEMO and FVCOM configurations, the open boundary of the river was placed upstream of the Long Reach and below the Spoon Island (see Figure 2). This location is close to the upstream extent of the salt wedge intrusion and limits the inclusion of complicated river morphology.

The river forcing was not included in the BoFSS1/36, but this did not cause a significant error in the BoFSS1/36 solution at the lateral open boundary of the BoF500. Although both models have the ability to include river discharge, suitable input data were not available, and thus the entry of the SJR was treated as an open boundary forced with the observed water levels at the Oak Point station in the river estuary system (Figure 2). A constant value of 0.7 m was subtracted from the observed Oak Point water levels prior to using them for OBC forcing. This constant is the estimated difference in the reference levels between the Oak Point station and the model [21].

At the sea surface, the models were forced using the forecast fields from the High-Resolution Deterministic Prediction System of the CCMEP, with a horizontal resolution of 2.5 km [39]. The forecast variables included hourly winds at 10 m height, air temperature and specific humidity at 2 m height, sea-level air pressure, precipitation, and surface incoming longwave and shortwave radiation. In NEMO, the surface wind stress, the sensible and latent heat fluxes, and the rate of evaporation were calculated using the forecast variables and the bulk formulae of the Coordinated Ice-Ocean Reference Experiments [40]. In FVCOM, the corresponding calculations were based on the Coupled Ocean-Atmosphere Response Experiment (COARE) version 3.0 algorithm [41], which was modified to use the specific humidity (instead of the relative humidity) as the input. Note that for the solutions provided for evaluation, the forcing of surface freshwater flux, due to evaporation and precipitation, was included in NEMO but not in FVCOM. According to the evaluation results (presented in the following sections), this difference was not a main factor for the difference in the model solutions.

Finally, both NEMO and FVCOM include parametrizations for unresolved sub-grid processes. The parameterizations used by the two models are similar but with some subtle differences, and are tuned separately to improve the performance of the models. NEMO uses a partial slip scheme for velocity near the lateral solid boundary, and the variable horizontal mixing for momentum computed with the scheme of Smagorinsky [42,43]. FVCOM specifies a zero normal component of the velocity on the lateral solid boundary and also uses the Smagorinsky eddy parameterization for the horizontal diffusion [44]. Both models adopt high-order closure schemes to compute the vertical mixing for tracers and momentum [45]: the k-ε configuration of the generic length scale scheme for NEMO and the 2.5 level scheme of Mellor-Yamada for FVCOM. Both models adopt the quadratic law for bottom drag parameterization but with different tuning of the drag coefficient to maintain the numerical stability and improve the tidal solutions. For NEMO, the drag coefficient has background values of 2.5×10^{-3}, 4×10^{-3} and 5×10^{-3} for the BoFSS1/36, BoF500 and SJAP100 components, respectively. In BoF500 and SJAP100, the drag coefficient was locally increased in the upper Bay of Fundy and in the Reversing Falls [46]. In FVCOM, the coefficient Cd is calculated by fitting a logarithmic bottom boundary layer to the model at a specified height zab above the bottom with a bottom roughness scale of z0, Cd = max(k2/ln(zab/z0)2,Cdmin) where k is the von Karman constant, z0 = 0.001, and Cdmin = 0.02. To obtain stable runs and simulate the strong dissipative effects of the rapids in the Reversing Falls, the sponge layer feature of FVCOM was used to reduce the horizontal currents in the Falls and approximately 1 km downstream using a damping coefficient of 0.015 m and a sponge radius of 500 m.

3.4. Evaluation of the Tidal Water Level and Currents

The tides dominate the variations of water levels and currents in the BoF and are a major concern for navigational safety in SJH. The tidal currents have a direct impact on the drift of spilled and other hazardous materials in the water.

The evaluation of the tidal water level and currents was based on their harmonic constants using the following metrics:

Difference in amplitude and phase:

$$D = X_m - X_o \tag{1}$$

Percent difference in amplitude:

$$D_\% = \frac{X_m - X_o}{X_o} \times 100\% \tag{2}$$

Vector difference for water level:

$$D_v = \left| A_o e^{i\varphi_o} - A_m e^{i\varphi_m} \right| \tag{3}$$

Vector difference for currents:

$$D_v = \frac{1}{T} \int \left| Q_o(t) - Q_m(t) \right| dt \tag{4a}$$

with:

$$Q(t) = u + iv = M[\cos(\omega t - \psi) + \varepsilon i \sin(\omega t - \psi)]e^{i\theta}. \tag{4b}$$

In the above equations, the subscript "*o*" denotes the observations and "*m*" denotes the model results. All the variables are defined for a single constituent. In Equations (1) and (2), X represents either the amplitude or phase. In Equation (3), A and φ denote the amplitude and phase of water level, respectively [26]. In Equations (4a) and (4b), Q is the 2D tidal flow (u, v) represented in complex form (with $i = \sqrt{-1}$), t is the time and T is the period of a specific tidal constituent with frequency ω. The other symbols are related to the tidal current ellipse: M is the length of the semi-major axis, ε is the eccentricity (defined as the ratio of the minor axis to the major axis), θ is the orientation, and ψ is the phase of the complex current [46].

For an area like the BoF, where there is significant spatial variability in the tides, percent difference is a more useful metric than the absolute difference. Vector difference (Equations (3) and (4)) combines the errors in amplitude and phase into one metric. Vector difference was the primary metric for evaluating the models' overall skill in reproducing the tides, but evaluating the amplitude and phase separately was useful for understanding the root of the errors.

Harmonic analysis on the water level and currents used the t_tide package for [47] with the Rayleigh criteria equal to 1. For stations with water level observations during 1 May 2015–30 April 2016, which include the real-time gauges and the moored ADCPs, the model output was extracted for the same time period as the observations. We also obtained harmonic constants derived from historical observations at the "constituent-only" stations. At these stations, the harmonic constants of the model solutions were obtained by performing the analysis on the full 1 year model output, regardless of the length of observations that were used to compute the observed constituents. For tidal currents, the model output was interpolated to the observed depths. Constituents (major and minor axis, inclination and phase) were computed for each depth bin of the observations. The evaluation focused on selected depths: the "surface", and 5, 10 and 20 m from the surface. Here, the "surface" is the uppermost bin with >75% ping return for the entire time series.

The polar plots in Figure 3 visually report the tidal water level metrics for each station in the inner harbour (Figure 3a) and outer harbour (Figure 3b), for the M2 constituent. The tidal ellipses in Figure 3 are representative examples of the M2 current at the surface from one station in the inner harbour (Figure 3c) and one station in the outer harbour (Figure 3d). Table 2 summarizes the evaluation results for the M2 constituent, for several selected metrics: D% (amplitude), D (phase) and Dv for water levels; and D (major axis), D (phase), and Dv for currents. The mean and the standard deviation of the mean for each metric, across the stations in the inner harbour and outer harbour, separately, are listed. Note that the means and standard deviations were computed using the absolute value of the metric so that the values reflect the average error in the model, regardless of whether the error was positive or negative. Similar tables were produced for other tidal constituents, but, for brevity, they are not presented here.
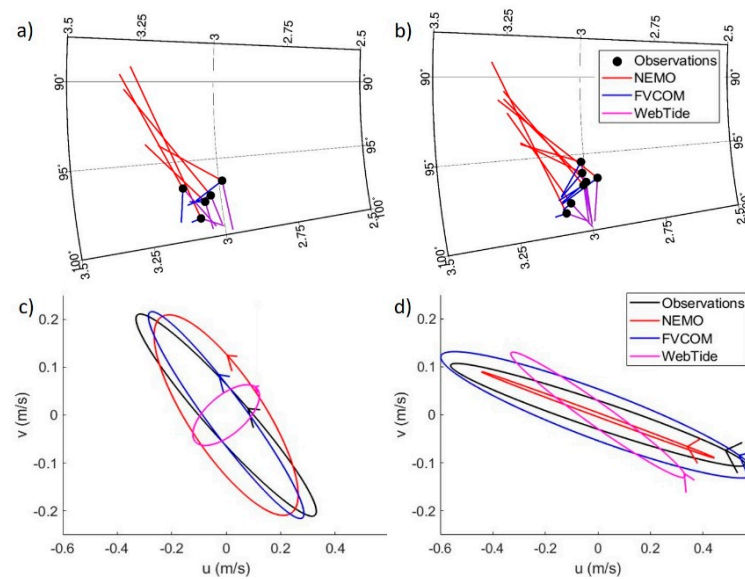
Figure 3. (**a**,**b**) Polar plot of the amplitude (m) and phase (degrees) of the observations (black circles) versus the model results in the inner harbour (**a**) and outer harbour (**b**). The size and direction of the vectors (NEMO, red; FVCOM, blue; and WebTide, purple) indicate the difference in the amplitude and phase between the observations and the model. Regional ice-ocean prediction system (RIOPS) results were omitted from these plots. (**c**,**d**) Observed and modelled M2 tidal ellipses for the surface currents from one station in the inner harbour (**c**) and one in the outer harbour (**d**). Position of the arrow on the ellipse indicates the phase, and the direction of the arrows indicates the counter-clockwise rotation of the complex current.

Table 2. The evaluation metrics for the M2 tidal water level and currents (see definitions in Equations (1)–(4) and in the text), averaged for the stations in the inner and outer Saint John Harbour (SJH). Numbers in brackets are the standard deviations across the stations. The metrics are computed for the NEMO, FVCOM, RIOPS and the WebTide solutions, separately. Note that for D and $D_\%$, the mean and standard deviations are calculated using the absolute values of the D and $D_\%$ values at each station.

		NEMO	FVCOM	RIOPS	WebTide
	Water Level (M2)				
Inner Harbour	$D_\%$ (amplitude)	7.68 (1.59)	2.48 (1.20)	21.78 (2.38)	0.93 (0.94)
	D (phase, degrees)	5.15 (2.69)	1.03 (0.67)	56.31 (0.80)	2.48 (1.09)
	D_v (m)	0.38 (0.12)	0.09 (0.04)	2.61 (0.01)	0.14 (0.04)
Outer Harbour	$D_\%$ (amplitude)	6.99 (1.17)	1.89 (1.08)	21.85 (2.55)	1.43 (1.21)
	D (phase, degrees)	6.03 (2.80)	0.72 (0.88)	55.58 (0.82)	2.07 (0.97)
	D_v (m)	0.40 (0.14)	0.08 (0.04)	2.60 (0.07)	0.12 (0.04)
	Tidal Current (M2)				
Inner Harbour	D (maj. axis, m s^{-1})	0.04 (0.03)	0.03 (0.02)	-	0.15 (0.16)
	D (phase, degrees)	38.77 (36.40)	13.21 (6.19)	-	72.05 (46.56)
	D_v (m s^{-1})	0.08 (0.03)	0.07 (0.03)	-	0.18 (0.09)
Outer Harbour	D (maj. axis, m s^{-1})	0.16 (0.10)	0.04 (0.03)	-	0.25 (0.09)
	D (phase, degrees)	7.65 (4.04)	7.95 (4.77)	-	23.40 (8.11)
	D_v (m s^{-1})	0.13 (0.06)	0.08 (0.03)	-	0.21 (0.06)

For the tidal water levels, according to Dv, FVCOM performed the best overall, followed by WebTide, NEMO, and then RIOPS. The performance of the models was not significantly different between the inner harbour and the outer harbour. Compared to WebTide, FVCOM had a slightly larger error in amplitude but a smaller error in phase. The larger error in NEMO was due to an error in the input files that was identified toward the end of the evaluation and corrected afterward [25]. RIOPS significantly underperformed compared to the other models, mainly due to the coarse spatial resolution that cannot capture the resonance characteristics of the M2 tide in the Bay of Fundy and uncalibrated tides in general.

For tidal currents, according to the Dv metric, FVCOM and NEMO performed similarly in the inner harbour, but FVCOM performed slightly better than NEMO in the outer harbour. FVCOM had smaller errors in the phase in the inner harbour, and smaller errors in the major axis in the outer harbour. Both FVCOM and NEMO significantly outperformed WebTide, mainly because the WebTide does not include the baroclinic dynamics. The errors for the RIOPS currents were not quantified but were expected to be significantly larger.

3.5. Evaluation of the Non-Tidal (Residual) Water Level and Currents

The non-tidal (or residual) component of the water level is important for predicting extreme water level events like storm surges. Residual currents dominate the net (after averaging out tidal excursions) evolution of the drift trajectory and contribute to extreme events of concern for navigational safety.

The residual water level and currents were computed by removing the tidal component (as determined by t_tide) from the observations and model output. An additional filter was applied to remove the energy in the tidal period bands (22–28 h, 11–14 h, 5–7 h). For the currents, all energy at periods < 7 h was removed.

The following metrics were used to evaluate the residual water level and the currents at each station (and depth for currents):

Mean bias:

$$\overline{D} = \overline{X_m} - \overline{X_o} \tag{5}$$

Root mean square error:

$$RMSE = \left\{ \frac{1}{N} \sum_{i=1}^{N} (X_m - X_o)^2 \right\}^{1/2} \tag{6}$$

Relative average error:

$$RAE = 100\% \frac{\sum_{i=1}^{N} (X_m - X_o)^2}{\sum_{i=1}^{N} \left(\left| X_m - \overline{X_o} \right|^2 + \left| X_o - \overline{X_o} \right|^2 \right)} \tag{7}$$

Correlation:

$$R = \frac{\sum_{i=1}^{N} \left(X_m - \overline{X_m} \right) \left(X_o - \overline{X_o} \right)}{\left[\sum_{i=1}^{N} \left(X_m - \overline{X_m} \right)^2 \left(X_o - \overline{X_o} \right)^2 \right]^{\frac{1}{2}}} \tag{8}$$

Skill:

$$Skill = 1 - \frac{\sum_{i=1}^{N} |X_m - X_o|^2}{\sum_{i=1}^{N} \left(\left| X_m - \overline{X_o} \right| + \left| X_o - \overline{X_o} \right| \right)^2} \tag{9}$$

In the above equations, X denotes the time series of the water level or a velocity component, and the overbar denotes its time mean. The other notations are consistent with those used in Equations (1)–(4). In conjunction with the above metrics, we also compared the power spectra of the time series.

Figure 4 presents the observed and modelled time series of the residual water level at the Saint John tide gauge location and the associated spectra. The shaded areas in Figure 4a–c highlight the time frame of the Fall freshet (September 2015) and the Spring freshet (April 2016), which are shown in detail in Figure 4d,e, respectively. NEMO and FVCOM obtained consistent results for the two freshet events, including the deficiencies after the peak water levels. That is, the modelled water levels were lower than those observed after the peak of the fall freshet, and higher than that observed after the peak of the spring freshet. The causes of these discrepancies were not further explored. The three models (NEMO, FVCOM and RIOPS) all captured the main characteristics of the observed variations of residual water level over a full year period, including similar spectral distributions.
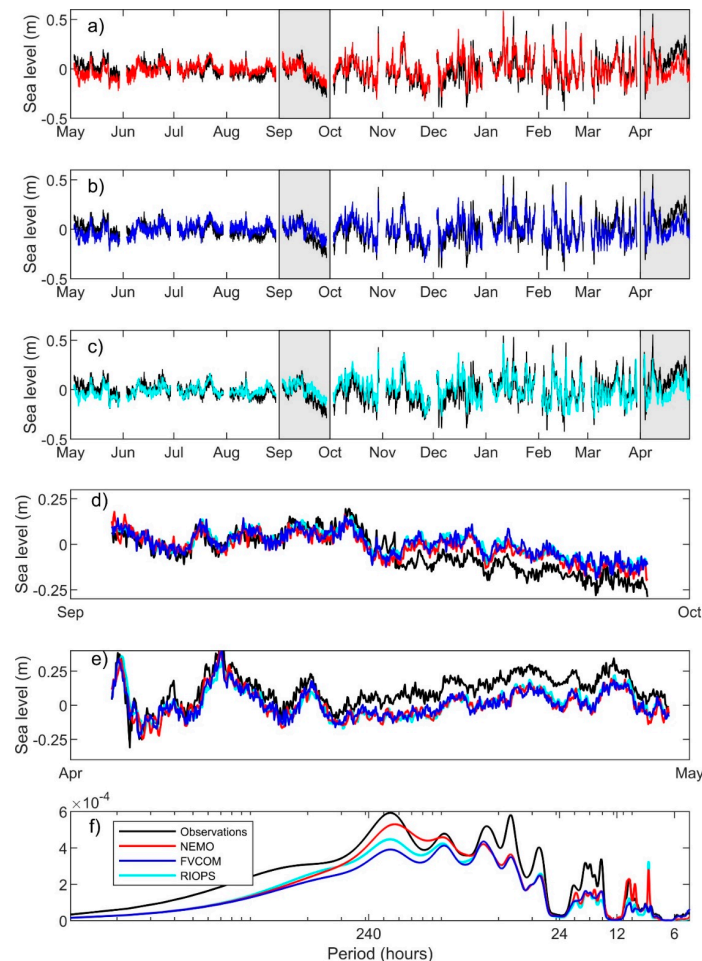
Figure 4. (**a**–**c**) One year time series of the residual water level at the Saint John tide gauge location from (**a**) the NEMO, (**b**) the FVCOM, and (**c**) the RIOPS. Observations are shown in black in each panel. The shaded areas in (**a**–**c**) highlight the time frame of the Fall freshet (September 2015) and the Spring freshet (April 2016), and are expanded in (**d**) and (**e**), respectively. (**f**) The spectra for the time series in (**a**–**c**) in variance-conserving form. The legend in panel (**f**) applies to all panels. The residual water level was computed by removing the tidal component (as determined by t_tide). An additional filter was applied to remove the extra energy in the tidal period bands (22–28 h, 11–14 h, 5–7 h).

Table 3 summarizes the results as the mean and standard deviation across all stations (and the depths for currents) for each metric. For the residual water level, the metrics were computed on the demeaned times series. The mean bias for the water level across all the stations (including the ADCP measurements) was not computed. NEMO, FVCOM and RIOPS showed very similar values for the three metrics (RMSE, RAE and correlation) listed. There are also no significant differences in the skills between the inner and outer harbours. The correlation of both models with the observed water level

was high; on average, 0.74 in the inner harbour, and 0.78 in the outer harbour, but as much as 0.9 at some stations. The metrics showed no significant difference in model performance between the inner harbour and the outer harbour, but the spectra from the Saint John tide gauge indicates that, in the inner harbour, the NEMO model had better representation of the energy in the 5–10 day period band, and the FVCOM captured more of the high-frequency (<6 h) energy.

Table 3. The evaluation metrics for the residual water level and currents (see definitions in Equations (5)–(8) and in text), averaged for the stations in the inner and outer SJH, separately. Numbers in brackets are the standard deviations across the stations. RMSE, RAE, R and Skill are defined in Equations (6)–(9). Note that for the water levels, the mean bias is not computed and the other metrics are computed for the demeaned time series. Water level metrics are computed for NEMO, FVCOM and RIOPS. Current metrics are computed for NEMO and FVCOM only.

			NEMO	FVCOM	RIOPS
			Residual Water Level		
Inner Harbour		RMSE (m)	0.08 (0.02)	0.08 (0.02)	0.08 (0.02)
		RAE (%)	26.60 (14.85)	27.78 (13.33)	29.41 (12.32)
		R	0.74 (0.15)	0.74 (0.13)	0.73 (0.12)
		Skill	0.85 (0.09)	0.84 (0.08)	0.82 (0.08)
Outer Harbour		RMSE (m)	0.09 (0.03)	0.09 (0.03)	0.10 (0.03)
		RAE (%)	26.66 (17.83)	28.57 (18.28)	28.17 (17.13)
		R	0.78 (0.12)	0.78 (0.12)	0.77 (0.12)
		Skill	0.85 (0.11)	0.83 (0.12)	0.84 (0.11)
			Residual Currents		
Inner Harbour	u	Mean Bias (m s^{-1})	−0.05 (−0.05)	−0.04 (0.04)	-
		RMSE (m s^{-1})	0.08 (0.02)	0.07 (0.03)	-
		RAE (%)	77.68 (13.46)	73.33 (17.19)	-
		R	0.42 (0.19)	0.47 (0.23)	-
		Skill	0.51 (0.10)	0.51 (0.10)	-
	v	Mean Bias (m s^{-1})	0.04 (0.07)	0.02 (0.05)	-
		RMSE (m s^{-1})	0.08 (0.04)	0.07 (0.02)	-
		RAE (%)	68.75 (14.28)	74.12 (10.71)	-
		R	0.53 (0.17)	0.49 (0.20)	-
		Skill	0.58 (0.10)	0.55 (0.05)	-
Outer Harbour	u	Mean Bias (m s^{-1})	−0.04 (0.02)	−0.08 (0.02)	-
		RMSE (m s^{-1})	0.07 (0.01)	0.10 (0.02)	-
		RAE %)	80.82 (11.85)	88.69 (10.08)	-
		R	0.26 (0.12)	0.32 (0.14)	-
		Skill	0.52 (0.08)	0.50 (0.06)	-
	v	Mean Bias (m s^{-1})	0.06 (0.03)	0.00 (0.02)	-
		RMSE (m s^{-1})	0.07 (0.02)	0.04 (0.01)	-
		RAE (%)	95.08 (7.22)	82.45 (9.33)	-
		R	0.15 (0.10)	0.23 (0.11)	-
		Skill	0.38 (0.06)	0.49 (0.06)	-

Figure 5 presents examples of the observed and modelled residual current at the surface from representative ADCPs in the inner harbour (a,b) and the outer harbour (c,d). As reflected in the figures, both NEMO and FVCOM captured the timing of large-scale events but underestimated the magnitude of the currents in the inner harbour. The magnitude of the currents was better predicted by the models in the outer harbour. The metrics in Table 3 indicate that both NEMO and FVCOM showed a better performance in the inner harbour (with average correlations in u an v between 0.42 and 0.53, respectively) compared to the outer harbour (average correlation between 0.15 and 0.32). We did not evaluate the models against RIOPS because its coarse horizontal resolution poorly represents the spatial variations of currents in SJH.
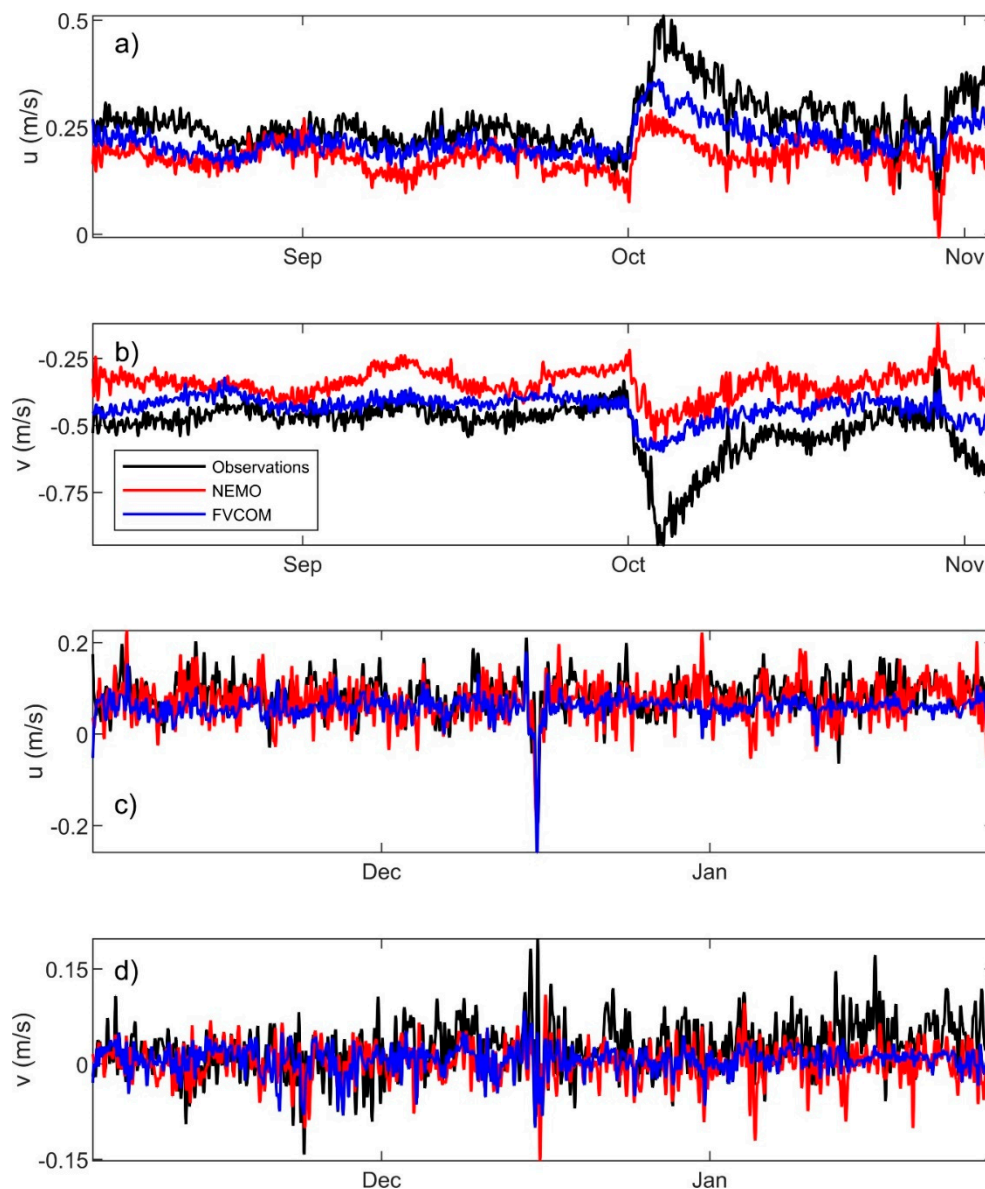
Figure 5. Observed (black) and modelled (NEMO, red; FVCOM, blue) time series of the u and v components of the residual currents, at the representative stations in the inner harbour (**a**,**b**) and outer harbour (**c**,**d**). The residual currents were computed by removing the tidal component (as determined by t_tide). An additional filter was applied to remove the extra energy in the tidal period bands (22–28 h, 11–14 h) and all energy at periods < 7 h.

3.6. Evaluation of the Temperature and Salinity Fields

Variations in the water temperature and salinity cause changes in density, density-driven currents, the stratification of the water column, and the chemistry (which subsequently influences the fate and behavior of spilled oil) of the water.

Variations in the temperature and salinity in SJH were first evaluated qualitatively for three cases: low water, spring freshet and fall freshet. Salinity variations are mainly attributed to the input of freshwater from the Saint John River and the mixing of freshwater and seawater in the harbour. Because NEMO and FVCOM were forced by the observed water level at Oak Point station (in the river estuary system), we first qualitatively diagnosed the freshwater flux (river runoff) across a section a few kilometers downstream of Oak Point and determined that the diagnosed fluxes were similar in both models. Figure 6 presents the surface density fields from NEMO and FVCOM, as well as

satellite-derived images of suspended particulate matter (provided by E. Devred, DFO), during two different phases of the tide. These figures show that the models obtained a similar extent of the river plume, consistent with the interpretation of satellite images.
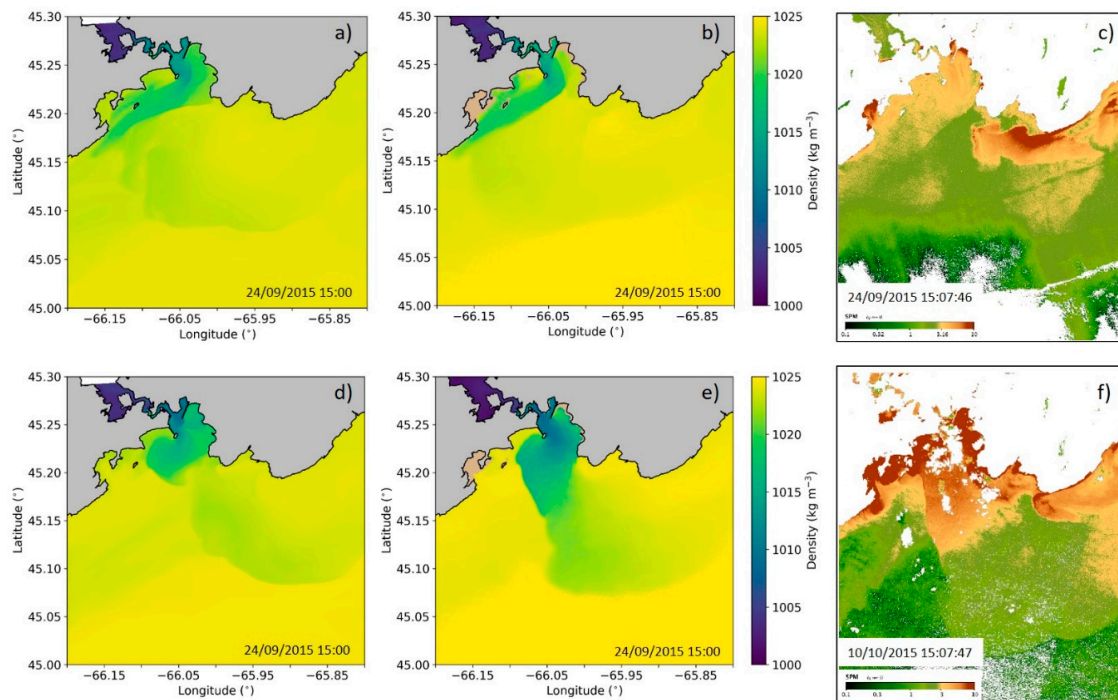

Figure 6. Snapshot of the surface density (kg m^{-3}) from the NEMO (**a**,**d**), and the FVCOM (**b**,**e**) within 15 min of the satellite-derived images of suspended particulate matter (**c**,**f**) provided by E. Devred, Fisheries and Oceans Canada (DFO). The dates and times of the satellite images, as well as the modelled surface density field are indicated.

The SmartAtlantic Buoy provided the only time series measurement of SST in SJH and was essential for evaluating the annual cycle of the SST. Figure 7 presents the comparison of the observed and modelled SST from the NEMO and FVCOM. The mean bias, RMSE, RAE, correlation and skill (as described in Equations (5)–(9)) were computed for the time series of SST from NEMO, FVCOM and RIOPS, and are presented in Table 4. Because the temperature fields from RIOPS were only available as daily means, the metrics for RIOPS were computed using the daily mean of the observations. Both NEMO and FVCOM captured the observed seasonal cycle as well as the high-frequency variations, but there was a warm bias in the FVCOM model. FVCOM performed slightly better than NEMO in terms of correlation (0.96 for FVCOM and 0.94 for NEMO), and both models outperformed RIOPS across all metrics except mean bias.

Figure 8 compares the observed vertical profiles of temperature and salinity from the CTD casts with the model simulations at two representative stations: one in the inner harbour (a,b), and one in the outer harbour (c,d). The temperature and salinity profiles fluctuate significantly with the phase of the tide in SJH, particularly for salinity due to the interaction of river runoff, tidal flow, and mixing. The observed profiles were compared to the nearest model output in time (t) and space. The figures also include shaded areas that represent the profiles within t − 3 h and t + 3 h which show how the modelled profiles varied over a 6 h period. As expected, the figures show more variation in the modelled T and S profiles in the inner harbour than in the outer harbour. For the qualitative evaluation of the profiles, the model results were "acceptable" if the observed profile fell within the shaded area. Generally speaking, NEMO showed a better representation of the vertical stratification. FVCOM obtained more uniform temperature and salinity profiles, probably due to too strong vertical mixing that needs to be tuned.
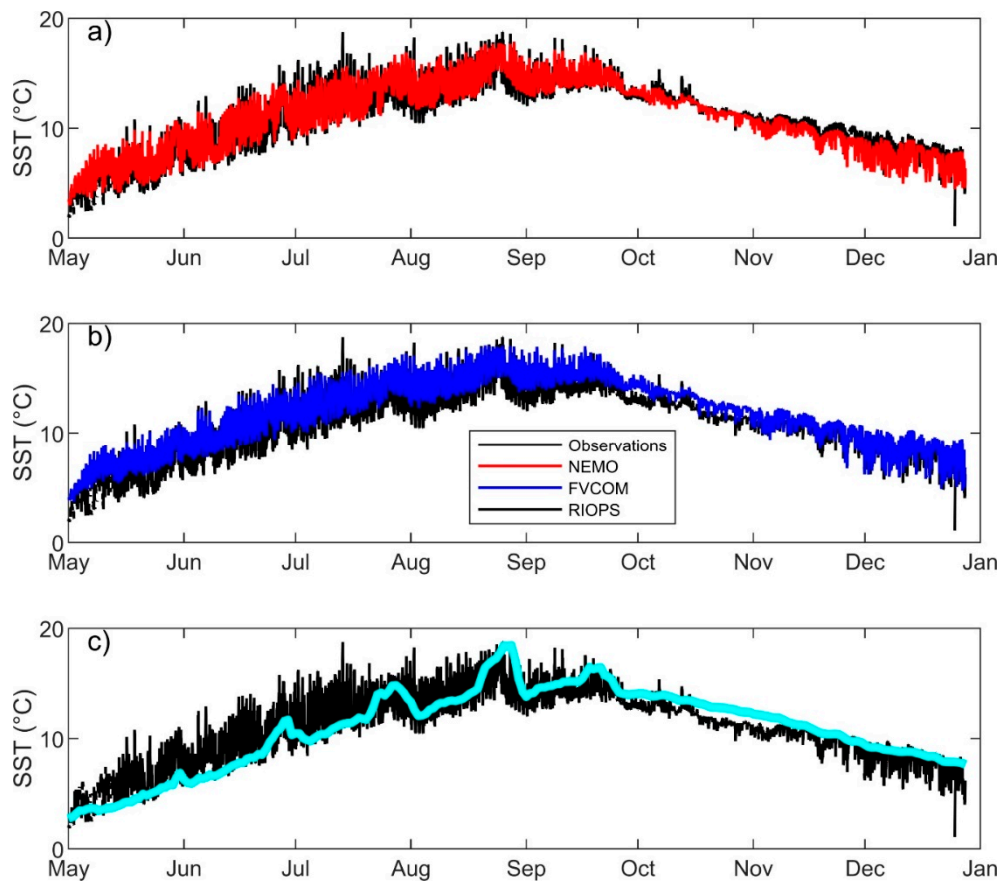
Figure 7. Time series of the sea surface temperature (SST) from (**a**) the NEMO, (**b**) the FVCOM, and (**c**) the RIOPS compared to the observed SST at the SmartAtlantic Buoy. RIOPS times series is based on daily output.
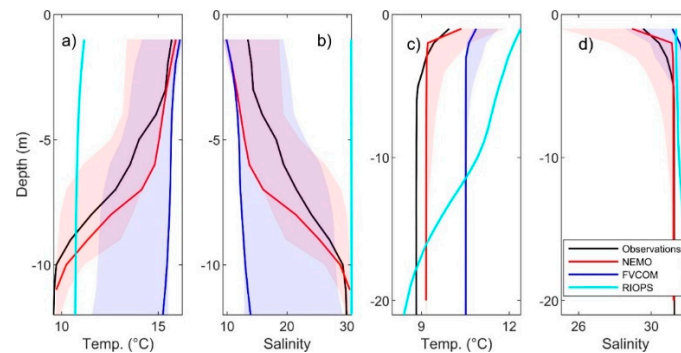
Figure 8. Observed (black) and modelled (NEMO, red; FVCOM, blue; RIOPS, cyan) temperature (**a,c**) and salinity (**b,d**) profiles from the representative stations in the inner harbour (**a,b**) and outer harbour (**c,d**). The model profiles are taken at the nearest hour to the observations, h. The shaded area around the NEMO and FVCOM profiles (light pink for the NEMO and light blue for the FVCOM) represents the profiles within $t - 3$ h and $t + 3$ h to show how the modelled profiles vary over a 6 h period. Where the areas overlap, the shading appears more purple.

The mean bias, RMSE, and RAE were computed for each profile, and the correlation and skill were computed using all the profile data in the inner harbour and outer harbour, separately. Table 3 summarizes the results, listing the mean and standard deviation across all stations in the inner harbour and outer harbour, separately, for mean bias, RMSE, and RAE. According to the correlation metric, NEMO and FVCOM performed better in the inner harbour than the outer harbour. With respect to salinity, the summarized metrics indicate that the errors for NEMO were lower than FVCOM,

but a t-test shows no significant difference. The correlation metric indicates that NEMO had a slight advantage over FVCOM in the inner harbour, while the opposite was true in the outer harbour.

Table 4. Evaluation metrics for the time series of the SST (compared to the observations from the SmartAtlantic (SA) Buoy), and the water temperature and salinity (compared to the observed CTD profiles). See the definitions of the metrics in Equations (5)–(9) and in the text. The mean bias (D), RMSE and RAE are averaged for the stations in the inner and outer SJH. Numbers in brackets are the standard deviations across the stations. Correlation and skill were computed using all the profile data in the inner and outer harbour, separately.

		NEMO	FVCOM	RIOPS
		SST (time series)		
SA Buoy	D (°C)	0.08	0.84	0.00
	RMSE (°C)	1.05	1.23	1.63
	RAE (%)	5.62	7.82	11.81
	R	0.94	0.96	0.89
	Skill	0.97	0.96	0.94
		Water temperature		
Inner Harbour	D (°C)	0.22 (0.55)	1.62 (0.72)	0.06 (1.91)
	RMSE (°C)	0.70 (0.32)	1.76 (0.90)	1.64 (1.36)
	RAE (%)	43.03 (38.76)	82.12 (20.20)	90.23 (9.08)
	R	0.98	0.96	0.80
	Skill	0.98	0.91	0.89
Outer Harbour	D (°C)	0.21 (0.45)	1.47 (0.34)	0.86 (0.59)
	RMSE (°C)	0.50 (0.44)	1.56 (0.27)	1.13 (0.48)
	RAE (%)	55.28 (30.08)	90.75 (16.47)	93.35 (7.17)
	R	0.99	0.99	0.98
	Skill	0.99	0.95	0.98
		Salinity		
Inner Harbour	D (psu)	−0.89 (1.70)	−2.04 (3.04)	5.80 (7.33)
	RMSE (psu)	2.28 (1.46)	3.28 (3.15)	7.48 (7.01)
	RAE (%)	29.00 (31.49)	44.13 (36.87)	98.80 (1.12)
	R	0.97	0.93	0.08
	Skill	0.98	0.96	0.44
Outer Harbour	D (psu)	−0.06 (0.96)	0.76 (0.94)	0.72 (1.17)
	RMSE (psu)	1.31 (1.32)	1.45 (1.30)	1.92 (1.82)
	RAE (%)	39.54 (30.47)	50.65 (31.13)	97.35 (3.45)
	R	0.75	0.79	0.33
	Skill	0.86	0.86	0.32

The high-variability of temperature and salinity in Saint John Harbour, particularly when close to the mouth of the river, was difficult to evaluate with point measurements, but nonetheless, the analysis revealed the presence of a warm/salty bias in the FVCOM model. The warm/salty bias in FVCOM has since been corrected and was due in small part to a technical issue in the implementation of the COARE 3.0 algorithm that was identified toward the end of the evaluation, and in large part due to the placement of the outer boundary extending past the shelf break. That being said, both models performed better than RIOPS due to the inclusion of the freshwater flux from the Saint John River.

3.7. Evaluation of Drifter Simulations

Improving drift prediction is one of the primary objectives of the oceanography component of the OPP. Hence, our evaluation process includes a simulation of the surface drift trajectories and a comparison with the observed trajectories of the four types of surface drifters deployed in the SJH.

Drifter simulations were performed using the Canadian Oil Spill Modelling Suite (COSMoS) [48]. COSMoS is a Lagrangian displacement model that solves the equations of motion for drift using surface currents and surface wind. It takes surface currents from the results of models on the structured or unstructured native grids. Note that currents are computed at the centre of the FVCOM grid elements

so that the solutions had to be interpolated to the nodes. The computation of the drifter trajectories uses a fourth order Runge-Kutta scheme. The simulations were run with surface currents from the NEMO (SJAP100 and BoF500), FVCOM, and WebTide. A simulated drifter was "released" at the same time and location of each observed drifter, and the simulation was terminated at the end of the observed drifter release. The drift evaluation was restricted to the smallest common domain, which is the SJAP100 domain. Even though BoF500 has a lower resolution in this area, the drift simulations were completed using currents from both SJAP100 and BoF500 to help understand the influence of increased resolution on drift prediction accuracy.

The inclusion of surface wind in the drifter simulations is to account for supplemental wind effects on the drifting object that are not included in the surface currents. However, the true wind effects may differ for different types of drifters. Simulations were run with 0% and 3% of the wind speed (10 m winds from HRDPS) to evaluate the effect. The results consistently showed that adding 3% of wind to the drifter simulations slightly improved the overall skill of the NEMO-based trajectories, but degraded the FVCOM-based solutions. Thus, the results presented here focus on the solutions with no added wind effect.

The first aspect of the drift evaluation was to qualitatively compare the distributions of surface velocities derived from the observed and modelled drifter trajectories. Drifter velocity is approximated as a forward Euler difference between the successive positions and reported in along-shore (241 °T positive) and cross-shore (151°T positive) components. Note that without including the wind effect, the "drifter velocity" from the models are very close to the surface currents. Figure 9a–e shows the distributions of drifter velocities for the Davis drifters, but the results were similar for all drifter types. The mean and RMS speeds derived from these distributions were within a few percent of each other for all cases, and the shapes of the distributions were generally similar. However, the distribution of the observed drift velocities appears to have heavier tails than the modelled drift velocity distributions, which is consistent with the underprediction of currents noted in Section 3.5. As expected, the distributions based on WebTide were more bi-modal (consistent with ebb and flood tide) in the along-shore direction and more peaked around 0 in the cross-shore direction. They showed no mean speed, and underpredicted the variability in the drift speeds. Such consequences of WebTide having no baroclinic flow was evident here, as well as across all the evaluations of the drifter trajectories (panels d, i, n, s, w, aa of Figure 9). Therefore, the WebTide-based results are not discussed further in this paper.

The second aspect of the drift evaluation was to compare the absolute dispersion of the observed and modelled drifters by plotting distributions of along-shore and cross-shore drifter displacements as a function of time lag, τ. Mathematically:

$$\Delta X(\tau) = X(t - \tau) - X(t) \tag{10}$$

where ΔX is the change in along-shore or cross-shore displacement during time lag τ. To account for all the possible displacements in the area sampled by the drifters, all non-overlapping segments of drifter tracks were used in the analysis. This is equivalent to assuming homogeneous and stationary absolute dispersion statistics.

Figure 9f–o shows the dispersion of the observed and modelled Davis drifters for up to 24 h. Displacements in the along- and cross-shore directions (y axis) are binned at hourly intervals (x axis) and the density of the observations in each bin is plotted to visualize the temporal evolution of the drifter 'cloud'. The statistics of along-shore dispersion display a clear 'tidal' character, that is, the expansion and contraction of the 'displacement cloud' on a ~12 h cycle. This is superimposed on a mean down-coast (along-shore positive) drift. Oscillations were not as evident in the cross-shore displacements. These characteristics were consistent across all four drifter types.

Generally, the distributions of displacement produced by the NEMO and the FVCOM were similar to the observed statistics, though extreme displacements vary slightly between models and observations. Both FVCOM and NEMO retained the general shape of the tracks, but on average the

modelled along-shore displacements were lower than the observed values, indicating that the motion was characteristically similar but occurred closer to Saint John in the models than was observed. In the cross-shore direction, both the FVCOM and the NEMO predicted a larger mean displacement than was observed, indicating that the modelled drifters travelled further offshore than the observed ones.
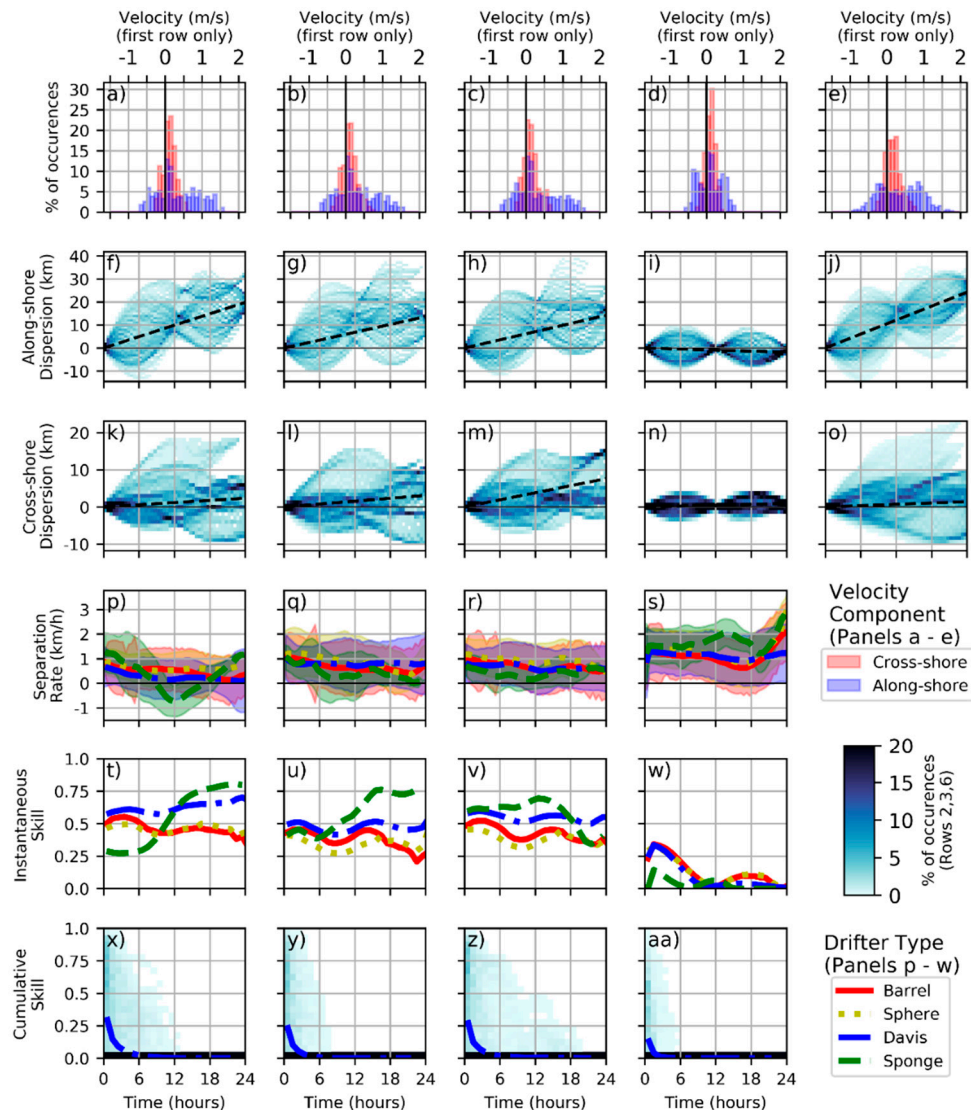
Figure 9. Results of the drifter analysis for the simulated drifters using the FVCOM (left column), SJAP100 (second from left), BoF500 (middle), WebTide (second from right), and observations where applicable (right column). Distributions of along-shore (blue) and cross-shore (red) drifter velocity (for the Davis drifters) are shown in the first row (**a**–**e**). Absolute dispersion in the along-/cross-shore directions (for the Davis drifters) is shown in the second and third row (**f**–**o**) with the mean drift superimposed on the density of the simulated/observed positions in each direction. Separation rate between the simulated drifters and the corresponding observations is shown in the fourth row (**p**–**s**) for all four drifter types. Instantaneous skill corresponding to the various model predictions is shown in the fifth row (**t**–**w**) for all drifter types, and the cumulative skill for the prediction of the Davis drifter trajectories is shown in the sixth row (**x**–**aa**), with the mean cumulative skill as a function of time is shown as a solid blue line.

The third aspect of the drift evaluation was to quantitatively assess the simulations using the following metrics:

Separation rate:

$$v_{sep} = \frac{\Delta\left(\vec{x}_{mod} - \vec{x}_{obs}\right)}{\Delta t} \tag{11}$$

Instantaneous skill:

$$Skill_{ins} = 1 - \min\left(\frac{d_i}{S_i}, 1\right) \tag{12}$$

Cumulative skill:

$$Skill_{cum} = 1 - \min\left(\frac{\sum d_i}{\sum D_{obs,i}}, 1\right) \tag{13}$$

Separation rate is defined as the relative rate of change between the modelled drifter position $\vec{x}_{mod}$ and the observed drifter position $\vec{x}_{obs}$. This is a basic metric for trajectory model performance that allows for the simple interpretation of the quality of future forecasts. It is visualized here by plotting the separation between the modelled and observed trajectories as a function of time (Figure 9p–s). The 0FVCOM-based trajectory predictions showed the slowest separation rate, though the results varied based on the drifter type. The FVCOM performed best when predicting the tracks of the CODE/Davis drifters, which suggests that the modelled currents averaged over the upper meter of the ocean correspond best with currents at 0.5 m depth (the approximate centroid of the subsurface portion of the CODE/Davis drifters). Separation rates from SJAP100 and BoF500 were mostly similar, but BoF500 performed slightly better at longer time scales.

Instantaneous skill (adapted from Molcard et al. [49]) aims to answer the question, 'To what extent can the model help us locate a drifting object given a last known position and associated time?'. It defines the model's skill as a function of absolute separation from the initial location at an instant in time, denoted by the ith discrete time interval. In Equation (12), d_i is the separation between the observed and modelled trajectories at a given point in time (denoted by subscript i), and S_i is the linear distance between the observed drifters' position and its deployment location. The normalization by S_i ensures that skill is assigned based on the accuracy of the prediction relative to the magnitude of the observed displacement. For example, a prediction that is within 500 m of the corresponding observation is assigned a higher skill if the observed drifter has travelled 5 km than if it has travelled 1 km. This implies that the model skill can increase with time if Si increases more rapidly than d_i. This behaviour is clearly evident in the results for Sponge drifters, which exhibited a large down-coast displacement. A non-zero instantaneous skill score indicates that the model improves our knowledge of the drifting object's position from the guess that it is still at the last known position. Therefore, all non-zero skill scores indicate that the model is useful, with higher skill scores indicating a narrower margin of error. FVCOM simulations had the highest instantaneous skill score, and perform best when simulating the tracks of the CODE/Davis drifters.

Cumulative skill (taken from Liu and Weisberg, [50]) answers the question, 'How well can the observed trajectory be reproduced when all the aspects of the trajectory evolution are considered?'. It quantifies the skill of the model using cumulative sums to account for the model's ability to reproduce the magnitude, direction, and timing of drift events, and the range of spatial scales in the flow that are encountered by a drifter along its trajectory. In Equation (13), d_i is as defined above, and $D_{obs,i}$ is the distance travelled by the observed drifter during the time i. Cumulative skill is a demanding test of the model's ability to reproduce the timing, magnitude, and direction of drift events, and any skill score above zero was considered an indication of useful model predictions. Figure 9x–aa shows the cumulative skill scores for the prediction of Davis drifter trajectories. Upon close inspection, the FVCOM model slightly outperforms SJAP100 with the solutions indicating non-zero skill for the forecasts up to (and slightly exceeding) 6 h for all drifter types. Again, the FVCOM's increased performance for the CODE/Davis drifters is evident in the cumulative skill scores (Figure 9x–y). We note that both models showed no cumulative skill for >50% of the observed trajectories, however,

this might be expected for a highly dynamic region such as SJH where small differences in predictions near the start of the trajectory can lead to substantial deviations at later times.

Interestingly, across all metrics, simulations using BoF500 outperformed the SJAP100 configuration, and according to the cumulative skill scores, BoF500 produced the longest skillful forecast overall—a significant non-zero skill up to 19 h for the CODE/Davis drifters (Figure 9z). This suggests that increased spatial resolution does not necessarily improve the representation of the currents in the uppermost layer of the water column, and a smoothed solution may provide better trajectory predictions. This is related to the unconstrained turbulence in the higher resolution configuration, explained further in Jacobs et al. [51].

3.8. Evaluation of Model Efficiency

The evaluation of model efficiency included a comparison of (1) the number of cores and run-time required to run the models, and (2) the scalability of the models which describes the model's ability to run faster on more cores. The performance of the two models can be directly compared because both were run on the same high-performance computing platform of the Government of Canada. Both the NEMO and FVCOM models for Saint John Harbour met the operational requirement (a 48 h simulation in less than 30 min), but the FVCOM model had much less computational cost overall. For a 48 h simulation, the FVCOM used eight slots (24 cores per slot) with a run-time of 19.5 min; and the three-level nested NEMO model used 31 slots with a run-time of 29.1 min. The high computational cost of NEMO was mainly due to the 100 m configuration (SJAP100). Figure 10 shows that the run time of SJAP100 rapidly decreased when the number of slots increased from 10 to 18, and then gradually decreased with an increasing number of slots. For the FVCOM, the rapid decrease in run time occurred as the number of slots increased from 5 to 8. The scalability curve for the 500 m configuration of the NEMO (BoF500) is closer to that of the FVCOM. The requirements on the number of computer slots are related to the difference in the numbers of computed nodes (grids) among different models: the FVCOM has 56,635 nodes and 108,301 elements, while SJAP100, BoF500 and BoFSS1/36 have 466,718, 161,670 and 39,600 grid points, respectively. This demonstrated a major computational efficiency in using the unstructured grid FVCOM over the structured NEMO for nearshore applications. Despite having a higher resolution in SJH, the FVCOM ran faster than SJAP100.
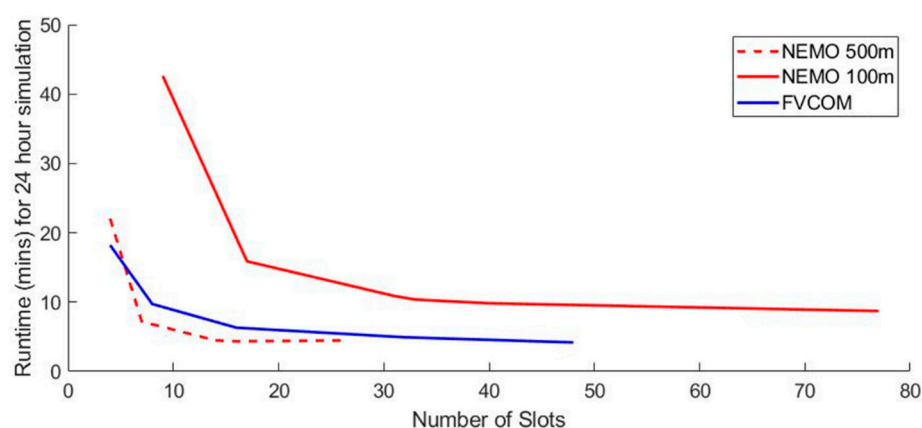
Figure 10. The run time for a 24 h simulation versus the number of slots on the General Purpose Science Computing platform of Shared Services of Canada, with the FVCOM (solid blue curve), the NEMO's 500 m (BoF500, dashed red curve) and 100 m (SJAP100, solid red curve) configurations.

4. Conclusions

This paper represents a significant collaborative effort that took place during the first year of a major Canadian research and development program, the Oceans Protection Plan. The oceanography sub-initiative of the OPP was tasked to develop port-scale models to fit into the multi-scale operational

ocean prediction systems in Canada, i.e., to extend the prediction capability from global, basin and coastal scales to nearshore waters. The work included three components: (1) the development of an evaluation process, (2) the development of model configurations with two state-of-the-art open source community models, and (3) the evaluation of the performances of the two model configurations including the comparison of each with the existing operational models. Note that the second and third components were carried out simultaneously, and both model configurations were constantly improved and fine-tuned with the guidance of the evaluation.

The evaluation process includes (1) the selection of the study area, (2) the requirements for the model setup, and (3) the metrics for evaluating the models. The selection of the study area was specific to the nature of the OPP project in that one of the six OPP pilot ports was selected, but the choice of port enabled a general assessment of the models for near-shore port-scale applications. The requirements for the model setup were quite general for evaluating the performance of different source models under the same setting, but by ensuring that both models were configured using the same sources of input data for the model setup and forcing, the differences in the model performances can be mostly attributed to the differences in grid structure, model numerics, and software technology. Thus, the results of this study are valuable for further improvements of the models. The metrics can be used for evaluating any model. They were selected to assess the various aspects of the model results, and collectively, they were able to detect minor contrasts between the two models.

The chosen study area, Saint John Harbour in the Bay of Fundy, features the presence of strong tides, significant river runoff and a narrow tidal-river channel (Reversing Falls). The complicated regional oceanography posed challenges to both the NEMO and FVCOM configurations with the full baroclinic dynamics included. The common challenge was to ensure that the models ran stably, particularly in the narrow channel with strong currents and spatial gradients of temperature and salinity. For this purpose, both models included local treatment of bottom drag parameterization. Previous configurations of the FVCOM model for the Saint John Harbour area did not include atmospheric forcing. Adding atmospheric forcing, particularly the surface heat-flux, was a challenge, but the issue was later attributed to the previously mentioned bug in the code. A challenge for NEMO was the lack of experience in creating a multiple-level nested configuration for nearshore waters. Despite the urgent timeline, both configurations were created and improved during the first year of OPP.

The evaluation metrics were defined based on the existing expertise of the team, through expert consulting and literature research, as well as the available ocean observational data. Evaluation with these metrics led to continuous improvements of both models, that is, errors in model settings (bathymetry, model parameters and forcing) were constantly identified and corrected. The evaluation results presented in Sections 3.4–3.8 were based on the model results obtained toward the end of the first year of OPP. In terms of model accuracy, compared to the observational data, NEMO and the FVCOM achieved comparable metrics for the tidal and non-tidal components of sea level and currents, seasonal variation of the sea surface temperature, vertical profiles of water temperature and salinity, and the trajectories of surface drift. Note that for the evaluation of the drifter trajectories, the impacts of winds were not fully considered. This aspect needs to be included in future work. A major difference identified was that the FVCOM required less computer resources and ran faster than the NEMO. One possible solution to reduce the required number of the computer slots is to reduce the domain size of the finest model configuration (SJAP100), and using the two-way nesting approach to ensure the dynamic interaction across different configurations are properly simulated. Initial testing on two-way nested configuration was achieved after the completion of the evaluation in the first year (results not shown). The results of the model efficiency evaluation suggest that both the NEMO and FVCOM meshes could be further refined, but in doing so, the NEMO model would not meet the operational requirement for runtime due to the limitations of the model numerics and computer efficiency.

Finally, the evaluation of the models was limited by the existing observational data. With increasing resolution, the models are reaching the limits to which the ocean is observed in terms of small scale,

rapidly varying, and nonlinear processes. To support the envisaged future of port-scale operational e-navigation systems, we need to reassess what observational capacities are required to support this, both from a monitoring point of view (i.e., real-time and high spatial resolution data), as well as for model development (e.g., delayed-mode data such as moorings).

References

1. Government of Canada. Protecting our Coasts: Oceans Protection Plan. Available online: https://www.tc.gc.ca/en/campaigns/protecting-coasts.html (accessed on 28 May 2020).
2. Government of Canada. CONCEPTS. Available online: http://www.science.gc.ca/eic/site/063.nsf/eng/h_97620.html (accessed on 28 May 2020).
3. Madec, G.; Delecluse, P.; Imbard, M.; Levy, C. *NEMO Ocean Engine*; Note du Pole de modèlisation, IPSL: Paris, France, 2015; ISSN No. 1288-1619.
4. NEMO Community Ocean Model. Available online: https://www.nemo-ocean.eu (accessed on 28 May 2020).
5. Chen, C.; Beardsley, R.C.; Cowles, G. An unstructured grid, finite-volume coastal ocean model (FVCOM) system. *Oceanography* **2006**, *19*, 78–79. [CrossRef]
6. The Unstructured Grid Finite Volume Community Ocean Model. Available online: http://fvcom.smast.umassd.edu/fvcom (accessed on 28 May 2020).
7. Smith, G.C.; Roy, F.; Reszka, M.; Surcel Colan, D.; He, Z.; Deacu, D.; Belanger, J.-M.; Skachko, S.; Liu, Y.; Dupont, F.; et al. Sea ice forecast verification in the Canadian Global Ice Ocean Prediction System. *QJR Meteorol. Soc.* **2015**, *142*, 659–671. [CrossRef]
8. Dupont, F.; Higginson, S.; Bourdallé-Badie, R.; Lu, Y.; Roy, F.; Smith, G.; Lemieux, J.-F.; Garric, G.; Davidson, F. A high-resolution ocean and sea-ice modelling system for the Arctic and North Atlantic oceans. *Geosci. Model. Dev.* **2015**, *8*, 1577–1594. [CrossRef]
9. Dupont, F.; Chittibabu, P.; Fortin, V.; Rao, Y.R.; Lu, Y. Assessment of a NEMO-based hydrodynamic modelling system for the Great Lakes. *Water Qual. Res. J. Can.* **2012**, *46*, 198–214. [CrossRef]
10. Chen, C.; Huang, H.; Beardsley, R.C.; Xu, Q.; Limeburner, R.; Cowles, G.W.; Sun, Y.; Qi, J.; Lin, H. Tidal dynamics in the Gulf of Maine and New England Shelf: An application of FVCOM. *J. Geophys. Res.* **2011**, *116*, C12010. [CrossRef]
11. Wu, Y.; Chaffey, J.; Greenberg, D.A.; Smith, P.C. Environmental impacts caused by tidal power extraction in the upper Bay of Fundy. *Atmos. Ocean* **2016**, *54*, 326–336. [CrossRef]
12. Han, G.; Ma, Z.; DeYoung, B.; Foreman, M.; Chen, N. Seasonal variability of the Labrador Current and shelf circulation off Newfoundland. *Ocean Model.* **2011**, *40*, 199–210. [CrossRef]
13. Foreman, M.G.G.; Stucchi, D.J.; Garver, K.A.; Tuele, D.; Issac, J.; Grime, T.; Guo, M.; Morrision, J. A circulation model for the Discovery Islands, British Columbia. *Atmos. Ocean* **2012**, *50*, 301–306. [CrossRef]
14. Lin, Y.; Fissel, D.B. The ocean circulation of Chatham Sound, British Columbia, Canada: Results from numerical modelling studies using historical datasets. *Atmos. Ocean* **2018**, *56*, 129–151. [CrossRef]
15. Drozdowski, A.; Jiang, D. Modelling internal tides in the Strait of Canso. *Atmos. Ocean* **2020**. [CrossRef]
16. Wu, Y.; Hannah, C.; O'Flaherty-Sproul, M.; MacAulay, P.; Shan, S. A modeling study on tides in the Port of Vancouver. *Anthropoc. Coasts* **2019**, *2*, 101–125. [CrossRef]

17. Church, I.W. Multibeam sonar ray-tracing uncertainty evaluation from a hydrodynamic model in a highly stratified estuary. *Mar. Geod.* **2020**. [CrossRef]

18. Wu, Y.; Chaffey, J.; Greenberg, D.A.; Colbo, K.; Smith, P.C. Tidally-induced sediment transport patterns in the upper Bay of Fundy: A numerical study. *Cont. Shelf Res.* **2011**, *31*, 2041–2053. [CrossRef]

19. Huang, H.; Chen, C.; Cowles, G.W.; Winant, C.D.; Beardsley, R.C.; Hedstrom, K.S.; Haidvogel, D.B. FVCOM validation experiments: Comparisons with ROMS for three idealized barotropic test problems. *J. Geophys. Res.* **2008**, *113*, C07042. [CrossRef]

20. Shchepetkin, A.F.; McWilliams, J.C. The regional oceanic modeling system (ROMS): A split-explicit, free-surface, topography-following-coordinate oceanic model. *Ocean Model.* **2005**, *9*, 347–404. [CrossRef]

21. Trotta, F.; Fenu, E.; Pinardi, N.; Bruciaferri, D.; Giacomelli, L.; Federico, I.; Gi, C. A structured and unstructured grid relocatable ocean platform for forecasting (SURF). *Deep Sea Res. II Top Stud. Oceanogr.* **2016**, *155*, 54–75. [CrossRef]

22. Umgiesser, G.; Melaku Canu, D.; Cucco, A.; Solidoro, C. A finite element model for the Venice Lagoon. Development, setup, calibration and validation. *J. Mar. Syst.* **2004**, *51*, 123–145. [CrossRef]

23. Biastoch, A.; Sein, D.; Durgadoo, J.V.; Wang, Q.; Danilov, S. Simulating the Agulhas system in global ocean models—Nesting vs. multi-resolution unstructured meshes. *Ocean Model.* **2018**, *212*, 117–131. [CrossRef]

24. Wang, Q.; Danilov, S.; Sidorenko, D.; Timmermann, R.; Wekerle, C.; Wang, X.; Jung, T.; Schröter, J. The Finite Element Sea Ice-Ocean Model (FESOM) v.1.4: Formulation of an ocean general circulation model. *Geosci. Model Dev.* **2014**, *7*, 663–693. [CrossRef]

25. Paquin, J.-P.; Lu, Y.; Taylor, S.; Blanken, H.; Marcotte, G.; Hu, X.; Zhai, L.; Higginson, S.; Nudds, S.; Chanut, J.; et al. High-resolution modelling of a coastal harbour in the presence of strong tides and significant river runoff. *Ocean Dyn.* **2019**, *70*, 365–385. [CrossRef]

26. Dupont, F.; Hannah, C.G.; Greenberg, D. Modelling the sea level of the upper Bay of Fundy. *Atmos. Ocean* **2005**, *43*, 33–47. [CrossRef]

27. Shaw, J.; Amos, C.; Greenberg, D.; O'Reilly, C.; Parrott, D.; Patton, E. Catastrophic tidal expansion in the Bay of Fundy, Canada. *Can. J. Earth Sci.* **2010**, *47*, 1079–1091. [CrossRef]

28. Metcalfe, C.D.; Dadswell, M.J.; Gillis, G.F.; Thomas, M.L.H. *Physical, Chemical, and Biological Parameters of the Saint John River Estuary, New Brunswick, Canada*; Technical Report No. 686; Department of the Environmental Fisheries and Marine Service: Halifax, NS, Canada, 1976.

29. Leys, V. 3D flow and sediment transport modelling at the Reversing Falls—Saint John Harbour, New Brunswick. *IEEE Oceans Conf.* **2007**. [CrossRef]

30. Vouk, I.; Murphy, E.; Church, I.; Pilechi, A.; Cornett, A. Three-dimensional modelling of hydrodynamics and thermosaline circulation in the Saint John River Estuary, Canada. In Proceedings of the 38th IAHR World Congress, Panama City, Panama, 1–6 September 2019. [CrossRef]

31. Neu, H.A.; Hydrographic survey of Saint John Harbour, N.B. *National Research Council of Canada, Division of Mechanical Engineering*; Mechanical Engineering Report No. MH-97; National Research Council of Canada: Vancouver, BC, Canada, 1960. [CrossRef]

32. Delpeche, N. Observations of Advection and Turbulent Interfacial Mixing in the Saint John River Estuary, New Brunswick Canada. Master's Thesis, University of New Brunswick, Fredericton, NB, Canada, 2006.

33. Smartatlantic. Available online: https://www.smartatlantic.ca (accessed on 28 May 2020).

34. Drakkar Group. Eddy-permitting ocean circulation hindcasts of past decades. *CLIVAR Exch.* **2007**, *12*, 8–10.

35. Levier, B.; Tréguier, A.; Madec, G.; Garnier, V. *Free Surface and Variable Volume in the NEMO Code*; MERSEA IP report WP09-CNRS-STR03-1A; MERSEA: Paris, France, 2007.

36. Becker, J.; Sandwell, D.; Smith, W.; Braud, J.; Binder, B.; Depner, J.; Fabre, D.; Factor, J.; Ingalls, S.; Kim, S.-H.; et al. Global bathymetry and elevation data at 30 arc second resolution: SRTM30_PLUS. *Mar. Geod.* **2009**, *32*, 355–371. [CrossRef]

37. Flather, R.A. A tidal model of the northwest European continental shelf. *Mem. de la Soc. R. des Sci. de Liège* **1976**, *10*, 141–164.

38. Engedahl, H. Use of the flow relaxation scheme in a three-dimensional baroclinic ocean model with realistic topography. *Tellus A Dyn. Meteorol. Oceanogr.* **1995**, *47*, 365–382. [CrossRef]

39. Milbrandt, J.A.; Bélair, S.; Faucher, M.; Vallée, M.; Carrera, M.L.; Glazer, A. The pan-Canadian high resolution (2.5 km) deterministic prediction system. *Weather Forecast.* **2016**, *31*, 1791–1816. [CrossRef]

40. Large, W.G.; Yeager, S.G. *Diurnal to Decadal Global Forcing for Ocean and Sea-Ice Models: The Datasets and Flux Climatologies*; NCAR Technical Note TN-460+ STR; NCAR: Boulder, CO, USA, 2004. [CrossRef]

41. Fairall, C.W.; Bradley, E.F.; Hare, J.E.; Grachev, A.A.; Edson, J.B. Bulk parameterizations of air-sea fluxes: Updates and verification for the COARE Algorithm. *J. Clim.* **2003**, *16*, 571–591. [CrossRef]

42. Smagorinsky, J. General circulation experiments with the primitive equations: I. The basic experiment. *Mon. Weather Rev.* **1963**, *91*, 99–164. [CrossRef]

43. Smagorinsky, J. Some historical remarks on the use of nonlinear viscosities. In *Large Eddy Simulation of Complex Engineering and Geophysical Flows*; Cambridge University Press: Cambridge, UK, 1993; Volume 1, pp. 69–106.

44. Chen, C.; Beardsley, R.C.; Cowles, G.; Qi, J.; Lai, Z.; Gao, G.; Stuebe, D.; Liu, H.; Wu, L.; Lin, H.; et al. *An Unstructured Grid, Finite-Volume Community Ocean Model FVCOM User Manual, v.3.1.6*, 4th ed.; SMAST/UMASSD-13-0701; Marine Ecosystem Dynamics Modelling Laboratory, School for Marine Science and Technology, University of Massachusetts-Dartmouth: New Bedford, MA, USA, 2013.

45. Umlauf, L.; Burchard, H. A generic length-scale equation for geophysical turbulence models. *J. Mar. Res.* **2003**, *61*, 235–265. [CrossRef]

46. Aldridge, J.N. Hydrodynamic model predictions of tidal asymmetry and observed sediment transport paths in Morecambe Bay. *Estuar. Coast. Shelf Sci.* **1997**, *44*, 39–56. [CrossRef]

47. Pawlowicz, R.; Beardsley, B.; Lentz, S. Classical tidal harmonic analysis including error estimates in MATLAB using T_TIDE. *Comput. Geosci.* **2002**, *28*, 929–937. [CrossRef]

48. Marcotte, G.; Bourgouin, P.; Mercier, G.; Gauthier, J.P.; Pellerin, P.; Smith, G.; Brown, C.W. Canadian oil spill modelling suite: An overview. In Proceedings of the 39th AMOP Technical Seminar on Environmental Contamination and Response, Halifax, NS, Canada, 7–9 June 2016; pp. 1026–1034.

49. Molcard, A.; Poulain, P.M.; Forget, P.; Griffa, A.; Barbin, Y.; Gaggelli, J.; Rixen, M. Comparison between VHF radar observations and data from drifter clusters in the Gulf of La Spezia (Mediterranean Sea). *J. Mar. Syst.* **2009**, *78*, 79–89. [CrossRef]

50. Liu, Y.; Weisberg, R.H. Evaluation of trajectory modelling in different dynamic regions using normalized cumulative Lagrangian separation. *J. Geophys. Res.* **2011**, *116*, C09013. [CrossRef]

51. Jacobs, G.A.; D'Addezio, J.M.; Bartels, B.; Spence, P.L. Constrained scales in ocean forecasting. *J. Adv. Space Res.* **2019**. [CrossRef]